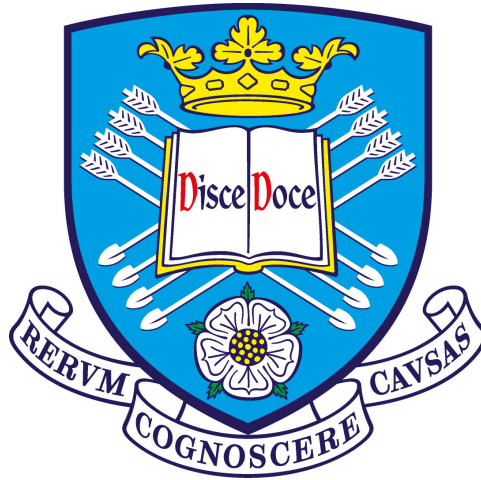


# Using ejected stars to constrain the initial conditions of young star-forming regions

Christina Schoettler

Department of Physics & Astronomy  
The University of Sheffield



*A dissertation submitted in candidature for the degree of  
Doctor of Philosophy at the University of Sheffield*

Submission September 2021



“

Space is big. Really big. You just won't believe how vastly, hugely, mind-bogglingly big it is.

”

— Douglas Adams

# Declaration

I declare that, unless otherwise stated, the work presented in this thesis is my own. No part of this thesis has been accepted or is currently being submitted for any other qualification at the University of Sheffield or elsewhere.

Most of the work presented here has been published and can be found in [Schoettler et al. \(2019\)](#), [Schoettler et al. \(2020\)](#), [Schoettler & Parker \(2021\)](#) and [Schoettler et al. \(2022\)](#). These publications form the basis of Chapters 4, 5, 6 and 7 of this thesis.

The analysis and results presented in Chapter 5 have been updated since they were published in [Schoettler et al. \(2020\)](#), based on an updated method introduced in [Schoettler et al. \(2022\)](#).

# Acknowledgements

This PhD thesis brings a long period of study and research to a close. It all started with the Open University, where I had the chance to study Astronomy while working as a Product Manager. At the beginning, I studied only out of personal interest. I will forever be thankful to a work colleague, who asked if I was sure I was doing it just for “fun” or if I actually wanted to study for a degree. All in all, it took me almost 12 years to reach the point of submitting this PhD thesis and there are many people I would like to thank.

Thanks to Stefan, the work colleague who asked the question. Tricia and Charles, two of my fellow distance-learning students at LJMU who both embarked on their PhDs as well.

I would like to say a huge thanks to Richard Parker, my PhD supervisor. He has always been there to give advice and insight, answer questions, provide feedback and support me in developing into an independent researcher. Thanks to Richard, I also had the chance to spend six amazing months as a visiting researcher at ESTEC (European Space Research and Technology Centre) in the Netherlands.

I would like to thank Jos de Bruijne, who was my supervisor at ESTEC and continues to be a key collaborator for my research. He taught me everything I know about *Gaia*. At ESTEC, I was integrated into the *Gaia* team and would also like to thank everyone there for their support.

Coming back to the people in Sheffield, I would like to thank the whole Astronomy group: staff, postdocs and my fellow PhDs. You are a great and supportive group of people. I always enjoyed the chats and the cake. In particular, I would like to thank two of my collaborators on my first paper, Becky and Liam. Also, my fellow star formation and desk neighbour Emma who started at the same time and was always there for a virtual chat throughout our long working from home period due to the pandemic.

Finally, thanks to my partner István who has been there for me throughout it all. Let's see where Astronomy takes us next!

# Abstract

Star-forming regions have fascinated observers for many centuries. These regions are the locations where most stars form and are considered to be a fundamental unit of star formation. However, what we observe of these regions now is not necessarily how they formed as dynamical evolution can quickly change the initial density, spatial and kinematic substructure. Knowledge about these initial conditions is essential to be able to constrain star formation theories.

Star-forming regions and their likely initial conditions have been studied in the past with different methods, all of which focused on stars that are considered to still be members. In this thesis, I investigate if we can use ejected stars that are often found on the outskirts of young star-forming regions to infer their initial properties. These stars are commonly known as runaway stars, having “run away” from their birth region.

These fast-moving stars have been studied since their discovery in the 1940s using simulations and observations. With improvements in computing capabilities, simulations have made huge advances in the past 50 years, allowing us to make our models of the formation and evolution of these star-forming regions increasingly realistic. For many decades, runaway star observations were limited to bright massive stars, but this changed with the launch of the *Gaia* telescope. This mission has already provided accurate measurements of the positions and velocities for over 1.5 billion stars in our Galaxy. With this data, discoveries of countless new runaways have been made and we now know that they occur across all stellar masses.

In this work, I use a combination of simulations and *Gaia* observations and show that ejected stars can be used to constrain the initial conditions of the ONC and NGC 2264. Both regions appear to have formed from high stellar density, substructured and subvirial initial conditions.

# Contents

<b>Declaration</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Abstract</b>	<b>v</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation for the study . . . . .	2
1.2 Star-forming regions . . . . .	4
1.2.1 Formation mechanisms . . . . .	4
1.2.2 Evolution of young star-forming regions . . . . .	6
1.2.3 Existing methods to constrain initial conditions . . . . .	10
1.2.4 In-situ star formation . . . . .	13
1.3 Star Formation . . . . .	15
1.3.1 The initial mass function . . . . .	15
1.3.2 Models of star formation . . . . .	19
1.3.3 Low-mass stars . . . . .	22
1.3.4 High-mass stars . . . . .	28
1.3.5 Binary systems . . . . .	31
1.4 Protoplanetary discs . . . . .	33
1.4.1 Formation and evolution of protoplanetary discs . . . . .	33
1.4.2 Observations of protoplanetary discs . . . . .	37
1.4.3 Destruction/dispersal of the discs and/or planet formation . . . . .	38
1.5 Runaway stars and their formation mechanisms . . . . .	40
1.5.1 Binary supernova scenario - BSS . . . . .	45
1.5.2 Dynamical ejection scenario - DES . . . . .	53
1.5.3 Can we distinguish between the formation scenarios? . . . . .	58
1.5.4 Hyper-runaway and Hyper-velocity stars . . . . .	60
1.6 Work in this thesis . . . . .	63

<b>2</b>	<b>N-body simulations</b>	<b>64</b>
2.1	<i>N</i> -body interactions in simulations . . . . .	66
2.1.1	Different <i>N</i> -body techniques . . . . .	66
2.1.2	Softening in <i>N</i> -body simulations . . . . .	69
2.2	The Starlab software . . . . .	70
2.2.1	<i>N</i> -body integrator <i>kira</i> . . . . .	70
2.2.2	Stellar and binary evolution package <i>SeBa</i> . . . . .	74
2.2.3	Other <i>N</i> -body software . . . . .	76
2.3	Setting up <i>N</i> -body simulations to search for RW/WW stars . . . . .	77
2.3.1	The box-fractal method . . . . .	77
2.3.2	Initial mass function . . . . .	79
2.3.3	Primordial binaries . . . . .	80
2.3.4	Stellar and binary evolution . . . . .	83
2.3.5	Other parameters defined in the initial conditions . . . . .	83
<b>3</b>	<b>Observations with <i>Gaia</i></b>	<b>84</b>
3.1	Astrometry and the <i>Gaia</i> telescope . . . . .	85
3.2	Observations with <i>Gaia</i> . . . . .	86
3.3	Using <i>Gaia</i> DR2 to find ejected stars . . . . .	88
3.3.1	Filtering data to achieve high quality . . . . .	89
3.3.2	Tracing back stars using their astrometric data . . . . .	91
3.3.3	Constructing CAMDs using photometry . . . . .	94
3.3.4	Age estimates using PARSEC isochrones . . . . .	97
<b>4</b>	<b>Unbound stars in <i>N</i>-body simulations</b>	<b>100</b>
4.1	Introduction . . . . .	101
4.2	Methods . . . . .	101
4.2.1	Initial conditions . . . . .	101
4.2.2	Unbound stars and fractions by mass class . . . . .	102
4.3	Results . . . . .	104
4.3.1	Cumulative 2D-velocity distributions of all stars . . . . .	104
4.3.2	Unbound fractions of stars from initially subvirial and virialized regions . . . . .	108
4.3.3	2D-velocity of unbound stars from initially subvirial and virialized star-forming regions . . . . .	113
4.3.4	Runaway and walkaway stars . . . . .	121
4.4	Discussion . . . . .	123
4.5	Conclusions . . . . .	126
<b>5</b>	<b>Ejected stars from the ONC in <i>Gaia</i> DR2 observations</b>	<b>127</b>
5.1	Introduction . . . . .	128
5.2	ONC specific information used in the analysis . . . . .	128
5.3	<i>N</i> -body simulations of the ONC . . . . .	132
5.3.1	Simulation set-up . . . . .	132



5.3.2	Predictions from the simulations . . . . .	133
5.4	Results from <i>Gaia</i> DR2 . . . . .	136
5.4.1	2D-candidates . . . . .	136
5.4.2	3D-candidates . . . . .	143
5.5	Discussion . . . . .	147
5.6	Conclusions . . . . .	154
<b>6</b>	<b>Protoplanetary systems that may experience more than one dense star-forming environment</b>	<b>156</b>
6.1	Introduction . . . . .	157
6.2	Circumstellar discs in the ONC . . . . .	157
6.3	Method . . . . .	158
6.4	Results . . . . .	159
6.5	Discussion and Conclusion . . . . .	162
<b>7</b>	<b>Ejected stars from NGC 2264 in <i>Gaia</i> DR2 observations</b>	<b>166</b>
7.1	Introduction . . . . .	167
7.2	NGC 2264 specific information used in the analysis . . . . .	167
7.3	Results from <i>Gaia</i> DR2 . . . . .	176
7.3.1	RW/WW stars from S Mon . . . . .	176
7.3.2	RW/WW stars from IRS 1 and IRS 2 . . . . .	179
7.3.3	Confirming previously identified ejected stars . . . . .	182
7.3.4	3D-candidates with a protoplanetary disc . . . . .	182
7.4	RW and WW stars from <i>N</i> -body simulations . . . . .	183
7.4.1	Simulation set-up . . . . .	184
7.4.2	Numbers from the simulations . . . . .	187
7.5	Discussion . . . . .	188
7.5.1	3D-candidates tracing back to more than one subregion . . . . .	188
7.5.2	S Mon . . . . .	189
7.5.3	IRS 1 and IRS 2 . . . . .	192
7.6	Conclusions . . . . .	197
<b>8</b>	<b>Conclusions and future work</b>	<b>199</b>
8.1	Conclusions of the work presented in this thesis . . . . .	200
8.1.1	Do unbound stars in <i>N</i> -body simulations show differences depending on initial conditions? . . . . .	200
8.1.2	Can ejected stars be used to constrain the initial conditions of the ONC? . . . . .	201
8.1.3	Can star-disc systems experience more than one dense environment? . . . . .	201
8.1.4	Can ejected stars be used to constrain the initial conditions of the NGC 2264? . . . . .	202
8.1.5	Final remarks . . . . .	203
8.2	Outlook at potential future work . . . . .	204
	<b>Bibliography</b>	<b>205</b>

<b>A</b>	<b>Additional information for <i>Gaia</i> DR2 identified stars from the ONC</b>	<b>233</b>
A.1	Excluded candidates from ONC analysis . . . . .	234
A.2	2D-candidates with multiple, varying RV measurements for the ONC . . .	235
A.3	Past visitors to the ONC . . . . .	236
A.4	Red giants visiting the ONC . . . . .	238
A.5	ONC high-velocity stars without a circumstellar disc . . . . .	239
<b>B</b>	<b>Additional information for <i>Gaia</i> DR2 identified stars from NGC 2264</b>	<b>241</b>
B.1	RW/WW predictions from simulations for the full mass range . . . . .	242
B.2	2D-RW and WW candidates from S Mon and IRS 1/2 . . . . .	244
B.3	Past visitors to NGC 2264 . . . . .	250
B.4	Future visitors to NGC 2264 . . . . .	252
B.5	Red giants visiting NGC 2264 . . . . .	256
<b>C</b>	<b>List of Publications</b>	<b>258</b>

# List of Figures

1.1	Image of the ONC and Cygnus OB2 . . . . .	2
1.2	Formation model of star-forming regions by hierarchical star formation . . . . .	6
1.3	Early dynamical evolution of regions with different initial conditions . . . . .	9
1.4	Different variations of the IMF . . . . .	16
1.5	Comparison of different forms of the IMF . . . . .	18
1.6	Shu et al. (1987) star formation model . . . . .	19
1.7	Stages of low-mass star formation . . . . .	23
1.8	SEDs of young stellar objects in Class 0+I . . . . .	25
1.9	SEDs of young stellar objects in Class II+III . . . . .	26
1.10	Evolutionary stages of a typical disc . . . . .	34
1.11	Observational tracers for protoplanetary discs . . . . .	36
1.12	Stellar masses and space velocities of RW star candidates . . . . .	43
1.13	Bow shock produced by the massive RW star $\zeta$ Ophiuchi . . . . .	44
1.14	Evolution of a massive binary from an initially close binary . . . . .	46
1.15	Example of the dynamical interactions between a hard binary and a single star . . . . .	54
1.16	Outcomes of binary-single star interactions . . . . .	55
1.17	Outcomes of binary-binary interactions . . . . .	56
2.1	Evolution of particle numbers in $N$ -body simulations . . . . .	65
2.2	Block timesteps in $N$ -body simulations . . . . .	72
2.3	Different initial levels of spatial substructure . . . . .	78
2.4	Flat mass ratio sampling vs IMF sampling . . . . .	82
3.1	Illustration of <i>Gaia</i> scanning procedure . . . . .	87
3.2	Map of the Milky Way using <i>Gaia</i> DR2 fluxes . . . . .	87
3.3	Measuring parallax with <i>Gaia</i> . . . . .	89
3.4	Distribution of the ONC data according to RUWE value . . . . .	90
3.5	CAMD plots showing the ONC data before/after filter and extinction and reddening correction . . . . .	95
3.6	CAMD plots with PARSEC isochrones of different ages . . . . .	98
4.1	Unbound fractions from ten simulations . . . . .	104
4.2	Cumulative 2D-velocity distributions at four different simulation times . . . . .	105

4.3	Long-term evolution of the cumulative 2D-velocity distributions at four different simulation times . . . . .	107
4.4	Unbound fractions by mass class for initially subvirial and virialized star-forming regions . . . . .	109
4.5	Cumulative 2D-velocity distributions for unbound stars with different initial conditions . . . . .	113
4.6	Violin plots showing the 2D-velocity distributions of unbound stars at 10 Myr . . . . .	115
4.7	Violin plots showing 2D-velocity (xy-plane) and 3D-velocity distributions of unbound stars . . . . .	118
4.8	Violin plots showing the 2D-velocity distributions of unbound stars at 10 Myr split by mass class . . . . .	120
4.9	RW stars by mass for four fractal dimensions and three initial virial ratios	121
4.10	WW stars by mass for four fractal dimensions and three initial virial ratios	122
5.1	Masses and velocities of all RW/WW stars from 20 simulations after 4 Myr	133
5.2	CAMD showing traced-back 2D RW candidates from the ONC . . . . .	137
5.3	CAMD showing traced-back 2D WW candidates from the ONC . . . . .	138
5.4	CAMD showing traced-back 3D RW/WW candidates from the ONC . . . . .	144
5.5	Location of the identified 3D RW and WW stars around the ONC . . . . .	152
6.1	Location and motion of the five identified stars with possible discs around the ONC . . . . .	164
7.1	<i>Gaia</i> DR2 right ascension and declination of NGC 2264 members . . . . .	171
7.2	<i>Gaia</i> DR2 right ascension and distance (inverted parallax) of NGC 2264 members . . . . .	172
7.3	CAMDs of all 2D and 3D RW and WW candidates from S Mon . . . . .	174
7.4	CAMDs of all 2D and 3D RW and WW candidates from IRS 1/2 . . . . .	178
7.5	Maximum number of RWs for four selected initial conditions from simulations . . . . .	190
7.6	Location and direction of motion of RW/WW stars from S Mon and IRS 1/2 . . . . .	195

# List of Tables

2.1	Binary fraction in the $N$ -body simulations . . . . .	81
2.2	Mean binary separations in the $N$ -body simulations . . . . .	83
5.1	ONC centre parameters . . . . .	130
5.2	Ejected RW and WW stars from $N$ -body simulations . . . . .	135
5.3	RW star 2D candidates sorted by decreasing 2D-velocity . . . . .	140
5.4	WW star 2D candidates sorted by decreasing 2D-velocity . . . . .	141
6.1	Stars around the ONC possibly with a circumstellar disc . . . . .	160
7.1	NGC 2264 centre parameters . . . . .	169
7.2	S Mon 3D RW and WW stars sorted by decreasing 3D-velocity . . . . .	175
7.3	IRS 1/2 3D RW and WW stars sorted by decreasing 3D-velocity . . . . .	180
7.4	Overview of all initial condition combinations used in the $N$ -body simulations . . . . .	185
7.5	Ejected RW and WW stars from $N$ -body simulations . . . . .	186
A.1	Excluded runaway and walkaway star 2D-candidates around the ONC . . . . .	235
A.2	Slow 2D-candidates with multiple, varying RV measurements . . . . .	236
A.3	Past visitors to the ONC . . . . .	237
A.4	2D and 3D red giant tracebacks to the ONC . . . . .	239
A.5	Stars ejected from the ONC without a circumstellar disc . . . . .	240
B.1	Ejected RW and WW stars from $N$ -body simulations within 100 pc . . . . .	243
B.2	S Mon RW star 2D candidates . . . . .	245
B.3	S Mon WW star 2D candidates . . . . .	246
B.4	IRS 1/2 RW star 2D candidates . . . . .	248
B.5	IRS 1/2 WW star 2D candidates . . . . .	249
B.6	Past visitors to NGC 2264 . . . . .	251
B.7	Future 3D-RW visitors to NGC 2264 . . . . .	253
B.8	Future WW visitors to NGC 2264 . . . . .	254
B.9	Red giants visiting NGC 2264 . . . . .	257

# Chapter 1

## Introduction



Figure 1.1: Image of the ONC as an example of a bound cluster (left) and Cygnus OB2 as an example of an unbound association (right). Image credit: ONC - ESO/M.McCaughrean et al. (AIP), Cygnus OB2 - X-ray: NASA/CXC/SAO/J.Drake et al, Optical: Univ. of Hertfordshire/INT/IPHAS, Infrared: NASA/JPL-Caltech.

## 1.1 Motivation for the study

Star clusters have been a source of fascination for observers of the night sky throughout human history and even found themselves embedded into mythology (e.g. Orion and the Pleiades, [Rappenglück 2001](#); [Hard & Rose 2004](#)). With the advancements in observational and theoretical/numerical techniques, methods and instruments, our understanding of the importance of these regions in star formation has increased immensely but many open questions remain.

Stars often form in grouped or clustered environments where we observe higher stellar densities than in the Galactic field. Stellar densities are  $\sim 5000$  stars  $\text{pc}^{-3}$  in a bound star cluster like the Orion Nebula Cluster (ONC) and  $\sim 6$  stars  $\text{pc}^{-3}$  in an unbound association like Taurus-Auriga compared to the field stellar density (in the solar neighbourhood) of only  $\sim 0.1$  stars  $\text{pc}^{-3}$  ([Lada & Lada 2003](#); [Korchagin et al. 2003](#); [Bressert et al. 2010](#); [King et al. 2012](#)). Depending on the initial conditions in these grouped environments, these star-forming regions either evolve as bound star clusters or unbound associations. Figure 1.1 shows images of an example of each of these two star-forming environments, with the ONC as an example of a bound cluster and Cygnus OB2 as an example of an unbound association.

To allow us to better constrain our theories of how these regions form and in turn

constrain star formation theories, we need to improve our knowledge about their initial conditions. The initial properties of a star-forming region play an important role in their evolution. We can observe regions at different stages of their evolution in the night sky and infer from these observations that star-forming regions are unlikely to have formed from the same initial conditions. However, to constrain the initial conditions exactly is a difficult task as the regions can evolve rapidly over a short period of time and erase much of their initial properties.

The initial conditions that we are mainly interested in are related to the initial stellar density, the initial spatial distribution of stars and the initial kinematic structure, e.g. the velocity distribution. The initial properties of the star-forming regions do not only affect the star formation but also have implications for planet formation (e.g. [Bonnell et al. 2001a](#); [Adams et al. 2006](#); [Parker & Quanz 2012](#); [Vincke & Pfalzner 2016](#); [Nicholson et al. 2019](#); [Concha-Ramírez et al. 2021](#)).

Different methods have already been employed to constrain initial conditions and I describe several of these methods in Section 1.2.3. All of these methods have one thing in common, which is that they are focused on stars that are considered members of the star-forming regions. However, the evolution of these regions can lead to stars being ejected and the work in this thesis is centred on investigating, if we can use ejected stars to constrain the initial conditions of their birth regions. This work follows on from initial work done by [Allison \(2012\)](#) and is the first that uses *Gaia* observations in combination with results from  $N$ -body simulations to constrain these conditions based on ejected stars. The analysis and results of applying this method to two young star-forming regions (ONC and NGC 2264) are presented in Chapters 5 and 7.



## 1.2 Star-forming regions

The term star-forming region encompasses both clusters and associations. In clusters, the member stars are gravitationally bound, meaning that the sum of their kinetic and potential energies is less than zero. Bound clusters are usually more compact with smaller radii and higher densities, whereas associations often have smaller densities and larger radii making the stars in them generally unbound from each other (e.g. [Portegies Zwart et al. 2010](#)). Young stars that are not members of bound clusters are usually observed in unbound associations before they disperse into the Galactic field ([Blaauw 1964](#)). However, there are also observations of young, isolated stars, which appear to have formed outside of star-forming regions altogether (e.g. [de Wit et al. 2005](#); [Parker & Goodwin 2007](#); [Zinnecker & Yorke 2007](#); [Rate et al. 2020](#)).

In the following sections, I start by describing the two main theories used to explain the formation of star-forming regions. I continue with a brief section about their dynamical evolution, followed by a description of existing methods to constrain the initial spatial and kinematic structure as well as the initial density. Finally, I will give a brief overview related to in-situ star formation of apparently isolated stars.

### 1.2.1 Formation mechanisms

The main theories for the formation of star-forming regions can be divided into two simplified categories: the “top-down” and the “bottom-up” scenario. Regardless of the scenario, both start with the collapse of a large cloud of molecular gas, called a giant molecular cloud (GMC). Details about these clouds, such as bulk properties, how they can be observed and what drives the initial collapse are provided in Section 1.3.2.

The classic “top-down” scenario is known under many different names, such as the clustered, episodic, monolithic, starburst like, prompt assembly, gravoturbulent or in-situ scenario (e.g. [Lada et al. 1991](#); [Kroupa et al. 2001](#); [Kroupa 2011](#); [Longmore et al. 2014](#); [Banerjee & Kroupa 2015](#); [Krause et al. 2020](#)). It is based around the idea that all gas that is needed to form a star-forming region is present at the time of formation. This GMC then collapses and fragments in a single star formation episode to form stars in a centrally concentrated cluster (e.g. [Kroupa et al. 2001](#); [Banerjee & Kroupa 2015](#)). [Longmore et al. \(2014\)](#) suggested two different modes for this in-situ scenario in the formation of young massive clusters (YMCs). YMCs are much larger than the typical bound clusters and can rival the masses and densities of old, globular clusters. In the fast in-situ scenario, the gas accumulates in a short time with almost instantaneous star

formation once the required high densities are reached (Longmore et al. 2014). Most of the stars in this scenario would be born virtually at the same time in a starburst like event (e.g. NGC 3603, Banerjee & Kroupa 2015). The second mode suggested by Longmore et al. (2014) is the slow in-situ scenario of YMCs in which the gas accumulation proceeds slowly with no sign of star formation during this time. The authors suggested that this type of in-situ star formation occurs close to the Galactic centre, where gas clouds can reach high densities without forming stars.

The “top-down” scenario implies that most star formation occurs in bound clusters, which then expand and potentially become unbound as a result of disruption during their dynamical evolution, e.g. due to the first supernova (SN). Associations would then be the result of an expanded cluster that has become unbound with the stars in it eventually dispersing into the field (e.g. Kroupa 2011). However, recent results show that some associations were never bound star clusters and that these groups of stars were in fact born unbound (e.g. Wright et al. 2014; Wright 2020). This finding thus requires an alternative explanation for how these often-observed groups of stars could have formed, as it cannot be explained via the “top-down” scenario.

Another indicator for the classical scenario not being the only or even the dominant formation mechanism are observations of star formation having been initiated in denser regions that do not resemble a centrally concentrated cluster (e.g. Krause et al. 2020). Könyves et al. (2015) found a large percentage of stars in the process of formation in over-dense regions of molecular clouds that are not part of an already formed or collapsing cluster.

Both of these observations therefore suggest the existence of an alternative scenario for the formation of star-forming regions. This second “bottom-up” formation scenario is also known by several names describing similar ideas, such as the conveyor-belt, (global) hierarchical collapse or hierarchical assembly scenario (e.g. Bastian et al. 2007; Kruijssen 2012; Longmore et al. 2014; Grudić et al. 2018; Krause et al. 2020).

In contrast to the first scenario, the gas that is eventually converted into stars is not all present within a given region, but is more extended, e.g. in filaments and clumps (e.g. André et al. 2014). There is no single collapse event, but the collapse happens on smaller scales in sub-clouds that will then form smaller sub-clusters. These sub-clusters then merge hierarchically (or via a conveyor belt) into a bound cluster or an unbound association depending on the density of the sub-clusters (e.g. Bastian et al. 2007; Longmore et al. 2014; Grudić et al. 2018). Fig 1.2 illustrates this process. Grudić et al. (2018) suggested that the process starts with a collapse of the initial GMC leading to

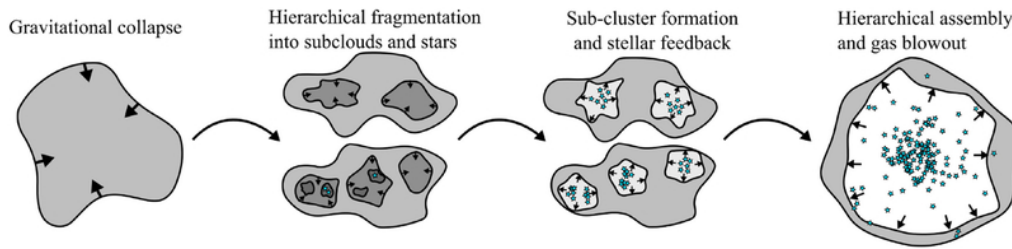


Figure 1.2: The hierarchical formation model for star-forming regions is illustrated in this figure. The first step requires a GMC to undergo gravitational collapse. Instead of forming a centrally bound cluster as in the classical “top-down” model, the cloud undergoes hierarchical fragmentation, which results in different levels of sub-clouds that will then form stars themselves, possibly on different density scales. Sub-clusters form from stars in the same sub-cloud, which can then assemble into a bound or unbound star-forming region. Figure credit: [Grudić et al. \(2018\)](#).

hierarchical fragmentation into sub-clouds in which stars then form at different densities. These regions can assemble, merge or converge into the final observed star-forming region. Unlike the first scenario, this scenario can explain observations of associations that were always unbound, if the initial densities in the sub-clusters were not big enough to form a bound cluster (e.g. [Wright 2020](#)).

### 1.2.2 Evolution of young star-forming regions

The evolution of a young star-forming region can be split into three general phases. In the first few million years (Myr), the regions evolve deeply embedded in their natal molecular cloud with on-going star formation and a significant amount of surrounding gas. These regions are commonly known as embedded clusters in this phase. The second phase of the evolution is a largely gas-free period, where stellar mass loss due to stellar evolution (in particular of more massive stars) plays an important role (e.g. [Portegies Zwart et al. 2010](#)). The lack of gas is a result of two processes. One is simple gas exhaustion, which will conclude the star formation in these regions. The other is gas expulsion due to stellar feedback, e.g. stellar winds or the first SN and this can even lead to a region’s rapid dissolution (e.g. [Tutukov 1978](#); [Lada et al. 1984](#); [Goodwin 1997](#); [Goodwin & Bastian 2006](#); [Longmore et al. 2014](#); [Shukirgaliyev et al. 2018](#)).

Most star-forming regions do not survive this second evolutionary stage, resulting in only 10% of stars being observed in bound clusters past an age of 10 Myr (e.g. [Lada & Lada 2003](#)). For any star-forming regions (now called an open cluster) surviving this second phase, the last stage of the evolution is dominated purely by stellar dynamical processes of the remaining lower-mass stars and can last several 100s Myr (e.g. [Portegies](#)

Zwart et al. 2010). Examples of such older open clusters are the Hyades and Coma Berenices ( $\sim 600\text{--}800$  Myr; Tang et al. 2018; Lodieu et al. 2019). There are a few open clusters that are even older than these, such as Messier 67 ( $\sim 3.5\text{--}4.0$  Gyr; Sandquist et al. 2021) or Berkeley 17, which is thought to be one of the oldest open clusters in the Milky Way ( $\sim 9\text{--}10$  Gyr; Bragaglia & Tosi 2006).

Unbound associations can generally be split into two different types, OB and T associations, depending on their majority stellar components. OB associations were historically identified by their massive stars, but also contain lower-mass stars. Their IMFs are the same as those we find in clusters (e.g. Bastian et al. 2010). T associations are mainly made up of lower-mass pre-main sequence (pre-MS) stars, e.g. T Tauri stars (e.g. Wright 2020). I will describe the formation of T Tauri stars and OB stars in Section 1.3.3 and 1.3.4. Due to their already unbound nature, many associations will not survive for much longer than  $\sim 10$  Myr (similar to the upper age for most bound clusters) and are torn apart by the tidal forces of their host galaxies and then disperse into the field (e.g. Kruijssen & Bastian 2016; Wright 2020). However, there are also older associations, but none of these are as old as the oldest open clusters. One example is the newly identified  $\mu$  Tau Association with a suggested age of  $\sim 60$  Myr (Gagné et al. 2020). A slightly older association was found in Volans-Carina and is suggested to have an age of  $\sim 90$  Myr (Gagné et al. 2018).

From theoretical models and observations, it is suggested that star-forming regions form with significant substructure (“clumpy”) from subvirial (“cool”) initial conditions in molecular clouds with turbulence in these clouds causing filaments and clumps (e.g. Carpenter & Hodapp 2008; Allison et al. 2010). However, even very young clusters like the ONC can present as smooth clusters with a central concentration of massive stars (e.g. Hillenbrand & Hartmann 1998). As a consequence, it must be possible for the substructure to be lost on short timescales, which is more likely to happen if the regions are subvirial from formation (e.g. Goodwin & Whitworth 2004; Allison et al. 2010; Jaehnig et al. 2015; Park et al. 2018).

The virial ratio  $\alpha_{\text{vir}}$  describes the ratio of total kinetic energy  $T$  of all stars to the modulus of the total potential energy  $\Omega$  of all stars,  $\alpha_{\text{vir}} = T/|\Omega|$ . A star-forming region in virial equilibrium has a virial ratio  $\alpha_{\text{vir}} = 0.5$ , with subvirial regions at values below and supervirial ones above.

Subvirial star-forming regions start their evolution by collapsing rapidly and undergo violent relaxation. Violent relaxation was termed by Lynden-Bell (1967), who described how dynamical systems rapidly evolve that formed in a state that is far from equilibrium.

The system as a whole has a global gravitational potential, which is out of equilibrium at formation. In the example of a star-forming region, this non-equilibrium state can occur when the region is initially substructured and/or subvirial. To reach equilibrium, these systems collapse, which leads to a change in the gravitational potential, which in turn affects the orbits (and velocity) of the individual stars. This change in orbit can be quite dramatic, e.g. leading to ejection of a star from the region. The change in orbit will then lead to another change in the overall gravitational potential. Violent relaxation is essentially the change of the orbit of the stars caused by the change in the gravitational potential (Ridpath 2018).

Following the violent relaxation, which can remove most if not all substructure, the regions then start expanding to establish virial equilibrium or even become supervirial (e.g. Goodwin & Whitworth 2004; Allison et al. 2010; Clarke 2015). In addition to the rapid erasure of initial substructure, the dynamical evolution can also cause a reduction in stellar density (e.g. Marks & Kroupa 2012; Parker 2014) and the destruction of primordial binaries/multiples (e.g. Kroupa 1995; Parker & Goodwin 2012; Marks & Kroupa 2012; Duchêne et al. 2018). It can even affect young planetary systems and protoplanetary discs around the young stars in these regions, as they can be disrupted during a dense phase (e.g. Bonnell et al. 2001a; Adams et al. 2006; Parker & Quanz 2012; Vincke & Pfalzner 2016; Nicholson et al. 2019).

Simulations of a star-forming region's early evolution show that regions with subvirial, sub-structured initial conditions mass segregate within a short period of time and can even eject some of their massive stars (e.g. Allison et al. 2010; de Mink et al. 2014). Mass segregation in a clustered environment describes the tendency of massive stars to be located towards or at the centre of the region, whereas low-mass stars are distributed all across the region (e.g. Clarke 2015). This segregation can either occur in parallel with massive star formation (primordial) or after massive star formation (dynamical) (e.g. McMillan et al. 2007; Allison et al. 2009; Moeckel & Bonnell 2009a; Moeckel & Bonnell 2009b; Parker et al. 2014).

Mass segregation occurs dynamically when massive stars form in other parts of the region, but then migrate to the centre due to interaction with other members. It has been shown in simulations that trapezium-like systems (like in the ONC) can form as a consequence of dynamical mass segregation (e.g. Allison & Goodwin 2011). Primordial mass segregation is the outcome of one of the star formation models for high-mass stars (competitive accretion model) that I discuss in section 1.3.4. This model suggests that protostars in the central dense regions can accrete more material than those located in

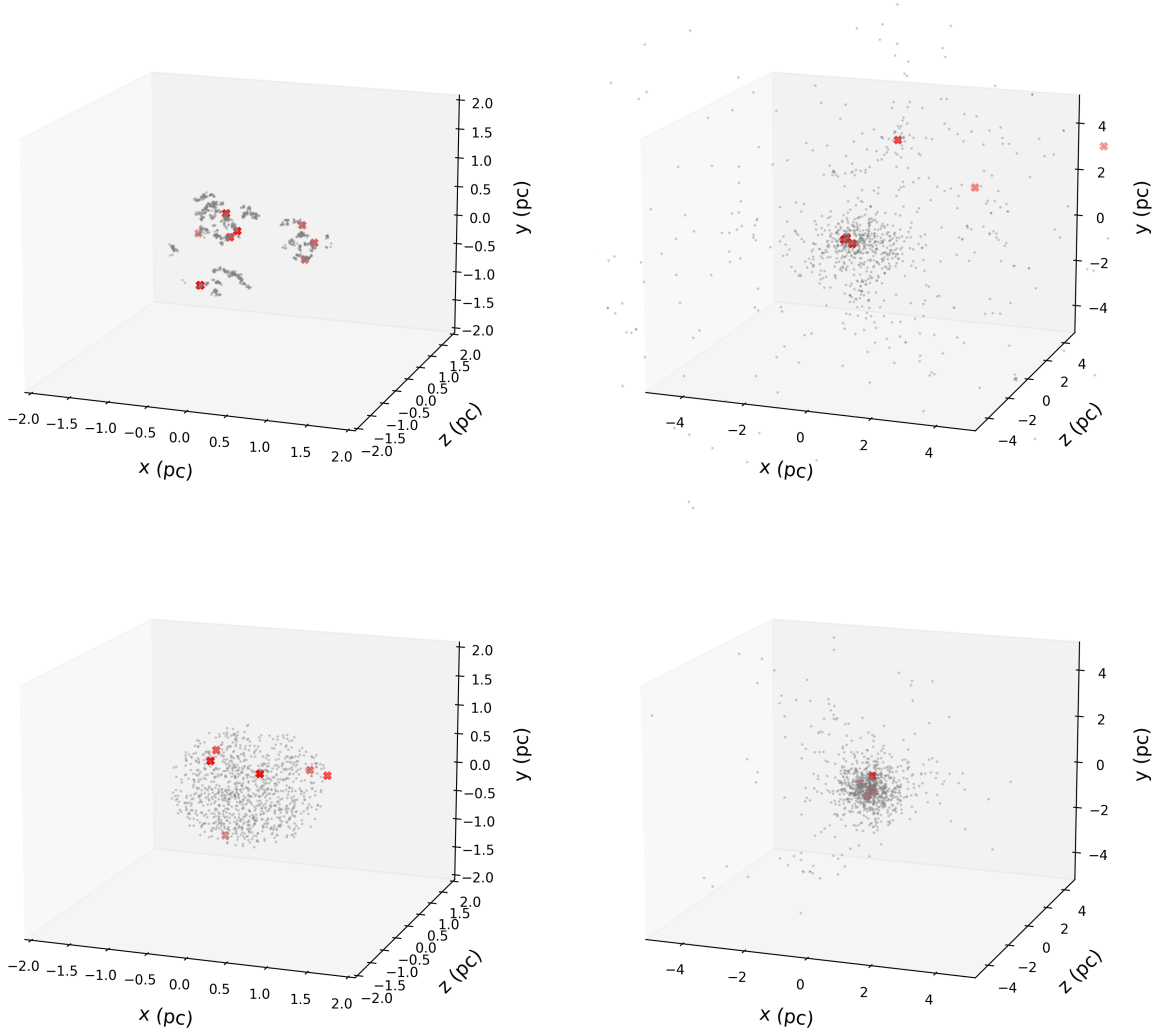


Figure 1.3:  $N$ -body simulations of the early dynamical evolution of two star-forming regions (1000 single stars sampled from a [Maschberger \(2013\)](#) IMF - see 2.3.2) with different initial conditions. Top left: fractal dimension of  $D = 1.6$ , i.e. a high amount of initial spatial substructure (“clumpy”) and an initial virial ratio  $\alpha_{\text{vir}} = 0.3$  at 0 Myr. Top right: The same initial conditions as top left at 5 Myr. Top row: This initially clumpy, subvirial simulation starts with no primordial mass segregation. After 5 Myr, the region has collapsed with the most massive stars (red “x”) segregating to the centre (some of them even get ejected due to interactions). Following the violent relaxation ([Lynden-Bell 1967](#)), the region then expands to establish virial equilibrium or become supervirial (e.g. [Goodwin & Whitworth 2004](#); [Allison et al. 2010](#); [Clarke 2015](#)). Bottom left: fractal dimension of  $D = 3.0$ , i.e. a low amount of initial spatial substructure (“smooth”) and an initial virial ratio  $\alpha_{\text{vir}} = 0.5$  at 0 Myr. Bottom right: The same initial conditions as bottom left at 5 Myr. Bottom row: The initially smooth, virialized simulation also starts without primordial mass segregation. After 5 Myr, most of the stars still occupy a similar volume that they did at 0 Myr. The most massive stars (red “x”) have changed positions slightly but are not more centrally concentrated than at 0 Myr. None of the massive stars got ejected during the dynamical evolution of this region.

the outskirts. This leads to mass segregation during the formation of massive stars in the centre, such as is postulated to have occurred in the young cluster Berkeley 59 (e.g. [Bonnell et al. 1997](#); [Bonnell et al. 2001b](#); [Panwar et al. 2018](#)). In initially smooth, relaxed star-forming regions there is very little dynamical mass segregation. Most of the more massive stars located in the centre have likely formed there and are therefore the result of primordial mass segregation (e.g. [de Grijs 2010](#)).

Figure 1.3 illustrates the early dynamical evolution of two star-forming regions with different initial conditions in  $N$ -body simulations at 0 Myr (panels in left column) and at 5 Myr (panels in right column). Both regions have no primordial mass segregation and an initial radius of 1 pc. In the top row, I show an  $N$ -body simulation with an initial fractal dimension of  $D = 1.6$ , i.e. a high amount of initial spatial substructure (“clumpy”) and an initial virial ratio  $\alpha_{\text{vir}} = 0.3$ . This region undergoes “cool” collapse during which the most massive stars (red “x”) segregate to the centre. Some of these massive stars even get ejected due to interactions in the centre. In the bottom row, I show an  $N$ -body simulation with an initial fractal dimension of  $D = 3.0$ , i.e. a low amount of initial spatial substructure (“smooth”) and an initial virial ratio  $\alpha_{\text{vir}} = 0.5$ . The region in the bottom row undergoes little dynamical evolution as most of the stars still occupy a similar volume at 5 Myr (bottom right) than at 0 Myr (bottom left). The most massive stars (red “x”) have changed positions slightly and are not more centrally concentrated at 5 Myr than at 0 Myr (e.g. [Parker & Dale 2017](#)). None of the massive stars got ejected during this dynamical evolution.

### 1.2.3 Existing methods to constrain initial conditions

Constraining the initial conditions of star-forming regions (i.e. initial density, spatial and kinematic structure) is important for testing the two formation scenarios for star-forming regions that I discussed earlier in this introduction chapter. There are a number of methods to constrain different aspects of these initial conditions.

The two-point correlation function (TPCF) is a method to detect spatial structure. It is a probability function depending on the distance between two objects, e.g. galaxies or stars, and it gives the excess probability of these two objects to be at a certain distance compared to a reference distribution. It has been used to investigate the spatial distribution of stars, e.g. in Taurus-Auriga ([Gomez et al. 1993](#); [Larson 1995](#)) and NGC 346 ([Gouliermis et al. 2014](#)). [Bastian et al. \(2009\)](#) used it to determine the timescale involved in the destruction of spatial structure in the Large Magellanic Cloud (LMC) in combination with another method, the  $Q$ -parameter.

The  $Q$ -parameter is used as an approach to quantify the amount of spatial structure present in a young star-forming region (e.g. Cartwright & Whitworth 2004, 2009). The method is based on using a minimum spanning tree (MST) approach, which uses straight lines to connect all points in a system and then minimises the total length of this tree avoiding closed loops (see description in Gower & Ross 1969). The  $Q$ -parameter is defined by

$$Q = \frac{\bar{m}}{\bar{s}}. \quad (1.1)$$

In this fraction,  $\bar{m}$  represents the normalised mean edge length of the MST of all stars in the region,  $\bar{s}$  represents the normalised correlation length measured by the average separation between the stars divided by the total radius of the region.  $\bar{m}$  depends on the number of stars in the system. The more stars there are in a defined region, the shorter their connection lines will be, therefore decreasing  $\bar{m}$ . The important value for the  $Q$ -parameter is 0.8, with spatially substructured regions showing a value below and smooth, centrally concentrated regions a value above it (e.g. Cartwright & Whitworth 2004, 2009). It has been applied to simulations to model the dynamical evolution of star-forming regions (e.g. Parker et al. 2014; Parker & Wright 2016) and also to observations to detect substructure in many tens of star-forming regions (e.g. Kumar & Schmeja 2007; Wright et al. 2014; Hetem & Gregorio-Hetem 2019).

Other methods are based around measuring the density, such as density maps that split a region in space into different bins and can identify more/less dense regions (e.g. Schmeja 2011). Nearest neighbour methods use distances between stars in star-forming regions (e.g. Gutermuth et al. 2009) or are based on local density (e.g. Schmeja 2011). The INDICATE (INdex to Define Inherent Clustering And TEendencies) method is a recent nearest neighbour approach, in which regions with more neighbours in a defined radius are given a higher index as those with fewer neighbours (Buckner et al. 2019).

The above describes different approaches to determine the spatial structure of a star-forming region. To investigate the kinematic structure in a young star-forming region, I will describe two approaches next. The first distinguishes between a bound (cluster) or unbound (association) state by using the effect of applying different initial virial ratios. Parker et al. (2014) combined the  $Q$ -parameter with a measure of mass segregation ( $\Lambda_{\text{MSR}}$ , Allison et al. 2009) and the local surface density ratio ( $\Sigma_{\text{LDR}}$ , Maschberger & Clarke 2011) to show that regions evolve differently in plots of  $Q$  against  $\Sigma$ . As a consequence, different initial conditions depending on the initial virial ratio (kinematic substructure) can be discerned.

A second approach that is often used involves the velocity. The three-dimensional



velocity of stellar objects, such as stars in young star-forming regions, can be used to describe their kinematics and can be divided into two components: the radial or line-of-sight velocity and proper motion or transverse velocity. Radial velocity (RV) is the one-dimensional velocity component of a stellar object moving towards or away from Earth in the line-of-sight. Transverse (or tangential) velocity describes the two-dimensional velocity component perpendicular to the radial velocity. Transverse velocity is calculated using the proper motion and the distance to the star or object in question. Observationally, the velocity components are often obtained using two different methods. Radial velocity is measured via spectroscopy using the Doppler effect, whereas proper motion is measured via astrometry using the change of position of an object over time (e.g. [Ridpath 2018](#)).

Velocities measured by one or more of the above methods allow an estimate of the velocity dispersion in a star-forming region and can inform about whether the region is expanding or contracting (e.g. [Foster et al. 2015](#)). The velocity dispersion is the spread of velocities in a region, such as a star cluster or a galaxy (e.g. [Ridpath 2018](#)). The virial mass equation (Eq. 7, [Parker & Wright 2016](#)) can be used to calculate the theoretical velocity dispersion for a region in virial equilibrium ( $\sigma_{\text{vir}}$ ) and then be compared to the observed dispersion ( $\sigma$ ):

$$\sigma_{\text{vir}} = \sqrt{\frac{2GM}{\eta R}} \quad (1.2)$$

In order to calculate the virial dispersion, assumptions about the radius  $R$ , mass  $M$  and structure parameter  $\eta$  are required. The resulting values can be used to derive the virial state of a star forming region - subvirial for  $\sigma < \sigma_{\text{vir}}$  or supervirial for  $\sigma > \sigma_{\text{vir}}$ , with subvirial/virialized states corresponding to bound clusters and supervirial states to unbound associations (e.g. [Parker & Wright 2016](#)). However, [Parker & Wright \(2016\)](#) did not consider this to be sufficient in all cases to inform about the current state as virialized regions can appear supervirial if they have previously undergone violent relaxation. Information about the structure, such as the interquartile range (IQR) dispersion should be used in addition to the virial state. To derive a bound/unbound state of a region, [Da Rio et al. \(2017\)](#) used a bound velocity dispersion  $\sigma_{\text{bound}} = \sqrt{2}\sigma_{\text{vir}}$  as an additional value to the virial state, which can show that a region can be supervirial, but still bound.

### Examples of star-forming regions and their likely initial conditions

The ONC in the Orion A molecular cloud is a well-studied young star cluster. It is one of the closest bound clusters containing over 3000 stars (e.g. [Hillenbrand 1997](#); [Menten](#)

et al. 2007). King et al. (2012) used  $N$ -body simulations to show that it must have been initially substructured to explain observations of its binary population. The work in Allison et al. (2010) and Allison & Goodwin (2011) showed that this cluster must have been both spatially and kinematically substructured at birth using the mass segregation measure. Fűrész et al. (2008) measured the radial velocity dispersion of about 1100 ONC stars with a value of  $3.1 \text{ km s}^{-1}$ . Due to its young age, they concluded that the cluster still shows some traces of the initial conditions and has not completely dynamically relaxed, retaining some of its substructure. Tobin et al. (2009) investigated individual radial velocities and dispersion of gas and stars in the region. They found across most of the ONC, the radial velocity dispersion of gas and stars to be between  $3.76$  and  $7.14 \text{ km s}^{-1}$ , agreeing with previous conclusions that it is not dynamically relaxed.

Cygnus OB2 is one of the most massive OB associations in the Milky Way with a total mass of  $1\text{-}3 \times 10^4 M_{\odot}$ . It has hundreds of high-mass stars with stellar masses up to  $100 M_{\odot}$  and is close enough to Earth to observe velocities of both high- and low-mass stars (e.g. Drew et al. 2008; Wright et al. 2010a, 2016). There is an age spread of  $1\text{-}7$  Myr, indicative of star formation happening over several epochs (e.g. Wright et al. 2010b). Wright et al. (2014) used the  $Q$ -parameter and two measures of the mass segregation to determine that it formed highly substructured with a low density and is unbound, strongly indicative of always having been an association. Wright et al. (2016) found a high 3D-velocity dispersion of  $\sim 17.8 \text{ km s}^{-1}$  and also concluded it to be unbound, as the virial mass is larger than the current stellar mass.

Cygnus OB2 is thought to not have dynamically evolved much since its birth. It likely featured a high amount of substructure at birth with a considerable amount of substructure still present with its subgroups now in approximately virial equilibrium. It is not thought to be the result of an expanded bound cluster, strongly suggesting it formed via the “bottom-up” scenario (e.g. Wright et al. 2014, 2016).

#### 1.2.4 In-situ star formation

While it is assumed that most stars form in grouped or clustered environments, there are observations of massive young, isolated stars. If they were born in a grouped environment, they would still be found surrounded by other young stars of different masses due to their short lifetimes. The simulations in Figure 1.3 illustrate that massive stars can be ejected from their birth environments in only a few Myr, so these apparently isolated stars could be ejected stars. In-situ star formation is commonly used to explain the formation of young stars that cannot be traced back to a star-forming region, i.e.

non-ejected stars. It has been suggested in several studies that  $\sim 5\text{-}10\%$  of O-stars in the Galaxy might have formed in isolation as based on their observable characteristics, no origin cluster or association could be located (e.g. [de Wit et al. 2005](#); [Parker & Goodwin 2007](#); [Zinnecker & Yorke 2007](#)).

However, with the improvements in observations, this percentage will likely shrink in the future. [Vargas-Salazar et al. \(2020\)](#) recently analysed field OB-stars in the Small Magellanic Cloud (SMC) and concluded that these stars are consistent with an almost exclusive origin as stars ejected from a star-forming region. Even the lack of observable signatures pointing to ejections (high velocity, trace-back to a star-forming region) should not necessarily lead to a conclusion of isolated star formation. [Pflamm-Altenburg & Kroupa \(2010\)](#) investigated the effects of a two-step ejection mechanism that can lead to an increase or decrease of the velocities commonly assumed from a single ejection event. In addition to differences in the velocities, the two-step ejection can also lead to reorientation of the direction of motion making a trace-back impossible. This two-step ejection event involves the ejection of a binary due to a dynamical interaction followed by a binary supernova ejection (for a detailed explanation of these two ejection mechanisms, see Section 1.5). This can result in a massive runaway star that cannot be traced back to its birth region. The authors even suggested that this mechanism could account for the number of isolated stars in [de Wit et al. \(2005\)](#), therefore making in-situ star formation potentially unnecessary.

Other effects that can mask an ejection origin is the rejuvenation of the companion in a binary due to mass transfer from the primary before it explodes in a SN. Isolated stars are then considered much younger than they actually are, leading to an underestimation of the distance they could have travelled in their lifetime (e.g. [Gvaramadze et al. 2012](#)). [Lucas et al. \(2018\)](#) suggested an alternative explanation for isolated high-mass stars in the LMC as a result of their analysis of  $N$ -body simulations. Through mergers of a lower-mass cluster with a high-mass cluster the first cluster can be tidally disrupted, and massive stars can be ejected at velocities below  $20 \text{ km s}^{-1}$ . At these low velocities at a location outside their birth clusters, they would not be classified as runaways and would be identified as isolated massive stars instead.

## 1.3 Star Formation

The formation of star-forming regions is directly connected with the formation of the stars within these regions. In this section, I will therefore discuss several aspects related to star formation, such as how stars of different masses form and whether they form alone or with companions (binaries or multiples). I start by describing how the different stellar masses are distributed within a stellar population, such as a star-forming region. I follow this by discussing the models for the formation process of stars at different masses, from very low-mass objects up to very high-mass stars, followed by a brief section on the formation of multiple systems.

### 1.3.1 The initial mass function

The initial mass function (IMF) describes the initial distribution of masses in stellar populations, such as star clusters, the Galactic field and even other galaxies. It is an empirical and very significant relationship in star formation that was first discovered by [Salpeter \(1955\)](#) and is thought to be almost universal in the local universe (e.g. [Kroupa 2001](#); [Chabrier 2003a](#); [Offner et al. 2014](#)). However, this universality might not apply all throughout the cosmos, especially at higher redshifts (e.g. [Bastian et al. 2010](#); [Kroupa & Jeřábková 2019](#); [Riaz et al. 2021](#)).

While the IMF describes a stellar mass distribution, this parameter itself cannot be directly measured and needs to be derived via relationships with other measurable parameters, such as luminosity or absolute magnitude. Another important aspect when deriving the IMF is the fact that the IMF itself can only be measured for very young star-forming regions, which contain most of their initial components before dynamical and stellar evolution remove stars from their populations. Dynamical evolution can lead to ejections and evaporation of the initial stellar population. Internal stellar evolution of the most massive stars will lead to SNe shortly after they were born (the first SNe can occur already after 4-5 Myr), removing these already rarer masses from the population. If one measures the masses of stars in a region that is a few Myr old, then only the Present-Day Mass Function (PDMF) can be measured. The IMF and the PDMF are linked by the star formation rate or the stellar creation function ([Miller & Scalo 1979](#)).

[Salpeter \(1955\)](#) used a data set of field stars in the Galactic disc with absolute magnitudes of -4.5 to 13.5 mag, which corresponds to masses between  $\sim 0.4$ -10  $M_{\odot}$ . This data did not directly yield the now known shape of the “Salpeter” IMF, but the author suggested that a single power-law with a slope of -1.35 in its logarithmic form (or -2.35

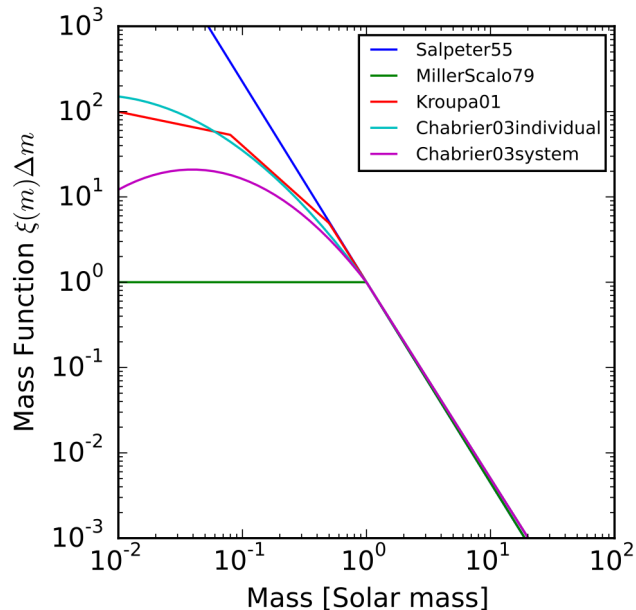


Figure 1.4: Different variations of the IMF (Salpeter 1955; Miller & Scalo 1979; Kroupa 2001; Chabrier 2003a,b). The mass function  $\xi(m)\Delta m$ , which is the number of stars with masses in the range  $m$  to  $m+dm$  within a defined volume of space, is plotted against the stellar mass (in  $M_{\odot}$ ). Figure credit: Johannes Buchner.

in its linear form) would provide an adequate fit to their logarithmic slope. The value of the slope was based on the then available observational data and assumptions about the age of the Galactic disc. With more recent observations and Galactic age assumptions, the slope would have a slightly different value nowadays (-1.05 instead of -1.35, Zinnecker 2011). However, the classical “Salpeter” IMF is still used for stars with masses  $> 1 M_{\odot}$  (e.g. Scalo 2005; Kroupa & Jeřábková 2019).

The negative value of the power-law slope implies that there is a much larger number of lower mass stars than there are high-mass stars in stellar populations. Different studies after the seminal Salpeter (1955) paper showed that the IMF is more complex than a single power law. Figure 1.4 depicts several different formulations of the IMF with the Salpeter (1955) IMF in blue as a straight line with a negative slope. In this figure, on the y-axis is the mass function  $\xi(m)\Delta m$ , which is the number of stars with masses in the range  $m$  to  $m+dm$  within a defined volume of space. The x-axis shows different stellar masses (in  $M_{\odot}$ ).

The first major change to the IMF was introduced by Miller & Scalo (1979). These authors suggested a log-normal form with a flattening of the IMF for stars below a mass of  $1 M_{\odot}$  (green line in Fig. 1.4). Further work was done by Scalo (1986), who

reviewed observational constraints on the IMF, showing that the high-mass end showed no variations within 1 kpc but that the exact slope of the high-mass IMF ( $>15 M_{\odot}$ ) could not be determined due to uncertainties.

The shape of the IMF is now either described following a multi-part, segmented (broken) power law (Kroupa 2001, 2002) or as a smooth log-normal segment combined with a power-law component at higher masses (Chabrier 2003a,b, 2005). The ‘‘Kroupa’’ single-star IMF (Kroupa 2001, 2002) is a 4-part power law with the same power-law index for the upper two mass classes (but differences in their errors), so that it is also referred to as a 3-part power law (red line in Fig. 1.4). The components of this IMF are described by:

$$\xi(m) \propto \begin{cases} m^{-\alpha_0} & \alpha_0 = 0.3 \pm 0.7 & 0.01 \leq m/M_{\odot} < 0.08, \\ m^{-\alpha_1} & \alpha_1 = 1.3 \pm 0.5 & 0.08 \leq m/M_{\odot} < 0.50, \\ m^{-\alpha_2} & \alpha_2 = 2.3 \pm 0.3 & 0.50 \leq m/M_{\odot} < 1.00, \\ m^{-\alpha_3} & \alpha_3 = 2.3 \pm 0.7 & 1.00 \leq m/M_{\odot}. \end{cases} \quad (1.3)$$

$\alpha$  is the power-law index for the slope of the mass function at different masses. This IMF does not have a directly specified upper mass limit in its original source material. However, an upper limit of  $120 M_{\odot}$  was implied in Kroupa (2001, 2002), which was the most massive star known at the time. Unresolved binaries can lead to a bias in the single-star IMF below  $1 M_{\odot}$ . Kroupa (2001) suggested a steepening of the power-law indices  $\alpha_1$  from 1.3 to 1.8 and  $\alpha_2$  from 2.3 to 2.7 to account for these systematic biases in what they called the ‘‘present-day star formation IMF’’.

The second commonly used form is the ‘‘Chabrier’’ IMF (Chabrier 2003a,b, 2005). The author formulated separate IMFs for individual single stars and multiple systems (e.g. binaries), which differ for masses below  $1 M_{\odot}$ , as can be seen in Fig. 1.4. The cyan coloured IMF for single stars has a steeper gradient below a mass of  $1 M_{\odot}$  than the violet-coloured system IMF (Chabrier 2003a,b).

In the work shown in this thesis, I use a Maschberger (2013) IMF, which is a combination of a Chabrier (2003a,b, 2005) log-normal IMF approximation for low-mass stars combined with the power-law slope of Salpeter (1955) for stars above  $1 M_{\odot}$ . It is very similar to the ‘‘Chabrier’’ IMF as can be seen in the comparison of the probability density functions in Fig. 1.5. A more detailed description of the ‘‘Maschberger’’ IMF can be found in Chapter 2.3.2.

As illustrated in Fig. 1.4, there are different formulations of the IMF for stars with

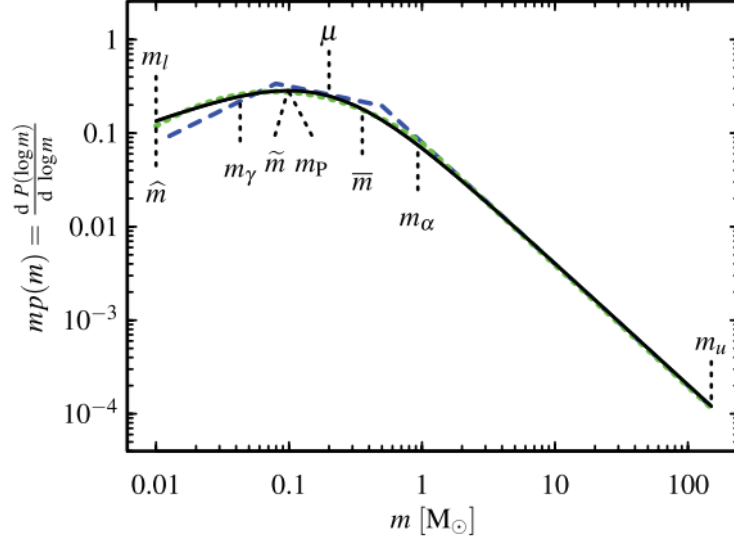


Figure 1.5: Probability density function for the [Maschberger \(2013\)](#) IMF (black solid line), the [Chabrier \(2003a,b\)](#) IMF (dotted green line) and the [Kroupa \(2001, 2002\)](#) IMF (dashed blue line) highlighting the similarity of the “Maschberger” IMF with the “Chabrier” IMF. Figure credit: Figure 6 from [Maschberger \(2013\)](#).

a mass below  $1 M_{\odot}$ , but all of these studies closely agree on the power-law slope for stars above  $1 M_{\odot}$ . While we are able to describe the IMF based on observations and theoretical modelling, the origin of the IMF itself is still poorly understood. Stars form from protostellar cores in clouds of gas and there are different models, either arguing for a correlation of the core mass function (CMF) with the IMF or claiming the independence of CMF and IMF. There appears to be a resemblance of the shape of the CMF with that of the IMF, both in observations as well as in simulations. This led to models claiming that the CMF determines the shape of the IMF, which implies that the mass of a protostellar progenitor core determines the mass of the forming star (e.g. [Hennebelle & Chabrier 2008](#); [Hopkins 2012](#)). This scenario of star formation is also referred to as the core, or monolithic collapse model for massive stars, which assumes that all gas from which an individual star forms is present before star formation commences. Recent work by [Pelkonen et al. \(2021\)](#) showed that there is only a statistical similarity between the CMF and stellar IMF. However, the authors did not find a direct correlation between these two mass functions on a star-by-star basis. Other authors have found that the IMF appears to be independent from the CMF as additional gas is required to produce the observed IMF (e.g. [Bonnell et al. 2001b](#); [Clark et al. 2007](#); [Takemura et al. 2021](#)).

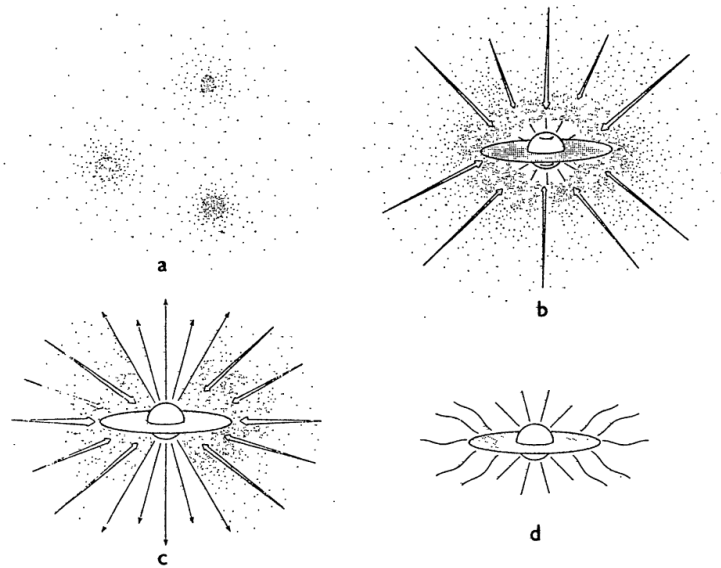


Figure 1.6: The four stages of star formation as described in the [Shu et al. \(1987\)](#) model. In the first stage (a), slowly rotating cores are formed inside molecular clouds that have previously assembled. In the second phase (b), the cloud core eventually becomes unstable and collapses to form an embedded protostar with a disc surrounded by an envelope of dust and gas that is infalling towards the protostar. In the third stage (c), the infall of material onto the protostar slows down and the material preferentially accretes on the disc due to the angular momentum. In the final stage (d), most of the material accretes onto the disc allowing the angle of the stellar wind to widen, eventually revealing the pre-MS star, which will over time lose its circumstellar disc. Figure credit: [Shu et al. \(1987\)](#).

### 1.3.2 Models of star formation

#### [Shu et al. \(1987\)](#) model of star formation

The [Shu et al. \(1987\)](#) model of star formation describes the formation of stars within molecular clouds as a slow, quasi-static process. Figure 1.6 illustrates the four stages of star formation that describe this model. In the first stage (Fig. 1.6 a), slowly rotating cores are formed inside molecular clouds that have previously assembled. This core formation happens once the magnetic field that has prevented the cloud from gravitationally collapsing starts to leak due to ambipolar diffusion. This diffusion occurs when hydrogen molecules decouple from the plasma, allowing the plasma to collapse leading to a density increase in the cloud. In the second stage (Fig. 1.6 b), the cloud core eventually becomes unstable and collapses to form an embedded protostar with a disc surrounded by an envelope of dust and gas that is infalling towards the protostar. The protostar accretes material from this envelope as well as the disc.



In the third stage (Fig. 1.6 c), the infall of material onto the protostar slows down and the material preferentially accretes on the disc due to angular momentum. This slowdown allows the stellar wind to pierce through at the stellar poles creating bipolar outflows and collimated jets. In the final stage (Fig. 1.6 d), most of the material accretes onto the disc allowing the angle of the stellar wind to widen, eventually revealing the pre-MS star, which will over time lose its circumstellar disc. [Shu et al. \(1987\)](#) also suggested a fifth evolutionary phase, which is the disappearance of the disc.

The timescale for star formation via the [Shu et al. \(1987\)](#) model is determined by the magnetic diffusion time, which can be a considerably longer timescale than the lifetime of a GMC itself (GMC lifetime  $\sim 10\text{-}30$  Myr; [Chevance et al. 2020a,b](#); [Benincasa et al. 2020](#)).

### Modern theories of star formation

The [Shu et al. \(1987\)](#) model was considered the standard model of star formation for several years until observations and simulations showed that star formation occurred on a much faster timescale. This faster timescale is around one to two crossing times, compared to tens of crossing times required in the [Shu et al. \(1987\)](#) model (approximate crossing time for a standard GMC is  $\sim 4$  Myr, [Elmegreen 2000](#); [Clark et al. 2005](#)). Relating the crossing time to the free-fall time (minimum time for a GMC to collapse purely under the effect of its self-gravity), it was found that these were possibly correlated and could be of similar value, depending on the gas surface densities in the cloud (e.g. [Chevance et al. 2020a,b](#)). [Chevance et al. \(2020b\)](#) also found that GMC lifetimes are at most a factor 2 of these timescales, similar to results of [Benincasa et al. \(2020\)](#).

In addition to these timescale-related issues, the [Shu et al. \(1987\)](#) model also prohibited the formation of binaries via fragmentation of cores. It relied instead on binaries having formed dynamically from the capture of another unbound protostar, which is considered to be an inefficient formation scenario (e.g. [Tohline 2002](#)).

The slow, quasi-static process of star formation was abandoned in favour of a fast process driven by turbulence. This turbulence could originate from the passage of spiral arms in galaxies, be caused by stellar feedback such as shock fronts from SNe, be due to cloud-cloud collisions or gravitational instabilities in the interstellar medium (ISM) itself. Any of these can cause a compression of the diffuse ISM on a large-scale, creating over-densities. These over-densities can then form GMCs of hydrogen with masses that are around  $\sim 10^5\text{-}10^6 M_{\odot}$  and sizes that spread across regions larger than 30 pc and up to 100-200 pc. The typical temperature of GMCs is around 10-20 K with a mean density of  $\sim 10^2$

particles per  $\text{cm}^3$  (density will be higher in denser regions, such as clumps and cores). GMCs are located in a warmer neutral medium with a temperature of  $\sim 8000$  K and a density of  $\sim 1$  particle per  $\text{cm}^3$ . It is suggested that the trigger mechanisms influence the size of the GMCs they can form, e.g. stellar feedback is expected to create GMCs with smaller masses ( $\sim 10^4 M_\odot$ ), whereas spiral arms passage is connected to those forming with larger masses (e.g. Elmegreen 2007; Chevance et al. 2020a; Ballesteros-Paredes et al. 2020).

The turbulence in the ISM can be passed down to the GMCs, which are supersonically turbulent and possibly Alfvénic. Turbulence is essentially a redistribution of energy and density over different length scales. The energy cascades down from larger to smaller scales and this can happen via shocks and/or eddies (e.g. Burgers 1939; Kolmogorov 1941; Pudritz & Kevlahan 2013; Boneberg et al. 2015; Ballesteros-Paredes et al. 2020).

Molecular hydrogen ( $\text{H}_2$ ) makes up the largest proportion of the gas in GMCs. However, this gas is very difficult to observe as it emits only weakly in the Infrared (IR) and Ultraviolet (UV), due to its lack of dipole moment. Molecular carbon monoxide (CO) is commonly used to trace the  $\text{H}_2$  in GMCs, as well as protoplanetary discs (see Section 1.4.2). However, there are open questions about the exact correlation of CO with  $\text{H}_2$  and when it is observed during the lifetime of a GMC (e.g. Glover & Mac Low 2011; Chevance et al. 2020a). Other methods to observe GMCs, in particular to measure their mass, are thermal dust emission and dust extinction in IR and mm-wavelengths (Lombardi et al. 2014).

Within the GMCs, there are regions of substructure where over-densities in the molecular gas have formed clumps with width of  $\sim 1$  pc and filaments with width of  $\sim 0.1$  pc. Within these denser substructures are the pre-stellar cores, which are made up of cold gas at a temperature of  $\sim 10$  K with sizes  $\lesssim 0.1$  pc. These cores are thought to be the result of further fragmentation of the filaments (e.g. di Francesco et al. 2007; André et al. 2014; Rosen et al. 2020).

The collapse of the GMC itself is governed by a length/mass scale, the Jeans length and Jeans mass. A GMC will only collapse if the gravitational pressure from its mass is no longer balanced by the gas pressure. The Jeans length describes the minimum length a cloud must have to be able to collapse and is mainly influenced by the density of the gas. This length measurement then feeds into the Jeans mass as a radius. The Jeans mass is the minimum mass at which gravity will overcome the pressure of the gas leading to gravitational collapse (Jeans 1902).

Panels (a) and (b) in Figure. 1.7 illustrate these first stages of the star formation

process. The continuing gravitational collapse of the pre-stellar cores leads to an increase in density. Initially the temperature of the collapsing core is about the same as that of the GMC as the heating caused by the collapse is slow enough so that the core temperature does not increase. Excess energy is instead radiated away, making this phase of the collapse isothermal (e.g. [McKee & Ostriker 2007](#), and references therein). Eventually the gas reaches a critical density of  $\rho \sim 10^{-13} \text{ g cm}^{-3}$ , at which point the heating cannot be compensated for anymore by radiative cooling as the core has become opaque. The size of this first “hydrostatic” core is a few astronomical units (au). At this point the collapse halts due to the increase in core temperature. This will lead to dissociation of the molecular hydrogen. Eventually a second contraction/collapse phase is started when the core densities are  $\rho \sim 10^{-8} \text{ g cm}^{-3}$ , creating the second “hydrostatic” core. This core eventually stops collapsing further due to thermal pressure and starts accreting material from its surrounding envelope, thereby entering the protostar phase (e.g. [McKee & Ostriker 2007](#); [Rosen et al. 2020](#), and references therein).

### 1.3.3 Low-mass stars

The separation between low-/intermediate-mass stars and high-mass stars is commonly done based on whether stars undergo core-collapse SNe at the end of their lives. This mass is thought to be  $\sim 8 \pm 1 M_{\odot}$  at Zero-Age Main Sequence (ZAMS, i.e. when the star first reaches the MS) with different factors affecting the minimum mass, e.g. metallicity (e.g. [Smartt 2009](#); [Ibeling & Heger 2013](#)). For simplicity, I will refer to any star that is below this SN mass limit but above the mass limit for a brown dwarf (BD,  $0.08 M_{\odot}$ ) as a low-mass star in this section.

Figure 1.7 illustrates the formation steps of low-mass stars. The first two stages (panel a and b) related to the collapse of the GMC into pre-stellar cores were discussed in the previous section. I start this section at the point where a collapsed core forms a protostar. The evolutionary phases leading up to the main sequence (MS) are classified based on observations made about the slopes of the spectral energy distribution (SED) of young stellar objects (YSOs) within a defined wavelength range in the infrared (IR,  $\sim 2\text{-}25 \mu\text{m}$ ). The SEDs can be used to infer the mass and temperature of the objects ([Lada 1987](#); [McKee & Ostriker 2007](#)). Figures 1.8 and 1.9 show the SED graphs for Classes 0 through III.

In the first classification scheme for YSOs, three classes (I, II, III) were separately identified ([Lada 1987](#)). At this point, it was not yet possible to detect the mm/sub-mm emission from the envelope surrounding the still fully embedded protostar, but this

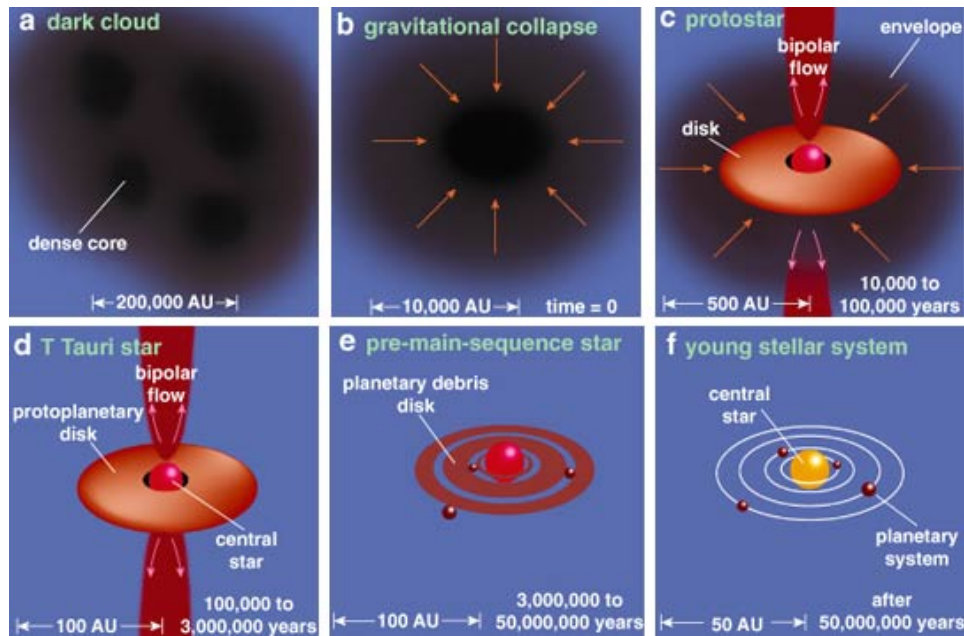


Figure 1.7: The stages of low-mass star formation start with the gravitational collapse of a GMC into pre-stellar cores (panel a and b). In panel c, the protostellar phase begins, which corresponds to Classes 0 and I, where a disc forms around the protostar and both are embedded deeply in the surrounding gas envelope. The protostar accretes material both from the envelope as well as the disc. In panel d (Class II phase), the gaseous envelope dissipates, and the gas-free phase begins, where the pre-MS star accretes from the disc and contracts to reach a temperature to start core hydrogen burning. In panel e, the disc is now a virtually gas-free debris disc, and the central star is now in its Class III phase continuing to contract. The result of the formation process is shown in panel f, where the star has reached the main sequence (MS) and is possibly surrounded by a young planetary system. Figure credit: Tom Greene (2001).

changed in 1993 and subsequently an earlier Class 0 phase was added (Andre et al. 1993).

Classes 0 and I are considered to cover the protostellar phase, whereas Classes II and III represent the pre-MS phase of a YSO. The Class 0 and I phases are illustrated in Figure 1.7 in panel (c). Class 0 is the shortest phase with a duration of only  $1-3 \times 10^4$  years, which increases to  $1-2 \times 10^5$  years for Class I (Andre et al. 2000). More recent estimates from Spitzer observations of the typical duration of these two phases are slightly longer at 0.13–0.26 Myr for Class 0 and 0.27–0.52 Myr for Class I (Dunham et al. 2015).

In the Class 0 phase, the young protostar is still deeply embedded in the envelope, with the mass of the envelope being larger than that of the protostar, which in turn is larger than the mass contained in the forming disc (for details on disc formation, see Section 1.4). During this earliest phase the protostar rapidly accretes mass from the surrounding envelope and a protostellar disc forms around it. The protostar is likely gaining most of its final mass during this phase. There is little emission in the optical and near-IR (nIR) and most of it is black-body-like emission in the sub-mm range, as only the cool gaseous envelope can be observed (e.g. McKee & Ostriker 2007; Williams & Cieza 2011). The top panel in Figure 1.8 illustrates the SED that would likely be observed at this stage of the evolution. The bolometric temperature  $T_{\text{bol}}$  of the protostar in this phase is  $< 70$  K and is used as a way of separating the different classes. The bolometric temperature is defined as being the effective temperature of a black body that has the same mean frequency as is observed in the SED of the YSO (Myers & Ladd 1993; Chen et al. 1995). Already in this phase, bipolar outflows can be observed to emerge from the envelope. The length of these early outflows is thought to be much less than 1 pc. They are of molecular type (e.g. CO, OH) with velocities from 50–150 km s<sup>-1</sup> and drive a mass loss rate of  $10^{-5} M_{\odot} \text{ yr}^{-1}$  (e.g. Bally 2016).

Protostars move from Class 0 to I, when they have accreted about half of the material in the envelope (e.g. Andre et al. 2000). Now the mass of the protostar is larger than the mass of the envelope, which is now similar to the mass in the disc. In this phase, the emission coming from the protostar starts to emerge from the envelope and is dominated by mid-IR to far-IR due the dust in the remaining envelope. An IR-excess is also present and provides evidence of disc emission. The slope of the SED in the mid-IR is increasing with  $\alpha > 0.3$  (e.g. McKee & Ostriker 2007; Williams & Cieza 2011; Hartmann et al. 2016). The bottom panel in Figure 1.8 illustrates the SED that will likely be observed at this stage of the evolution. In this phase, the protostar accretes both from the decreasing envelope as well as the protostellar disc. The bolometric temperature increases from

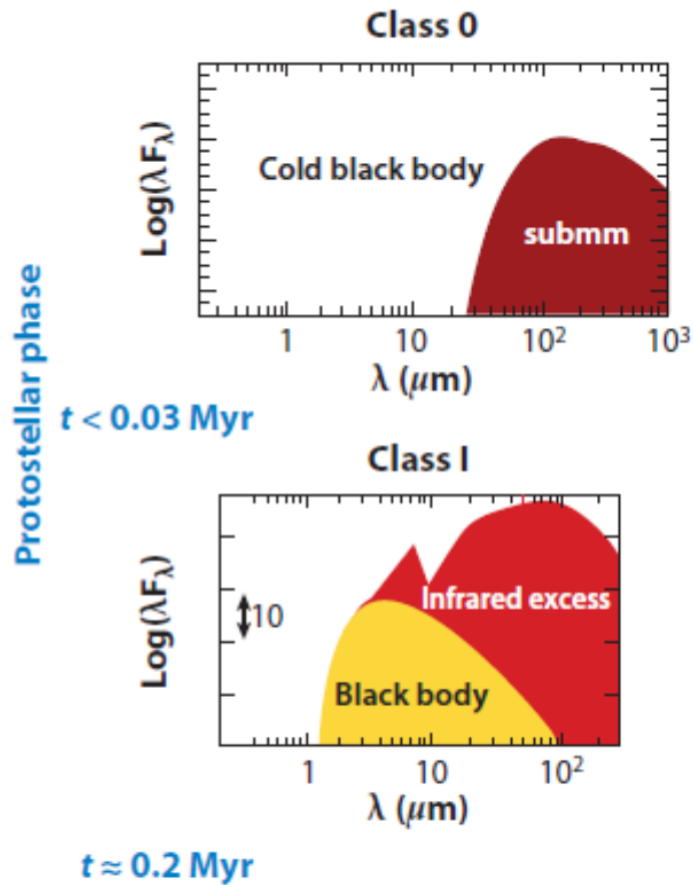


Figure 1.8: SEDs of young stellar objects in Class 0+I. The top panel shows the deeply embedded Class 0 phase, where the observed emission comes from the cold gaseous envelope and is in the sub-mm wavelengths. In Class I (bottom panel), the envelope starts to dissipate and emission from the central protostar becomes visible at IR-wavelengths. The slope of the SED between  $\sim 2\text{-}25 \mu\text{m}$  is increasing. There is IR-excess towards the longer wavelength pointing at disc emission. Figure credit: Swinburne University of Technology (2011).

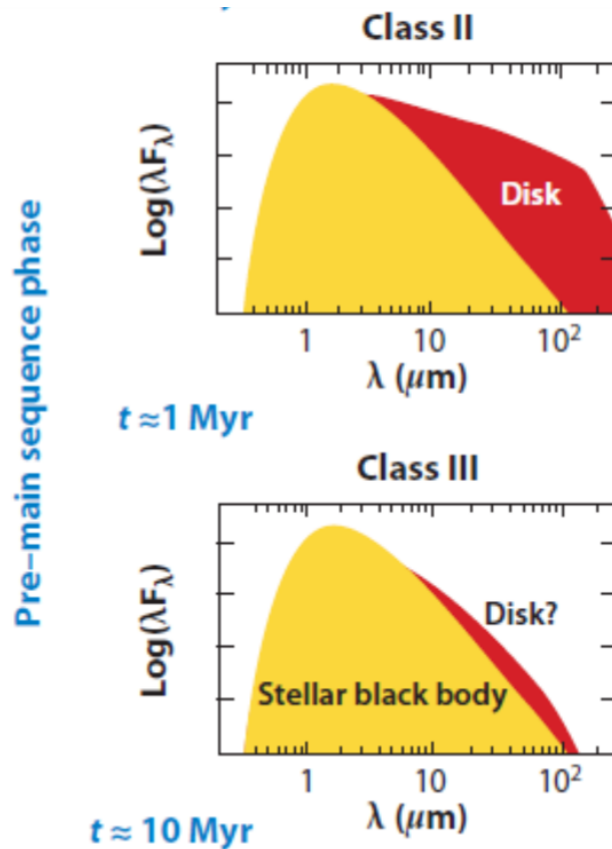


Figure 1.9: SEDs of young stellar objects in Class II+III. The emission of the pre-MS star in the Class II phase (top panel) peaks at shorter wavelength than it does in the Class I phase, pointing at increasing temperature of the YSO. The slope of the SED is now decreasing; however, it is not as steep as in the Class III phase, as there is still strong emission from the disc. In the Class III phase (bottom panel), this disc emission disappears, steepening the slope of the SED. Figure credit: Swinburne University of Technology (2011).

$\sim 70$  K to  $\sim 650$  K (Chen et al. 1995). The bipolar outflows in this phase show a high-velocity inner jet with a slower outer part with a wide angle (e.g. Bally 2016).

During the previous two phases, the protostar has accreted material from its surroundings, slowly contracting and increasing in core temperature. In the Class II phase, the envelope disappears and the YSO emerges as a pre-MS star (Fig. 1.7 panel d). In this phase, the star is accreting from its gas-rich protoplanetary disc and is also known as a Classical T Tauri star (CTTS) ( $< 2 M_{\odot}$ ) or a Herbig Ae/Be star ( $> 2 M_{\odot}$ ), which both show strong emission lines in their spectra. The top panel in Figure 1.9 illustrates the SED that will likely be observed at this stage of the evolution. The slope of the SED is now decreasing in the IR-wavelengths with  $-1.6 < \alpha < -0.3$ . The emission from

the central star is clearly visible and continues to move to shorter IR-wavelengths corresponding to ever increasing temperatures (e.g. [McKee & Ostriker 2007](#); [Williams & Cieza 2011](#); [Hartmann et al. 2016](#)). The bolometric temperature in this phase increases from  $\sim 650$  K to  $\sim 2800$  K ([Chen et al. 1995](#)). The duration of this phase is once again longer than the previous phases and will last  $\sim 2$  Myr (e.g. [Dunham et al. 2015](#)).

The final phase for the YSOs before they reach the MS is the Class III phase. In this phase, the stars are also called Weak-lined T Tauri stars (WTTSs). They show little or no accretion and have almost gas-free debris discs (Fig. 1.7, panel e). The bolometric temperature has reached a minimum of 2800 K ([Chen et al. 1995](#)). The bottom panel in Figure 1.9 illustrates the SED that will likely be observed at this stage of the evolution. The slope of the SED further steepens to  $\alpha < -1.6$  and the IR-excess from the surrounding disc gets smaller the older the pre-MS star gets (e.g. [McKee & Ostriker 2007](#); [Williams & Cieza 2011](#); [Hartmann et al. 2016](#)). This is the longest phase, and its exact length depends on the mass of the YSO. The higher the mass, the quicker the star will reach the temperature required to initiate core hydrogen burning ( $T \geq 10^6$  K).

Stars in the pre-MS phase (Class II and III) will follow [Hayashi \(1961\)](#) evolutionary tracks to approach the MS. Stars in this phase have gained virtually all of their final mass. However, their radii are still larger than they will be once the stars have reached the MS and so they will be slowly contracting. For stars with solar mass, this contraction phase will last a few  $10^7$  yr, for less massive stars this can take up to  $10^8$  yr. The more massive Herbig Ae/BE stars will be fully radiative around the start of their pre-MS phase and follow a faster route to the MS along the timescale of the Class II phase ( $\sim 10^6$  yr) or even less (e.g. [Waters & Waelkens 1998](#)).

While the above classifications and phases appear to be defined very clearly, aspects such as the geometry or viewing angle can make an unambiguous identification difficult. For example, Class I and II objects can be mistaken for each other depending on the inclination angle. Class 0 objects can be mistaken for Class I objects if the envelope is flattened due to magnetic fields. YSOs classified based on their SEDs or bolometric temperatures can also be mistaken for one another (e.g. [McKee & Ostriker 2007](#); [Williams & Cieza 2011](#); [Dunham et al. 2015](#)).

## Brown dwarfs

BDs are one of the most numerous objects in the sky and are often considered “failed” stars, i.e. stellar objects that never get massive enough to be able to start core hydrogen burning. It is thought that most BDs form via the same initial processes as low-mass stars



described in the previous section, but that the protostar accretion process is interrupted before these low-mass objects accrete sufficient mass to allow them to start core hydrogen fusion. Most of the models of BD formation are therefore centred around causes of the disruption of the accretion process. These disruption theories are also thought to apply to the formation of very-low mass stars with masses below  $\sim 0.2 M_{\odot}$  to explain why they do not gain more mass (e.g. [Luhman 2012](#)).

[Bonnell et al. \(2008\)](#) suggested that the high velocities present in young clusters and tidal shear, due to material moving in opposite directions, disrupt the accretion for many low-mass objects before they reach a mass required for hydrogen burning to start directly. Another mechanism suggested that photo-ionising radiation coming from rapidly formed massive stars (i.e. OB stars) destroys the accretion material surrounding the low-mass objects (e.g. [Hester et al. 1996](#); [Whitworth & Zinnecker 2004](#)). Fragmentation of massive accretion discs due to gravitational instability is suggested to lead to formation of low-mass companion stars (or BDs) around high-mass stars (e.g. [Rice et al. 2003](#); [Stamatellos et al. 2007, 2011a](#)). This disc fragmentation could also cause episodic accretion, limiting the total mass accreted, resulting in lower-mass objects (e.g. [Stamatellos et al. 2011b](#)).

In high-density environments such as young, embedded clusters, young objects interact dynamically with others leading to low-mass protostars being ejected from the massive, gas-rich cores (e.g. [Reipurth & Clarke 2001](#); [Bate et al. 2002, 2003](#)). [Reipurth & Mikkola \(2015\)](#) showed that they can even form from ejections of triple systems. Yet another model suggests that low-mass stars are not formed by interrupting accretion, but that pre-stellar, collapsing cores of different masses are produced as a direct result of gas fragmentation and turbulent compression. The smaller cores then evolve into BDs or very low-mass stars, predetermined in mass (e.g. [Padoan & Nordlund 2002, 2004](#); [Hennebelle & Chabrier 2008](#)).

### 1.3.4 High-mass stars

Unlike lower-mass stars, which approach the MS by continually contracting and increasing their core temperature until they reach the core hydrogen burning stage, high-mass stars follow a much more direct route. High-mass stars rapidly accrete material and will reach a mass where hydrogen burning can commence after only about  $10^5$  yr. They are essentially born on the MS as they emerge from their gaseous envelope while they are already fusing core hydrogen. Before these massive stars become visible, they go through the embedded phase, which lasts about 15% of their total lifetime. While the formation of B-stars ( $\sim 8-16 M_{\odot}$ ) can be considered to be lower-mass star-like, for O-stars (in

particular  $> 20 M_{\odot}$ ) this does not apply (e.g. [Zinnecker & Yorke 2007](#), and references therein).

In the high-mass range, several models are used to describe the formation process: the competitive accretion model (e.g. [Bonnell et al. 1997](#); [Bonnell et al. 2001b](#); [Liu et al. 2013](#); [Isequilla et al. 2021](#)), the core accretion or monolithic collapse model (e.g. [Yorke & Sonnhalter 2002](#); [McKee & Tan 2003](#); [Herpin et al. 2012](#); [Zhang et al. 2019](#)), the collisional merger model (e.g. [Bonnell et al. 1998](#); [Bonnell & Bate 2002](#); [Moeckel & Clarke 2011](#); [Banerjee et al. 2012](#); [Krumholz 2015](#)) and the fragmentation-induced starvation model (e.g. [Peters et al. 2010a,b](#)). These models are not mutually exclusive and all of them might occur in the universe ([Zinnecker & Yorke 2007](#)).

Star formation following the competitive accretion scenario happens in a clustered environment (e.g. [Bonnell et al. 1997](#); [Bonnell et al. 2001b](#)). In this environment, star formation takes place at different locations. The available gas fragments down to cores of about the same mass (i.e. Jeans mass, [Peters et al. 2010b](#)). Competitive accretion states that a star's ability to accrete material depends on the amount of available material in its surrounding accretion regions. The material to form stars is not all present when star formation starts. It continues to flow into the region while star formation is on-going. The more material is available, the more massive a young protostar can become, while competing with other YSOs for material. The region to accrete from is largest in the centre of a star-forming region due to the gravitational potential guiding the flow to this region ([Bonnell et al. 1997](#); [Bonnell et al. 2001b](#)).

[Liu et al. \(2013\)](#) observed the high-mass star-forming clump G10.6-0.4 and suggested that the emission and velocities are consistent with competitive accretion. [Isequilla et al. \(2021\)](#) observed another dense clump G34-MM1 located in a filamentary IR dark cloud. The authors found fragmentation on different scales and associated outflows, which they suggested to support the competitive accretion model.

Unlike the competitive accretion scenario, the core accretion or monolithic collapse model does not require massive star formation to occur within a clustered environment but can explain star formation in isolation. The available molecular gas fragments into individual cores of different masses. Each of these molecular cloud cores will collapse to form a single massive object (or a close binary) with an accretion disc around the object providing further material, which is similar to how lower-mass stars form (e.g. [Yorke & Sonnhalter 2002](#); [McKee & Tan 2003](#)).

The monolithic collapse scenario requires all material that will form a star to be present before the collapse starts. To explain massive star formation via the monolithic

collapse route requires huge gas reservoirs to be present, which is unlikely in clustered environments and more suited to isolated formation (e.g. [Zinnecker & Yorke 2007](#)). [Herpin et al. \(2012\)](#) observed a massive protostellar object in the star-forming region W43. They found evidence that the observed dynamics are predicted by the monolithic collapse model, but competitive accretion cannot be ruled out completely. Observations from another massive protostar G339.88-1.26 showed an infalling, rotating gaseous envelope transitioning into a Keplerian disc consistent with the core accretion model ([Zhang et al. 2019](#)).

Collisional mergers of intermediate-mass stars creating more massive stars can occur in extremely high-density environments. It is a complementary formation route to the competitive accretion model as the on-going accretion within a star-forming region can cause a contraction. Increasing the density leads to the possibility of mergers by collision to form more massive stars. Eventually, a star will be massive enough to lead to gas expulsion in the cluster due to its stellar feedback, causing the cluster to expand to a lower density preventing further collisions (e.g. [Bonnell et al. 1998](#); [Bonnell & Bate 2002](#); [Moeckel & Clarke 2011](#); [Banerjee et al. 2012](#); [Krumholz 2015](#)). Instead of a direct collision, the merger scenario can also occur during the hardening of eccentric binaries (e.g. [Bonnell & Bate 2005](#); [Banerjee et al. 2012](#)).

Another approach to massive star formation is the fragmentation-induced starvation model. Like in the competitive accretion model, star formation starts with massive stars forming more likely towards the centre of a star-forming region accreting surrounding material. However, in the starvation scenario, the accretion onto the most massive stars in the cluster centre stops when less massive companions at locations further out form due to fragmentation of the disc around the massive star. These initially lower-mass stars then intercept and accrete the material instead ([Peters et al. 2010a,b](#)). Simulations by [Klassen et al. \(2016\)](#) showed that the fragmentation does not lead to formation of companions on the smallest scales and occurs only on scales larger than the accretion disc scale.

The upper mass limit for high-mass stars has been investigated for decades. The creation of a  $50 M_{\odot}$  star requires a large amount of material of  $\sim 10^5 M_{\odot}$ , which is about the typical mass of a GMC (e.g. [Zinnecker & Yorke 2007](#); [Ballesteros-Paredes et al. 2020](#)). In the monolithic collapse model, all material to form a star has to be present before formation starts, so this would therefore limit the upper mass of massive stars that can form in isolation. [Weidner & Kroupa \(2004\)](#) suggested that very massive clusters like R136 in the LMC have the potential to form stars with masses up to  $\sim 750 M_{\odot}$ ; however,

the most massive stars observed at the time had masses  $<150 M_{\odot}$ .

[Crowther et al. \(2010\)](#) showed that this mass is not the limit for massive stars and located several stars in R136 that could be up to double that mass at  $\sim 300 M_{\odot}$ . [Schneider et al. \(2018\)](#) also probed this cluster and showed that it is well populated with a large number of massive stars between  $15\text{--}200 M_{\odot}$  and that the high-mass slope for the IMF that results from these observations will be shallower than the classical Salpeter IMF. Future observations of massive star-forming environments might lead to a need to re-evaluate the shape of the high-mass slope of the IMF, which is still virtually unchanged since its inception in 1955.

### 1.3.5 Binary systems

Stars exist in binaries and other multiple system orbiting together around their common centre of mass and they have been known for several hundreds of years ([Michell 1767](#); [Herschel 1803](#)). They are in fact a natural outcome of the star formation process and have been observed during the pre-MS phase (e.g. [Mathieu 1994](#)) and even at the earliest protostellar Class 0 phase (e.g. [Tobin et al. 2016b](#)) suggesting that they form in the earliest stages of the formation process. During the earliest stages, only  $\sim 9\%$  of Class 0 protostars have a visual companion. However, this low percentage is likely influenced by the difficulty in observing the central, heavily embedded object. At the less embedded Class I stage,  $\sim 44\%$  of the stars have a visual companion (e.g. [Duchêne & Kraus 2013](#), and references therein).

Binaries and multiple systems are observed for MS-stars in the field across the whole mass range and even for BDs. More than 80% of all high-mass O-stars are found in multiples, reducing to  $\geq 60\%$  for B-stars. This frequency decreases to  $\sim 45\%$  in stars with solar mass, dropping further for lower-mass M-dwarfs ( $0.1\text{--}0.5 M_{\odot}$ ) to around  $\sim 26\%$  and is lowest for BDs with a multiple frequency of  $\sim 20\%$ . The increase in multiple frequency with increasing mass suggests that when more material is available in larger GMCs to form higher-mass stars, the fragmentation process can create more than one seed in each core. This then leads to higher frequencies of multiple protostars within the same gaseous envelope for higher-mass stars (e.g. [Duchêne & Kraus 2013](#), and references therein).

Most models for the formation of multiple systems centre around the fragmentation that occurs early in the star-formation process forming pre-stellar cores resulting in primordial binaries. Binaries can also form via capture of another star (dynamical binaries). However, this formation route is considered to be inefficient (e.g. [Tohline 2002](#)).

The first formation model is turbulent fragmentation of the GMC (e.g. [Padoan & Nordlund 2002](#); [Goodwin et al. 2004](#); [Offner et al. 2010](#)). In this model, turbulence is caused by inhomogeneities in the velocity fields and the density within a GMC leading to fragmentation into more than one seed per core. The higher the turbulence, the higher the number of seeds in each core (e.g. [Goodwin et al. 2004](#)). The second model is driven by solid-body rotation and is the thermal fragmentation of a perturbed, rotating cloud during collapse (e.g. [Tohline 1981](#); [Burkert & Bodenheimer 1993](#); [Boss & Keiser 2013](#)).

The third model is the fragmentation of the circumstellar disc due to gravitational instabilities (e.g. [Adams et al. 1989](#); [Tobin et al. 2016a](#)). These instabilities can be caused for example by dynamical interactions with passing stars within a dense star-forming region. An already formed binary surrounded by a disc can also perturb its own disc and cause fragmentation to form an additional companion (e.g. [Bonnell & Bate 1994a](#)). A rotationally unstable protostar at the centre of the disc can also perturb the disc and initiate the disc's fragmentation (e.g. [Bonnell & Bate 1994b](#)). [Sigalotti et al. \(2018\)](#) showed that the disc can also become unstable and fragment simply due to it growing during the protostellar phase.

## 1.4 Protoplanetary discs

Discs form at almost the same time as the stars that they surround, while they are still located in higher density regions (e.g. [Haisch et al. 2001](#); [ALMA Partnership et al. 2015](#); [Andrews et al. 2018](#)). Discs have been observed around stars with masses ranging from very low-mass up to intermediate-mass stars ( $0.08\text{--}8 M_{\odot}$ ). There have also been observations of discs around BDs that suggest similar disc fractions as those for young T-Tauri stars (e.g. [Luhman et al. 2005](#); [Rilinger & Espaillat 2021](#)). Discs have not been observed around visible O-stars on the MS; however, they have been observed around them while these stars are still in their embedded protostellar phase (e.g. [Johnston et al. 2015](#)). It is unclear if the lack of discs around visible O-stars is due to the high photoevaporation rate clearing the visible components of the disc too quickly or due to a different star formation mechanism (e.g. [Zinnecker & Yorke 2007](#)). In their study of intermediate-mass stars, [Guzmán-Díaz et al. \(2021\)](#) showed how disc lifetimes correlate with stellar masses based on the photoevaporation scenario described in [Alexander et al. \(2014\)](#). In this scenario, the disc lifetimes drop to  $\sim 40$  kyr for stars with masses  $\sim 10 M_{\odot}$  ([Guzmán-Díaz et al. 2021](#)).

In the following sections, I will briefly discuss the formation, evolution and destruction/dispersal of the discs and how they can be observed.

### 1.4.1 Formation and evolution of protoplanetary discs

When GMCs collapse to form stars, the more distant material with higher angular momentum flattens to form a disc (e.g. [Terebey et al. 1984](#)). This disc forms rapidly within  $\sim 10^4$  yr during the Class 0 phase of the forming protostar (e.g. [Hueso & Guillot 2005](#)). This protostellar disc and the central young star remain embedded in the gaseous envelope throughout the Class I phase for  $\sim 0.5$  Myr (e.g. [Evans et al. 2009](#)). The average mass of the disc during these two phases is not dependent on the evolutionary state of the protostar at its centre. The discs show a median disc mass value of  $0.04 M_{\odot}$  and a range from  $\sim 0.02\text{--}0.1 M_{\odot}$ . 99% of the total mass of the disc at this stage is made up of gas, with the remaining 1% residing in dust (e.g. [Williams & Cieza 2011](#)).

The independence of the disc mass from the evolutionary stage suggests that the material that flows from the envelope onto the disc does not remain there but is accreted quickly onto the protostar from the inner disc edge, which can make the disc unstable (e.g. [Young et al. 2003](#)). This disc instability can lead to episodic accretion events, such as those observed for example in FU-Orionis-type events (e.g. [Vorobyov & Basu 2010](#)).

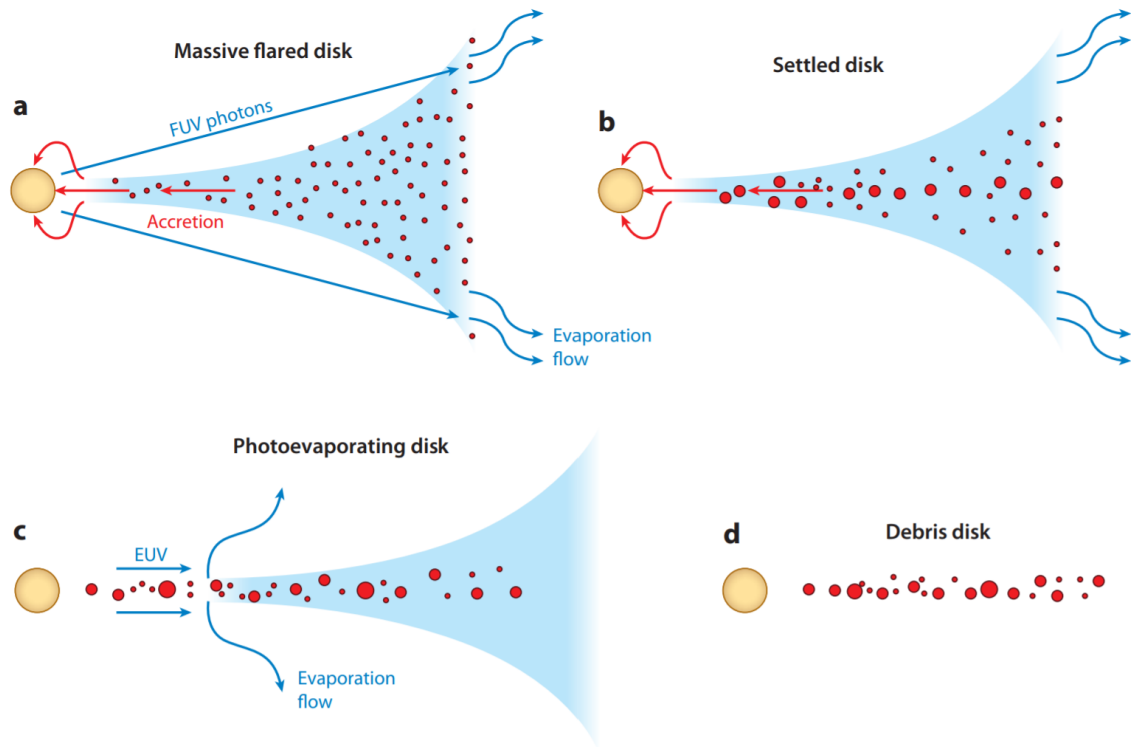


Figure 1.10: The four main evolutionary stages of a typical disc are shown in this figure from [Williams & Cieza \(2011\)](#). The gas is shown in blue; the dust is shown using the red points. The disc can be flared at the outer edges to larger scale heights due to hydrostatic balance. In panel (a), the disc goes through mass-loss due to accretion onto the star on the inner edge and the outer disc is photoevaporated by UV-radiation from the central star. Panel (b) shows a settled disc, with increasing dust grain sizes growing larger and settling in the disc's mid-plane. In panel (c), the mass of the disc and the accretion rate decreases and photoevaporation increases. The inner disc starts to dissipate from the inside as the outer disc is unable to resupply it with material. Panel (d) shows the final stage of the disc with no remaining gas. In this debris disc, smaller grains are removed and only larger bodies, such as planetesimals and possibly planets remain. Figure credit: [Williams & Cieza \(2011\)](#)

The size of the discs is typically a few 100s au, with the disc edge softening (decreasing surface density) between  $\sim 20$ -200 au. At the end of the embedded phase (end of Class I), the gaseous envelope disperses leaving a central star surrounded by the protoplanetary disc, with most of the mass concentrated in the central star and  $\sim 1\%$  left in the disc. At this point the central star enters the Class II phase as an accreting pre-MS star (CTTS or Herbig Ae/Be star) (e.g. [Williams & Cieza 2011](#), and references therein).

The evolutionary stages from a gas-rich protoplanetary disc to a gas-free debris disc are illustrated in Figure 1.10. In panel (a), there is continuing accretion of material from the disc onto the star. UV photoevaporation from the protostar itself slowly depletes the mass of the disc, which now cannot be further replenished via accretion onto the disc from the dispersed envelope. In panel (b), the disc evolves further with dust grain sizes growing larger and settling in the disc's mid-plane, possibly leading to planet formation (e.g. [Williams & Cieza 2011](#)).

The accretion of material from the disc onto the star requires a loss of angular momentum. However, the conservation of angular momentum requires this loss to be balanced out. While accretion moves material in the disc inwards towards the star, some material will be transferred outwards via viscous spreading/expansion, thereby increasing the disc size. The viscosity (internal friction) of the disc is proportional to the rate of accretion and in turn the rate at which angular momentum transfer has to occur (e.g. [Lynden-Bell & Pringle 1974](#); [Hartmann et al. 1998](#); [Hueso & Guillot 2005](#)). However, it is still not fully understood where this viscosity comes from and how the angular momentum transfer throughout the disc takes place (e.g. [Williams & Cieza 2011](#)).

A second mechanism that can preserve the angular momentum is a stellar wind (or outflow) that removes material from the disc but does not change the size of the disc itself (e.g. [Williams & Cieza 2011](#)). [Trapman et al. \(2020\)](#) studied the evolution of disc sizes over time to be able to distinguish between these two mechanisms of disc evolution. The authors concluded that current observations of discs are consistent with viscous evolution. [Armitage et al. \(2013\)](#) showed that it is possible that a combination of the two mechanisms can result in disc evolution on two different timescales. The first is a long phase of viscous evolution followed by a rapid phase of mass loss via winds, which can then disperse a disc on a much shorter timescale. Winds could be another effective mechanism to disperse the disc in addition to the dispersal effect of photoevaporation (e.g. [Clarke et al. 2001](#); [Armitage et al. 2013](#)).



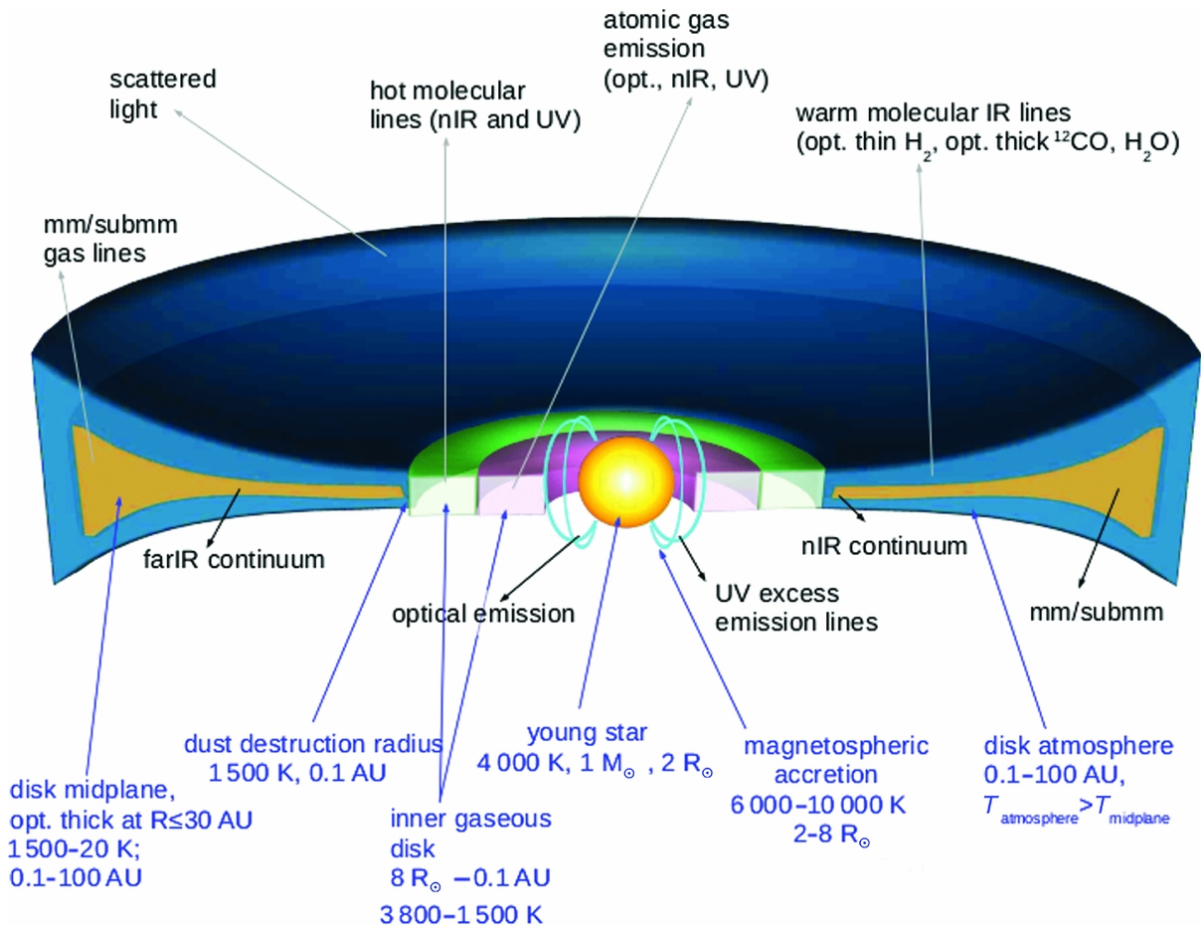


Figure 1.11: This figure from Sicilia-Aguilar et al. (2016) shows a cross-section of a protoplanetary disc with indication of how multi-wavelength observations trace different processes and regions in the disc. The three different categories of tracers are shown with scattered light tracing the surface of the disc where the grains can scatter the light from the central star. Continuum emission comes from the dust/solids located in the mid-plane of the disc with emission in the nIR found towards the inner disc and wavelengths increasing to far/IR and mm/sub-mm wavelengths towards the edges of the disc. Spectral line emission in different wavelengths is present across the whole disc with shorter wavelength emission located closer to the central star. Figure credit: Sicilia-Aguilar et al. (2016)

### 1.4.2 Observations of protoplanetary discs

The disc is hottest at the centre close to the protostar and gets cooler towards the edges. The dust emits black-body-like radiation, so the hotter inner regions of the disc will emit in the IR, whereas the cooler outer region will emit at mm/sub-mm wavelengths. Figure 1.11 shows a cross-section of a protoplanetary disc highlighting emission from different regions of the disc (Sicilia-Aguilar et al. 2016).

Three main categories of tracers are used to observe discs around young stars. Even though these discs are dominated by gas, the dust is much easier to observe. The solids absorb the light from the star and then emit thermal continuum emission across a wavelength range from 1  $\mu\text{m}$  - 1 cm, peaking in the sub-mm region. Scattered light is emission from the central star that is reflected off small grains in the gas at IR wavelengths. Both of these tracers allow the observation of the dust disc with thermal emission often used to determine disc masses using gas-to-dust ratios. The third tracer - spectral line emission - traces the gas disc. Most of the gas in the disc is molecular hydrogen, which cannot be directly observed. Spectral (molecular) line emission originates in the transitions of molecules between energy levels. One of the most commonly observed emission lines from discs is from CO, detected in mm/sub-mm wavelengths (e.g. Sicilia-Aguilar et al. 2016; Zhao et al. 2020; Andrews 2020).

The discs can be observed throughout most of the protostellar and pre-MS evolutionary phases. In the embedded phase, discs are observed easiest using mm/sub-mm and cm-wave interferometry, which measures the continuum emission of the dust in the disc. At this evolutionary stage, it is difficult to separate the line emission from the gaseous envelope from that coming from the gas components of the disc (e.g. Tobin et al. 2015b; Zhao et al. 2020). Further difficulties arise due to the smaller size of the discs during these embedded phases (e.g. Segura-Cox et al. 2018; Maury et al. 2019) and the general rarity of protostars (and protostellar discs) compared to pre-MS stars in nearby young star-forming regions (e.g. Tobin et al. 2020).

Protoplanetary discs become easier to observe when they are no longer embedded in the envelope but are still in their gas-rich phase with accretion from the disc onto the star (Class II). Spectral line emission from the gas is now no longer entangled with the envelope gas allowing the gas disc to be measured (e.g. Tobin et al. 2015b; Zhao et al. 2020). Close to the star, the high temperatures cause fluorescent electronic transitions that produce excess UV emission, which can be used to probe the accretion from the inner disc onto the star (e.g. Rebull et al. 2000; Herczeg et al. 2006). NIR and H $\alpha$  (optical) emission from the gas are other measures of the mass of the inner gaseous disc

and the accretion rate. The presence of this emission is commonly used to differentiate between accreting and non-accreting stars (e.g. [Galli et al. 2015](#)). Emission from the gas reduces in the later evolutionary stages of the disc, when the accretion rate drops and photoevaporation starts to dominate (Fig. 1.10, panel c). This can be observed as a rapid decline of the IR-excess from young stars, with almost 100% of all stars in young clusters ( $\sim 1$  Myr) showing an excess indicative of a disc, whereas only  $<10\%$  of stars do so at cluster ages larger than 5 Myr (e.g. [Haisch et al. 2001](#); [Mamajek 2009](#); [Kraus et al. 2012](#); [Richert et al. 2018](#)).

### 1.4.3 Destruction/dispersal of the discs and/or planet formation

Protoplanetary discs usually evolve quickly in just a few Myr into gas-poor debris discs with median lifetimes of about 3 Myr and a range from less than 1 Myr to  $\sim 10$  Myr (e.g. [Haisch et al. 2001](#); [Richert et al. 2018](#); [Andrews 2020](#)). Most stars completely lose their discs in the first 5 Myr. However, it has recently been suggested that discs can survive up to  $\sim 10$ -12 Myr ([Bell et al. 2013](#)) and observations exist of even older (several tens Myr) discs around stars that are still accreting (e.g. [Manara et al. 2013](#); [Murphy et al. 2018](#)).

[Silverberg et al. \(2020\)](#) recently introduced the term “Peter Pan” discs for discs around young, accreting M-stars that are older than 20 Myr. They observed several candidates that show evidence of both a debris and primordial disc. [Wilhelm & Portegies Zwart \(2021\)](#) followed up these results with  $N$ -body simulations and concluded that these discs can only survive around M-stars for longer than 20 Myr. In low-radiation environments around stars with masses  $\lesssim 0.6 M_{\odot}$  protoplanetary discs can survive up to 50 Myr. The estimated survival times of discs are also strongly influenced by the choice of model to estimate cluster ages, which in turn affects the disc lifetime estimates (e.g. [Bell et al. 2013](#); [Richert et al. 2018](#)).

There are different disc dispersal/destruction mechanisms. Panel (c) in Fig. 1.10 illustrates the stage when the disc quickly photoevaporates from the inside out as the outer disc is unable to resupply the inner disc with material. An inner hole forms around the star, which accelerates the internal photoevaporation. The dissipation of the disc is faster when the mass of the central star is higher. Very low-mass stars and BDs have disc lifetimes that are at least on the scale of that of solar mass stars (e.g. [Williams & Cieza 2011](#), and references therein).

The disc is not only affected by internal processes but also by external effects that can cause truncation or even the complete destruction of the discs (e.g. [Williams & Cieza](#)

2011; Hardy et al. 2015). One is the effect of photoevaporation due to UV radiation from nearby massive stars (e.g. Johnstone et al. 1998; Adams et al. 2004; Haworth et al. 2018; Nicholson et al. 2019; Concha-Ramírez et al. 2021). This external photoevaporation is likely to leave the interior of the disc intact. However, it can erode the outer part (at radii  $> 50$  au). More extreme effects are usually only expected for discs that are very close to a massive O star (e.g. Williams & Cieza 2011).

The second external process is the effect of dynamical interactions between the stars in regions of higher density (e.g. Adams et al. 2006; Olczak et al. 2008; Parker & Quanz 2012; Vincke & Pfalzner 2016). Flybys in these regions can cause warped discs, truncation, and they also have an effect on the size of the discs. When massive stars are present in the vicinity, photoevaporation is likely to dominate (e.g. Adams et al. 2006; Guarcello et al. 2016; Winter et al. 2018; Vincke & Pfalzner 2018; Cuello et al. 2019, 2020).

A third process affecting the lifetime of protoplanetary discs is the formation of planets (e.g. Stammler et al. 2019; Pinte et al. 2020). In the core accretion model of planet formation, the process starts with dust grains that grow into rocks via coagulation, which then form planetesimals and finally grow into protoplanetary cores that are massive enough to capture gas from the protoplanetary disc, while it still contains gas (e.g. Andrews 2020). It is commonly assumed that this process starts during the Class II stage of the evolution, when the disc is settled. However, in very massive discs with high gas and dust densities, this process can already start in the earlier embedded phases (e.g. Tsukamoto et al. 2017; Andrews 2020).

While the protoplanetary disc evolves and the inner hole forms, the disc stops emitting strongly in the nIR, which also marks the transition to a gas-poor debris disc (Fig. 1.10, panel d). Debris discs are virtually gas-free and small dust grains are being removed due to radiation pressure leaving only larger rocks, planetesimals and potentially planets (e.g. Williams & Cieza 2011, and references therein). At this point the central stars are WTTSs. While many WTTSs show no evidence of any remaining gas in their debris discs, WTTSs still showing weak accretion or discs in transition have been identified (e.g. Cieza et al. 2007, 2013; Hardy et al. 2015).

## 1.5 Runaway stars and their formation mechanisms

Fast-moving massive stars were first reported over 60 years ago (Humason & Zwicky 1947; Blaauw & Morgan 1954; Blaauw 1956). Humason & Zwicky (1947) found B-type stars at very large distances from the Galactic disc and suggested that they were unlikely born there. Amongst the first high-velocity stars discovered was AE Aurigae (AE Aur), a young star with an O9.5 spectral type. It was shown to move with a large velocity, both in its radial as well as tangential component (Blaauw & Morgan 1953). Shortly thereafter, Blaauw & Morgan (1954) discovered  $\mu$  Columbae ( $\mu$  Col), which moves with almost exactly the same velocity as AE Aur in virtually the opposite direction with both stars tracing back to the Orion star-forming region. The authors hypothesised that these two stars could have been created in the same event, but left a firm conclusion to that effect open in their 1954 paper (Blaauw & Morgan 1954).

The term runaway (RW) star for these high-velocity stars first appeared in literature in 1957 (Greenstein 1957) to describe high-velocity stars in the Galactic halo but is more commonly attributed to Blaauw (1961). Blaauw (1961) surveyed then available data on fast-moving O- and B-stars and proposed a theory for their origin. It is based on rapid mass loss from the primary component from either a SN or the contraction phase in the evolution of very massive stars destabilising the binary. The author suggested that this hypothesis explains the often single star status of RWs with large masses. This explanation is now considered one of the two main creation mechanisms for RW stars and is called the Binary-Supernova-Scenario (BSS). Section 1.5.1 provides a more detailed explanation of this first mechanism. In Section 1.5.2, I introduce the second formation mechanism, which is the Dynamical-Ejection-Scenario (DES), first proposed by Poveda et al. (1967). This mechanism explains ejections as a result of dynamical interactions in denser regions of star-forming environments.

Historically, ejected stars were considered to fall into the RW regime when their peculiar velocity (velocity in relation to a rest frame) exceeded  $\sim 30\text{-}40\text{ km s}^{-1}$  (e.g. Blaauw 1956, 1961; Stone 1991; Hoogerwerf et al. 2001). The original velocity cut-off ( $40\text{ km s}^{-1}$ ) used by Blaauw (1961) is based on the finding that “normal” OB-stars show a peculiar velocity of  $\sim 15\text{ km s}^{-1}$ . Placing a cut-off boundary between “normal” and RW stars at the  $3\sigma$  level results in the  $40\text{ km s}^{-1}$  boundary. Stone (1991) suggested that this boundary is a very conservative assumption, and one could miss a considerable proportion of lower-velocity RWs ( $\sim 38\%$  suggested in Stone 1979) and thereby underestimate the RW star frequency. They suggested a classification based on a group of stars (“High Velocity Group”), all of which are the result of one of the RW formation mechanisms

(regardless of velocity). “High Velocity Stars” are a sub-group of stars showing peculiar velocities of  $> 30\text{-}40\text{ km s}^{-1}$ , followed by another sub-group, which are the classical OB-RWs (Stone 1991). At the time, this peculiar velocity was mostly based on radial velocities. However, with the advent of larger-scale astrometry surveys, such as *Gaia* (see Chapter 3), RWs are nowadays more likely to be identified based on their tangential velocities rather than their radial velocities.

Stone (1991) suggested that setting velocity boundaries to identify RWs will lead to observers missing the full number of stars that are created via the two suggested formation mechanisms. Apart from finding young stars in locations where star formation is uncommon (e.g. Andersson et al. 2021), the best way to determine RW stars is still based on finding significant differences in their peculiar velocities compared to their surrounding stars or to the underlying stellar population. The classical velocity boundary was questioned more recently, and it was proposed to include slower ejected stars. These stars can also travel several tens of pc within the lifetime of star-forming regions, resulting in stars being found far outside their natal regions (e.g. de Wit et al. 2005; Eldridge et al. 2011; de Mink et al. 2014; Tetzlaff et al. 2011; Drew et al. 2018). Eldridge et al. (2011) suggested a lower limit of  $\sim 5\text{ km s}^{-1}$  for these slow RW stars and de Mink et al. (2012; 2014) subsequently termed them walkaway (WW) stars. These WWs are not part of a separate population of ejected stars but constitute the low-velocity end of the RW phenomena and are the outcome of the same formation mechanisms. In many studies, these slower stars are not even separately identified but included in the RW definition when choosing a lower cut-off velocity (e.g. Tetzlaff et al. 2011; Perets & Šubr 2012).

The velocity required to escape from a star-forming region depends on the total mass of this region and the location of the escaping star within the region. The escape velocity is given by:

$$v_{esc} = \sqrt{\frac{GM}{R}}. \quad (1.4)$$

$G$  is the gravitational constant,  $M$  is the total mass of the region and  $R$  is the distance of the escaping star in relation to the centre of the region it is escaping from. This equation highlights that the escape velocity does not depend on the mass of the escaping star, but only on the mass of the region it is escaping from. The gravitational potential well is deeper towards the denser parts of the star-forming regions, where more stars and therefore more mass is located. The closer an escaping star is located to these regions (smaller value for  $R$ ), the higher its escape velocity needs to be. This is also illustrated by the escape velocity of the Milky Way, which differs greatly with distance to its centre

(see Section 1.5.4).

At the time of discovery, it appeared that RW stars were an occurrence purely amongst high-mass O- and B-stars. This thinking continued for decades until in the early 2010s, [Tetzlaff et al. \(2011\)](#) suggested that this traditional assumption should be challenged and that lower-mass stars may also be able to become RW stars. These authors sampled the peculiar velocities of over 7000 stars (pre-MS and MS) from the HIPPARCOS catalogue in one, two and three dimensions to identify high-velocity members with different kinematics than normal population I stars. They identified over 2500 RW star candidates (some known but most of them new candidates), several of them with lower masses than the classical OB-RWs. They suggested that young stars ( $\leq 50$  Myr), regardless of their mass, can share similar kinematics. In Fig. 1.12, I use the data from [Tetzlaff et al. \(2011\)](#) to show the distribution of stellar masses of the RW-candidates against their 2D/3D velocities. The peculiar tangential (2D) velocity cut-off is  $20 \text{ km s}^{-1}$ , whereas the cut-off for the peculiar spatial (3D) velocity is higher at  $28 \text{ km s}^{-1}$ . The figure illustrates that there is a large number of low-mass stars with a lowest mass of  $0.6 M_{\odot}$  amongst the 2D- and 3D-RW candidates identified with over 90% probability. There is also a clear drop in RW-numbers between  $\sim 2\text{-}5 M_{\odot}$ . This is likely driven by the underlying selection of the young stars in [Tetzlaff et al. \(2011\)](#) as the masses of all stars analysed in their study show a similar gap between approximately  $3\text{-}5 M_{\odot}$ . This plot highlights that lower-mass stars can reach similarly high, if not higher velocities than the higher-mass stars.

[de la Fuente Marcos & de la Fuente Marcos \(2005\)](#) discovered the fast-moving star LP 543-25 and suggested it could be a low-metallicity M-dwarf former member of the old thick disc of the Milky Way. Using *Gaia* DR2, [de la Fuente Marcos & de la Fuente Marcos \(2018\)](#) revisited this star and used updated kinematic information to identify the star as bound to the Galaxy, moving with  $\sim 200 \text{ km s}^{-1}$  away from the Galactic disc. They confirmed their earlier conclusion of it being a low-mass K/M-dwarf. [McBride & Kounkel \(2019\)](#) searched for ejected stars from the Orion Nebula region and found several lower-mass RW/WW stars. Their findings are discussed in Chapter 5 in more detail. [Bischoff et al. \(2020\)](#) used the catalogue of [Tetzlaff et al. \(2011\)](#) as their basis to search for RWs in *Gaia* DR2 and determined their youth using isochrones and the presence of lithium in their spectra. They identified over 10 dwarf (low-mass) stars that fit their criteria of a young age. [Platais et al. \(2020\)](#) used data from the Hubble Space Telescope (HST) in combination with *Gaia* DR2 to identify four further RWs/WWs from the Orion Nebula region with masses  $< 1 M_{\odot}$ . All of these recent studies further

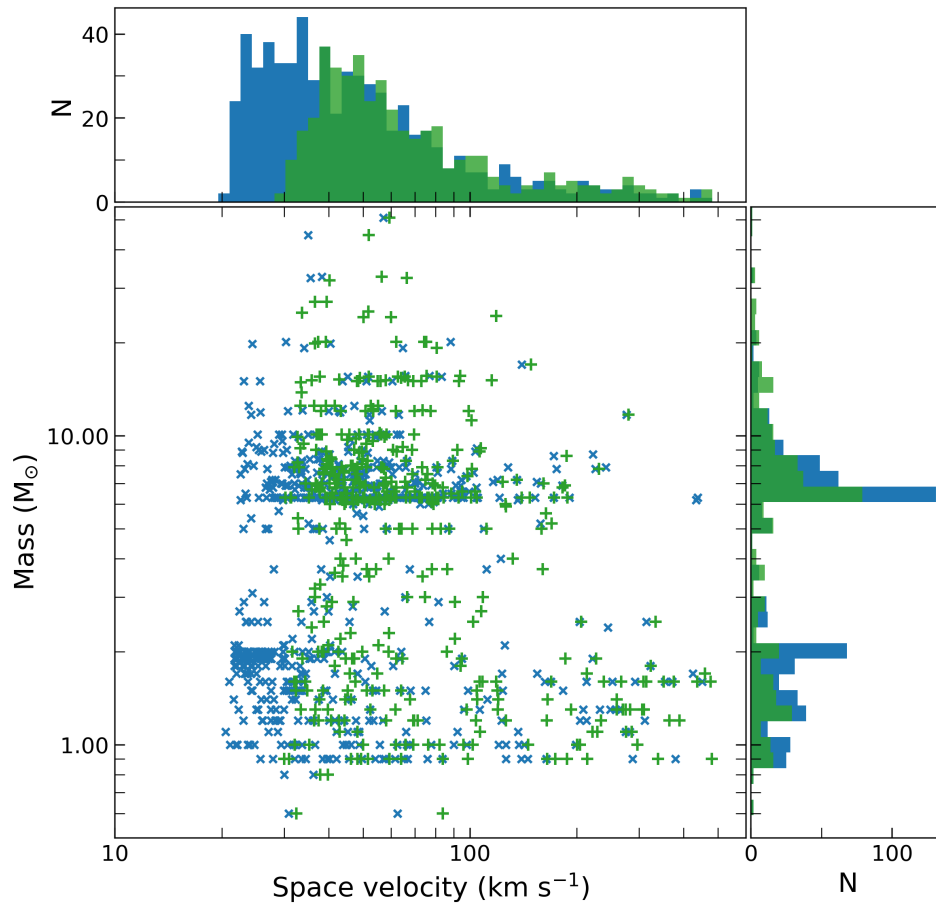


Figure 1.12: RW stellar masses in  $M_{\odot}$  vs. peculiar velocity in  $\text{km s}^{-1}$  (tangential or space velocity) based on data from [Tetzlaff et al. \(2011\)](#). I show RWs identified from their 2D peculiar tangential velocity with RW-probabilities  $> 90\%$  with a “blue x” and RWs identified from their 3D peculiar spatial velocity with RW-probabilities  $> 90\%$  with a “green +”. The 3D-RWs have a velocity cut-off at  $28 \text{ km s}^{-1}$ , whereas the 2D-RWs have a velocity cut-off at  $20 \text{ km s}^{-1}$ . This plot shows that the highest velocities appear to be reached by lower-mass stars. The plot also shows a lack of RWs between about  $2\text{--}5 M_{\odot}$ , which is likely driven by the underlying selection of the young stars in [Tetzlaff et al. \(2011\)](#). The masses of all stars analysed in this study show a similar gap between approximately  $3\text{--}5 M_{\odot}$ .



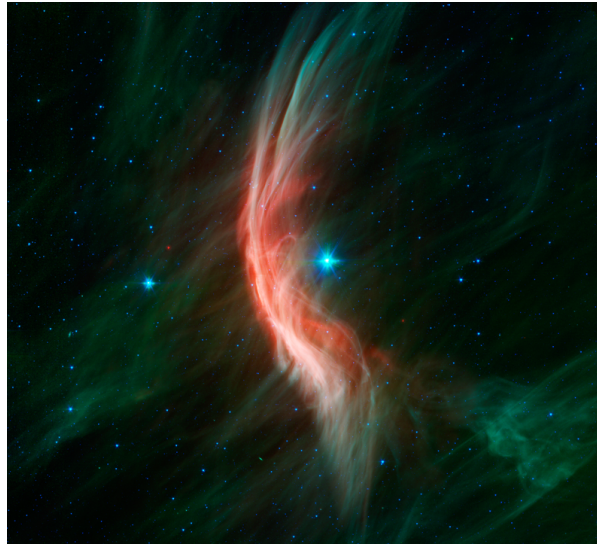


Figure 1.13: Bow shock produced by the 3 Myr old, massive ( $19.8 M_{\odot}$ ), O9-type RW star  $\zeta$  Ophiuchi. The high velocity of this star causes a perturbation of the ISM leading to the observed bow shock phenomena, which is scattered light or emission at the shock front. The direction of motion of the star in this image is from right to left. Figure credit: NASA/JPL-Caltech.

underpin the notion that RWs and WWs can occur at any mass and that the original assumption of RW stars purely occurring amongst OB-stars was likely due to observation (detection) bias.

OB-RWs can create an additional observable characteristic as a side effect of their high space velocities - a bow shock. Bow shocks form in the direction of motion of objects such as high-mass RW stars that perturb the ISM. The velocity of these stars will be above the sound speed of the ISM, i.e. at supersonic speed. When the fast stellar winds of the high-mass RW stars come into contact with the ISM, they decelerate very quickly leading to shocked gas. The RW stars compress the ISM matter into dense, thin shells, which can produce emission or scatter light thereby making the bow shock visible (Wilkin 1996). Peri et al. (2012) provided an overview of previously identified bow shocks and conducted an extensive search, suggesting that about 10% of all OB-RWs produce bow shocks. Fig. 1.13 provides an image taken with the Spitzer Space Telescope of such a bow shock created by the interaction of the massive O-type star  $\zeta$  Ophiuchi with the ISM. The peculiar space velocity of this star is below the classical boundary for RW stars at  $\sim 25 \text{ km s}^{-1}$  (velocity from Tetzlaff et al. 2011). Bow shocks can also be produced by lower-mass stars not moving at supersonic speeds, e.g. LL Orionis. It is a young (0.3 Myr) K-type star with a mass estimate of  $2 M_{\odot}$  located in the Orion Nebula

(Hillenbrand 1997). The reason for the presence of a bow shock around this type of star is linked to outflows, such as bipolar jets (e.g. Bally et al. 2006, 2012; Chick et al. 2020).

RWs have been used as an explanation for the apparent in-situ formation of isolated OB-stars (see Section 1.2.4) but have also been suggested as an explanation for other phenomena. Andersson et al. (2021) recently suggested that RW stars can be mistaken as star formation if they are found in uncommon places. The authors used hydrodynamical simulations to show that the presence of young massive stars that are ejected into low-density gas environments can mimic the appearance of star formation, where there is none. They concluded that RWs are not only contributing but could be a dominating factor explaining star formation observations in the outskirts of spiral galaxies (e.g. Bigiel et al. 2010). Fenske et al. (2021) used RWs, in particular the stellar winds of OB-RWs, as a potential explanation for the presence of elliptical ring structures around NGC 7538, a high-mass star-forming region  $\sim 2.7$  kpc from Earth.

In the following sections, I describe the two formation scenarios (BSS and DES) for RWs and WWs. I then briefly discuss ways of distinguishing between the formation scenarios. I conclude this section describing the faster hyper-runaway (HRW) stars and even faster hypervelocity (HVS) stars.

### 1.5.1 Binary supernova scenario - BSS

Theorised first by Zwicky (1957), the BSS was described in more detail by Blaauw (1961). This formation scenario starts with a close binary system with a primary of at least  $\sim 8 M_{\odot}$ . Fig. 1.14 from Renzo et al. (2019) illustrates the key stages in the evolution of a close massive binary leading to disrupted systems that can eject stars.

The more massive primary star will reach the end of its stellar evolution first and undergo a core-collapse SN. If the exploding star ejects more than half of the total system mass during this SN, the binary can be completely disrupted after this first SN, leading to the secondary, less massive companion star being ejected. The following derivation illustrates this and is based on a simplified case involving a binary with a circular orbit without pre-SN mass transfer. The binding energy  $E$  of a binary is described by:

$$E = -\frac{Gm_1m_2}{2a}, \quad (1.5)$$

where  $G$  is the gravitational constant,  $m_1$  and  $m_2$  are the stellar masses of the two stars and  $a$  is the semimajor axis. I assume a circular orbit so  $a$  is the radius  $r$ . The binding energy is the sum of the potential  $\Omega$  and kinetic energy  $T$ . Using the reduced mass

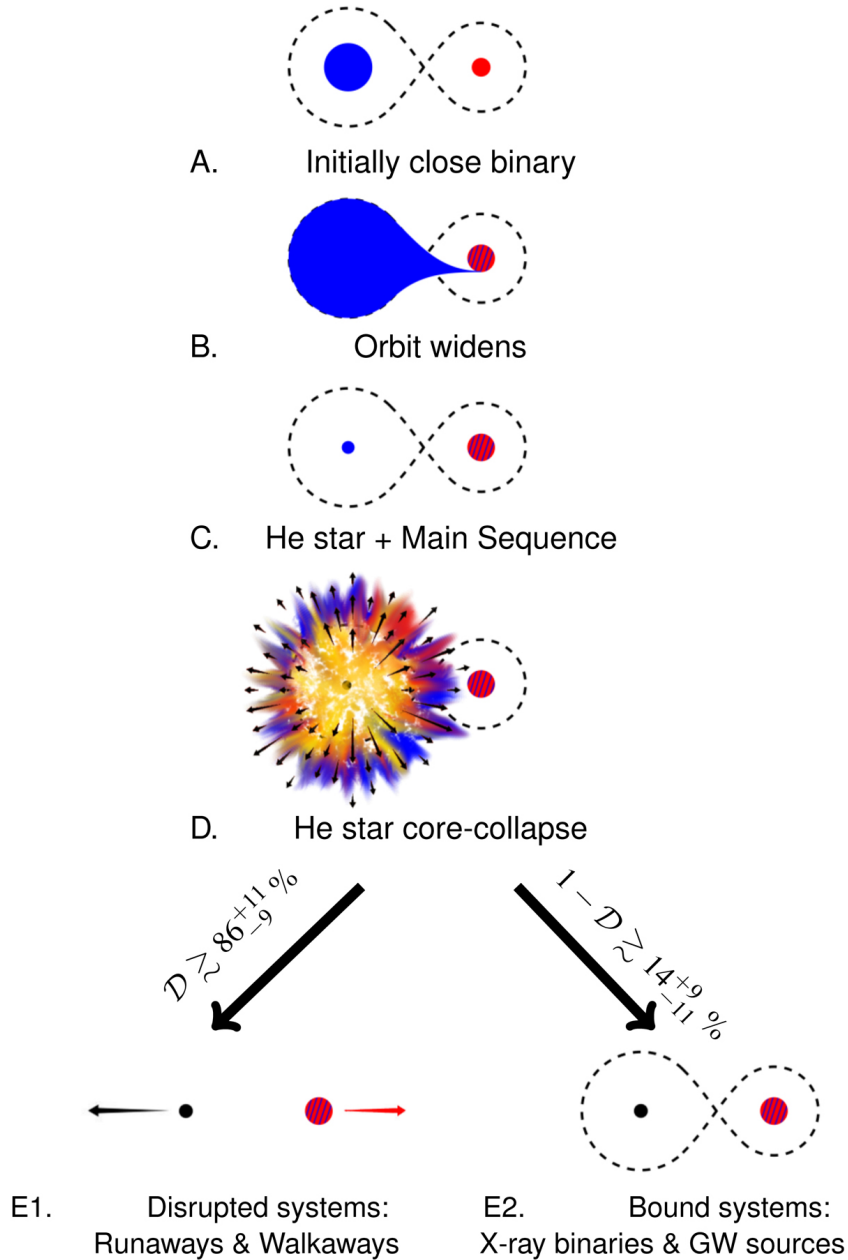


Figure 1.14: Evolution of a massive binary from an initially close binary (A). The evolution is shown for binaries that do not merge before the primary star undergoes SN. Stable mass-transfer via Roche-Lobe overflow from the more massive primary star onto the secondary star causes the orbit to widen (B). The primary star eventually undergoes a core-collapse SN (D). Following the results in [Renzo et al. \(2019\)](#), in the majority of cases ( $D =$  disruption fraction), this leads to the disruption of the binary and the formation of RWs and WWs (E1). Figure credit: [Renzo et al. \(2019\)](#).

$\mu = m_1 m_2 / (M)$  (assuming that  $m_2 \ll m_1$ ), where  $M = m_1 + m_2$  for the total binary mass, results in the following equations for the binding energy  $E_0$  of the system pre-SN (the index 0 indicates the situation before the SN):

$$\begin{aligned} E_0 &= -\Omega_0 + T_0 \\ -\frac{GM_0\mu_0}{2r_0} &= -\frac{GM_0\mu_0}{r_0} + \frac{1}{2}\mu_0 v_0^2. \end{aligned} \tag{1.6}$$

Solving this equation for the centre of mass velocity  $v_0^2$  :

$$v_0^2 = \frac{GM_0}{r_0}.$$

The binary will remain bound post-SN, if the total energy after the explosion  $E_1$  is less than zero with  $E_1 = 0$  as the bound/unbound boundary. Assuming that at the time of the ‘‘instantaneous’’ mass loss due to the SN neither the radius nor the velocity of the binary change:

$$\begin{aligned} 0 &= -\Omega_1 + T_1 \\ 0 &= -\frac{GM_1\mu_1}{r_0} + \frac{1}{2}\mu_1 v_0^2 \\ \frac{GM_1\mu_1}{r_0} &= \frac{1}{2}\mu_1 v_0^2 \\ \frac{GM_1}{r_0} &= \frac{1}{2}v_0^2. \end{aligned} \tag{1.7}$$

Substituting  $v_0^2$  from Eq. 1.6 :

$$\begin{aligned} \frac{GM_1}{r_0} &= \frac{1}{2} \frac{GM_0}{r_0} \\ M_1 &= \frac{M_0}{2}. \end{aligned}$$

The above result highlights that the binary will remain ‘‘just’’ bound if the total mass after the SN  $M_1$  is half of the mass of the binary  $M_0$  pre-SN. If more than half of the system mass is ejected then the total energy  $E_1$  becomes positive, thereby unbinding the binary. In this simplified case, I have assumed that the mass of the secondary  $m_2$  is much smaller than that of the exploding primary star. The more equal the masses of the binary stars are pre-SN, the higher the mass loss has to be to unbind the binary (e.g.

Blaauw 1961).

The relative velocity of the binary (centre of mass) before the SN is  $v_0$  as shown in Eq. 1.6. After the SN, it is  $v_1 = v_0 + v_{\text{kick}}$ . In the case of a symmetric explosion,  $v_{\text{kick}} = 0$ . If the mass loss is large enough to unbind the binary, the secondary star can be ejected from the binary due to the ‘‘Blaauw–Boersma recoil’’ or ‘‘Blaauw kick’’ (Blaauw 1961; Boersma 1961). This recoil occurs, as even though the SN is symmetric in the reference frame of the primary star, it is asymmetric in relation to the centre of mass of the binary. The velocity of the SN remnant  $v_{1,1}$  in its reference frame (centre of mass of the binary) remains unchanged in this symmetric explosion compared to its velocity pre-SN  $v_{0,1}$  (see Tauris & Takens 1998):

$$v_{1,1} = v_{0,1} = \frac{m_2}{m_1 + m_2} v_0 \quad (1.8)$$

The velocity of the secondary star post-SN  $v_{1,2}$  is different compared to its pre-SN value  $v_{0,2}$ . It has been derived in different forms in literature (e.g. Blaauw 1961; Boersma 1961; Gott et al. 1970; Tauris & Takens 1998; van den Heuvel et al. 2000) and I state the form used in Tauris & Takens (1998) here:

$$v_{1,2} = \frac{\sqrt{(m_{\text{shell}}^2 - m_2 - 1)}}{m_2} v_{0,1} \quad (1.9)$$

Using Eq. 1.8 to replace  $v_{0,1}$  results in :

$$v_{1,2} = \frac{\sqrt{(m_{\text{shell}}^2 - m_2 - 1)}}{m_1 + m_2} v_0$$

If not enough mass is lost in the SN to unbind the binary, the recoil can still occur but will affect the whole binary and can then create RW-binaries (e.g. Tauris & Takens 1998).

In the unbound case, high-velocities can be imparted not only to the secondary star, but also to the SN-remnant as it continues moving with its previous orbital velocity (e.g. Gott et al. 1970; Radhakrishnan & Shukre 1985; Tauris & Takens 1998). Tauris & Bailes (1996) suggested an upper velocity limit for SN remnants from symmetric SNe of  $\sim 270 \text{ km s}^{-1}$ . Observations of a broad velocity distribution of SN remnants with a mean velocity of  $\sim 450 \text{ km s}^{-1}$  (Lyne & Lorimer 1994) therefore required an alternative explanation. Iben & Tutukov (1996) suggested that these observations could still be explained by the disruption of binaries following SNe, but several additional assumptions

had to made, potentially weakening their argument (van den Heuvel & van Paradijs 1997).

However, to explain extremely high velocities of pulsars ( $>1000 \text{ km s}^{-1}$ , for recent examples, see Section 1.5.1) SNe had to be asymmetric, which will impart a substantial natal kick onto the SN remnant itself. This kick mechanism was first suggested by Shklovskii (1970). There are different explanations for how these asymmetric kicks are produced (e.g. rapid rotation, strong magnetic fields). There could be an asymmetry in the ejected matter as well as/or in the neutrino emission (e.g. van den Heuvel & van Paradijs 1997; Lai 2001; Höflich et al. 2011; Nordhaus et al. 2012). The dynamical effects of these natal kicks on the properties of the stars in the binary were discussed, e.g. in Tauris & Takens (1998). For example, the kick could stabilise a binary that would otherwise be disrupted or in an extreme case even turn the SN remnant “stationary” in its reference frame. Even just a small asymmetry in the angle of 1% could impart kicks of  $\sim 300 \text{ km s}^{-1}$  (Burrows & Vartanyan 2021).

### What percentage of stars become RWs via the BSS?

While the BSS scenario explains one of the formation mechanisms for RWs, Blaauw (1961) considered only O- or B-stars to have been the companion stars in these binaries and subsequently become RW stars. Because of this, the binaries pre-SN had to be of very large mass. The author also suggested that double or multiple stars are virtually absent in the RW population.

Hoogerwerf et al. (2001) searched for the parent groups of a sample of 56 RW candidates. For those that they could trace back, they estimated that the BSS is possible as the formation mechanism for about 2/3 of the observed OB-RWs. A later analysis of suspected RWs around the Scorpius–Centaurus OB association by Jilinski et al. (2010) that included the candidates investigated in Hoogerwerf et al. (2001) found that several of these candidates are in fact spectroscopic binaries instead of single RWs. Portegies Zwart (2000) suggested as a result of their binary population synthesis calculations that up to 30% of the O-type RWs and possibly all B-type RWs are created via the BSS.

Stone (1991) investigated the space frequencies of O- and B-type RW stars based on observational data from different past studies for the Milky Way. The author suggested two approaches of determining the space frequency, the first one is the number of RWs divided by the total number of stars of that spectral type. This is the same as the RW fraction used in other sources. The second approach involves differences in the spatial distribution, i.e. scale heights, of these stars (based on Stone 1979). Stone (1991) tried

to correct for systematic errors made in other studies that led to widely differing results. According to their analysis, the “true” frequency corrected for geometric effects and with a velocity cut off of  $0 \text{ km s}^{-1}$  for O-stars is  $\sim 46\%$ , whereas it is only  $\sim 4\%$  for B-type stars. They indicated that only the BSS, which was suggested to predict that 40% of all O-stars turn into RWs, could be used to explain this high occurrence of O-star RWs. In addition, they also argued that the decrease in space frequency with decreasing mass is another outcome of the BSS. However, the low percentage of B-star RWs could be explained by either formation scenario. Combining the percentages for the two types of massive stars gives a value of  $\sim 20\%$  (Stone 1991).

Using simulations based on stars with solar metallicities, Eldridge et al. (2011) predicted that  $\sim 80\%$  of binaries are disrupted as a consequence of a SN. The number of O-type stars in the Milky Way that are fast companions formed via the BSS is thought to be 0.5% (RWs) and 2.2% (WWs), and 2.2% (RWs) and 7.1% (WWs) for B-type stars. In comparison, simulation results from Renzo et al. (2019) suggested that  $\sim 86\%$  of the binary systems that do not merge ( $\sim 78\%$  of all binaries) end up producing RWs and WWs as a result of the BSS, with approximately 75% of these ejected companions being MS-stars. The RW fraction for O-stars is  $\sim 0.5\%$ , whereas a higher percentage of  $\sim 10\%$  are O-star WWs. Both percentages from Renzo et al. (2019) are consistent with the results from Eldridge et al. (2011), when considering the large uncertainties (up to 8.5%) in the former results. But these values are much lower than what is suggested from observations ( $\sim 10\text{-}20\%$ , e.g. Stone 1991; Tetzlaff et al. 2011; Maíz Apellániz et al. 2018).

Compared to the higher percentage ( $\sim 67\%$ ) of all binary companions in the BSS turning into RW or WWs in Renzo et al. (2019), Boubert et al. (2017) calculated a lower percentage of  $\sim 32.5\%$  (“about a third”) of core-collapse SNe to release a RW companion. Lux et al. (2021) searched for RWs as a result of the BSS from all spectral types in SN remnants and found one certain candidate in their list of 12 SN remnants. However, they considered their findings to be consistent with the percentage range suggested in the two sets of simulations (Boubert et al. 2017; Renzo et al. 2019) and suggested follow-ups for a list of over 70 further candidates.

Unlike the initial suggestion of Blaauw (1961) regarding the non-existence of RW binaries, Portegies Zwart (2000) showed in their work using binary population synthesis calculations that the BSS can produce RW binaries. Their results suggested that 20% (B-stars) to 40% (O-stars) of the RWs remain bound to the SN remnant. However, only 1% of the RW companions are expected to be visible due to the stellar winds from the RWs obscuring the remnant emission, which is the likely explanation for the single star

claim in [Blaauw \(1961\)](#). The SN remnant remaining bound to the ejected star is possible in an asymmetric SN or when not enough mass is ejected in the SN, as was discussed previously in this section.

Very little is known about the percentage of low/intermediate-mass stars that are ejected as a result of a SN in the BSS. For this to occur the pre-SN binary requires a smaller (more unequal) mass ratio, as the primary star still needs to have a mass of at least  $\sim 8 M_{\odot}$  to be able to explode in a SN. However, this lack of known low/intermediate-mass RWs from the BSS is not due to the binary ratio potentially being skewed towards more equal-mass binaries. [Sana et al. \(2012\)](#) showed in their statistical analysis of a sample of Galactic massive O stars that the intrinsic mass ratio distribution in their high-mass spectroscopic binaries is thought to be flat, which is an equal likelihood for the companion to have any mass. [Duchêne & Kraus \(2013\)](#) combined results from three other studies ([Preibisch et al. 1999](#); [Duchêne et al. 2001](#); [Peter et al. 2012](#)) to show that higher-mass stars appear to favour lower mass companions.

The results from all these studies show that there is still further work to be done to better constrain the percentage of stars that become RWs via the BSS. For the higher-mass O-stars, the studies show a range of predicted RW fractions from as low as 0.5% up to 46% ([Stone 1991](#); [Eldridge et al. 2011](#); [Renzo et al. 2019](#)), whereas the predicted fraction for B-stars is  $< 10\%$  in these studies.

### What velocities can RWs formed via the BSS achieve?

In the BSS, the velocity of the ejected companion star depends on the kick velocity from the SN and the orbital parameters of the binary pre-SN. [Portegies Zwart \(2000\)](#) showed that the BSS produces RWs with more moderate maximum velocities of up to  $300 \text{ km s}^{-1}$  and that greater velocities are achieved for secondary stars with lower mass. [Przybilla et al. \(2008\)](#) found a RW star with a velocity of  $> 500 \text{ km s}^{-1}$  (HD 271791), which could be fast enough to escape the Milky Way. These authors suggested that a SN (or hypernova) of a WR-star collapsing into a BH is responsible for the companion to have reached these velocities.

[Silva & Napiwotzki \(2011\)](#) analysed over 170 early-type stars located in the Galactic halo (far away from star-forming regions) from literature to evaluate their potential RW nature. Of these stars about 100 stars are MS-stars. Using proper motion information, their orbits were traced back to calculate their ejection velocities. The authors found two groups of stars, one with ejection velocities of  $400\text{-}500 \text{ km s}^{-1}$ , the other with lower velocities of  $\sim 300 \text{ km s}^{-1}$ . They argued that the bimodality in the velocities and the



upper limit are evidence of the BSS. However, they did not provide direct evidence that these RWs are due to the BSS.

[Tauris \(2015\)](#) used Monte Carlo simulations to investigate the maximum possible run-away velocities from the BSS. Their results suggested possible upper velocities between  $\sim 550\text{-}1280 \text{ km s}^{-1}$  for the companion stars. The maximum velocities in their simulations depends strongly on the initial mass of the ejected secondary star, with the lowest mass star ( $0.9 M_{\odot}$ ) having the highest ejection velocity. To reach these velocities via the BSS requires asymmetric SNe to produce sufficiently high natal kicks.

[Renzo et al. \(2019\)](#) used simulations to show that most ejected stars from the BSS do not become RWs but rather slower WW stars ( $\sim 12 \text{ km s}^{-1}$ ). They found that most of the RWs produced in this way do not reach velocities  $> 60 \text{ km s}^{-1}$ , but they also noticed a minor peak in the distribution of the ejection velocities  $v$  between  $100 \text{ km s}^{-1} \lesssim v \lesssim 400 \text{ km s}^{-1}$ . The fastest RW in their simulations reached a velocity of  $\sim 1100 \text{ km s}^{-1}$ .

[Eldridge et al. \(2011\)](#) suggested that RW binaries from the BSS have lower velocities than single RW stars as the velocity kick from the SN would have otherwise disrupted the binary. Their simulations predict a small number of lower-velocity neutron stars (NSs) ( $< 200 \text{ km s}^{-1}$ ) that form after the SN of the high-mass companion star disrupts the binary, which originally remains bound after the SN of the primary star creates the NS. [Renzo et al. \(2019\)](#) provided velocity distributions for the binaries that remain bound post-SN and showed that NS + MS binaries in their simulations have a median systemic velocity of  $\sim 20 \text{ km s}^{-1}$  and black hole (BH) + MS binaries have higher median systemic velocities of  $\sim 50 \text{ km s}^{-1}$ . This can possibly produce RW binaries but depends on the initial centre-of-mass velocity as their systemic velocities are in the frame of the original centre-of-mass. [Dorigo Jones et al. \(2020\)](#) indicated that the mass ratio reversal right before the SN and a widening orbit prevent bound binaries to achieve RW-velocities post-SN.

[Renzo et al. \(2019\)](#) also showed in their simulations that it was possible for both binary stars involved in the BSS to be ejected, thereby creating high-velocity SN remnants, i.e. NSs or BHs (for more details, see the beginning of Section 1.5.1). The natal kick imparted on the companion star during the SN of the primary star can accelerate the compact remnant to extremely high velocities, if the SN is asymmetric (e.g. [Tetzlaff et al. 2010](#); [Tauris 2015](#)). Observations have found NSs with velocities exceeding  $1000 \text{ km s}^{-1}$ , e.g. IGR J11014-6103, B2011+38, B2224+65 and B1508+55 (e.g. [Hobbs et al. 2005](#); [Chatterjee et al. 2005](#); [Tomsick et al. 2012](#); [Pavan et al. 2016](#)).

The results from these studies show that the upper velocities for RWs formed via the

BSS are likely to be in the region of  $\leq 300\text{-}400 \text{ km s}^{-1}$ . However, the BSS can produce RW stars (companions and SN remnants) that are fast enough to potentially escape the Milky Way in special circumstances. I will discuss these very fast RWs separately in Section 1.5.4.

## 1.5.2 Dynamical ejection scenario - DES

Blaauw & Morgan (1953, 1954) discovered two high-mass, high-velocity stars travelling in almost opposite direction with virtually the same velocity suggesting the creation of these RWs in the same event. Almost 50 years later, Hoogerwerf et al. (2001) showed that these two high-velocity stars (AE Aur and  $\mu$  Col) were highly likely ejected as the result of a dynamical interaction  $\sim 2.5$  Myr ago. This kind of interaction forms the second RW formation mechanism (DES), which does not involve SNe and can therefore explain ejections from very young star-forming regions.

Poveda et al. (1967) suggested that RWs can be created due to dynamical interactions during the gravitational collapse of a small cluster of massive stars, when the cluster is at its minimum radius. The 54 small clusters in their calculations are composed of 5 or 6 massive O-type stars. The authors found that most RWs (with velocities  $> 35 \text{ km s}^{-1}$ ) in their analysis are single stars but also found close RW binaries with separations in the order of a few au. For about half of the RWs in their analysis, another star is ejected in almost opposite direction often with higher velocities and several of these “opposite” ejections reach the RW-velocity themselves. These ejections are the result of dynamical interactions where close encounters between a single star and a binary (e.g. Hut & Bahcall 1983; Hut 1983, 1993) or a binary-binary interaction (e.g. Hoffer 1983; Mikkola 1983, 1984a,b) can lead to ejection of one or more of these stars.

Binaries can be divided into different “classes” based on their binding energy  $E_{bin}$  in relation to the average mass of a star  $m$  and the velocity dispersion  $\sigma$  of the region they are located in (e.g. a star-forming region). The seminal works by Heggie (1975) and Hills (1975a,b) resulted in the “Heggie-Hills”-law, which states that hard binaries get harder and soft binaries get softer after interactions with a single perturbing star. Hard binaries are characterised by  $|E_{bin}| m\sigma^2 > 1$ , whereas for soft binaries this is  $|E_{bin}| m\sigma^2 < 1$ .

When hard binaries encounter another star, this interaction transfers some of the energy from the binary to the passing star, which increases the kinetic energy of the single star. In turn the energy of the binary is decreased, therefore increasing its binding energy  $E_{bin}$ , i.e. it has become harder, possibly due to a decrease of the binary separation (semimajor axis). Another possibility to increase the binding energy is the exchange of

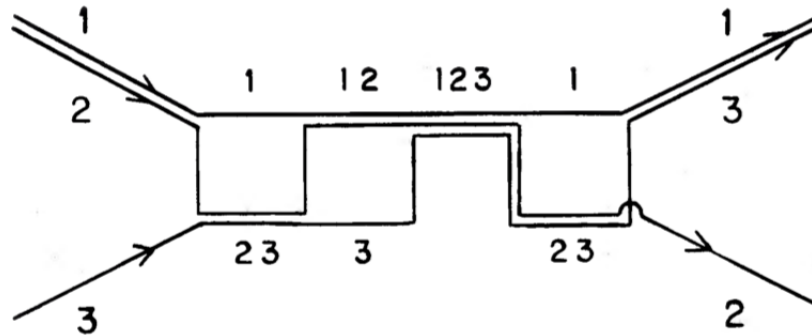


Figure 1.15: Example of the dynamical interactions between a hard binary (star 1 and 2) and a single star 3 (all equal mass) in a scattering experiment. The interactions show repeated swapping of the binary stars, the brief creation of an unstable triple system and the final result of the exchange of one of the binary stars with the incoming single star. Figure credit: extracted from Fig. 3 in [Hut & Bahcall \(1983\)](#).

one of the binary components with the passing star. This can occur when the single star has a higher mass than one of the binary stars. In these interactions often the least massive star will end up as a single (possibly) ejected star (e.g. [Heggie 1975](#); [Hills 1975a,b](#); [Heggie et al. 1996](#)).

When soft binaries encounter another star, the interactions transfer energy from the passing star onto the binary, thereby reducing the binding energy by increasing its total energy. This can lead to an increase of their separation (semimajor axis). If the velocity of the single star is high enough it can transfer so much energy to the binary leading the binary being ionised, i.e. become unbound. For this to happen, the velocity of the single star has to be larger than the critical velocity  $v_c$  (Eq. 2.4 from [Hut & Bahcall 1983](#)):

$$v_c = \sqrt{\frac{G m_1 m_2 (m_1 + m_2 + m_3)}{m_3 (m_1 + m_2) a}}, \quad (1.10)$$

where  $G$  is the gravitational constant,  $m_1$  and  $m_2$  are the masses of the binary stars,  $a$  the binary semimajor axis and  $m_3$  is the mass of the single perturbing star.

Fig. 1.15 shows the result of scattering experiments by [Hut & Bahcall \(1983\)](#) involving a binary and a single star all of equal mass. During the interactions shown here, the binary components are swapped multiple times, binary orbital parameters will likely change dramatically and the components even briefly form an unstable multiple system.

[Fregeau et al. \(2004\)](#) studied the possible results of binary-single star and binary-binary interactions based on a large number of numerical scattering experiments that

$n_{\text{coll}}$	Symbol	Description
0	[••] •	Preservation or exchange
0	•••	Ionization
0	[[••] •]	Stable hierarchical triple
1	[•:••]	Binary containing a two-star merger
1	•:••	Two-star merger and single star
2	•:•:•	Three-star merger

Figure 1.16: Possible outcomes of interactions between a binary and a single star. Bound binary/multiple systems are denoted with [], whereas mergers as a result of collisions (possible multiple collisions as indicated by  $n_{\text{coll}}$ ) are denoted by a colon. In bound systems the components can either be preserved or exchanged. Ionisation refers to the result of having three unbound stars. Figure credit: Table 1 from [Fregeau et al. \(2004\)](#).

can even involve collisions of the stars leading to merger products. Fig. 1.16 from [Fregeau et al. \(2004\)](#) shows the outcomes of binary-single star interactions based on the different number of collisions. These interactions can lead to preservation of the original binary or exchange of one the binary components with the single star, the formation of a stable triple system, as well as the ionisation of all the binary creating three single stars. When collisions are involved, binaries can be created with merger products or merged single stars [Fregeau et al. \(2004\)](#).

Binary-binary interactions can lead to even more combinations than the binary-single star case. Fig. 1.17 from [Fregeau et al. \(2004\)](#) provides an overview of the possible outcomes from the interaction of two binaries sorted by the number of collisions/mergers that occur. The general outcomes are the same as in the binary-single star interactions, i.e. preservation, exchange, ionisation, merger and stable multiple systems. Many of the interactions (binary-single star/binary-binary) can lead to the creation of single or binary RW/WWs depending on the parameters of the stars involved ([Fregeau et al. 2004](#)).

The DES formation mechanism did not gain traction for many decades as the use of a “single” formation mechanism was preferred. Only in the 1990s, when more powerful  $N$ -body simulations and binary scattering experiments were able to better predict the properties of RWs via the DES was a second formation mechanism accepted ([Leonard 1990](#)).

$n_{\text{coll}}$	Symbol	Abbreviation	Description
0	[••] [••]		Preservation or exchange
0	[••]••		Single ionization
0	••••		Full ionization
0	[[••]•]•		Stable hierarchical triple and single star
0	[[[••]•]•]•		Stable hierarchical quadruple
0	[[••] [••]]		Stable quadruple composed of two binaries
1	[••]••	(SS)D	Binary and two-star merger
1	[•••]•	(DS)S	Single star and binary containing two-star merger
1	•:•••	DSS	Two-star merger and two single stars
1	[[•:••]•]	((DS)S)	Stable hierarchical triple with two-star merger in inner binary
1	[[••]•:•]	((SS)D)	Stable hierarchical triple with two-star merger in outer binary
2	[••••]	(DD)	Binary composed of two two-star mergers
2	[•••••]	(TS)	Binary containing a three-star merger
2	•:•••	DD	Two two-star mergers
2	••:••	TS	Three-star merger and single star
3	••:•••	Q	Four-star merger

Figure 1.17: Possible outcomes of interactions between two binaries. Bound binary/multiple systems are denoted by  $[[ ]]$ , whereas mergers as a result of collisions (possible multiple collisions as indicated by  $n_{\text{coll}}$ ) are denoted by a colon. In bound systems the components can either be preserved or exchanged. Ionisation refers to the result of having binaries become unbound stars. Figure credit: Table 2 from [Fregeau et al. \(2004\)](#).

### What percentage of stars become RWs via the DES?

In the analysis in [Poveda et al. \(1967\)](#), the authors found that up to 15% of the O-stars in their 54 clusters (each composed of 5-6 OB-stars) turn into RWs, which are 38 RWs. They also found examples where their small clusters lose more than half of the stars leaving only a close binary behind. This was likely an effect of their choice of a 6-star cluster. In their work using  $N$ -body simulations, [Leonard & Duncan \(1990\)](#) showed that about 10% of RWs from their simulations are binaries. However, the authors did not provide RW fractions for their simulations and only suggested that cluster with 100-300 members are most effective at producing one RW within 25 crossing times.

[Fujii & Portegies Zwart \(2011\)](#) used  $N$ -body simulations of massive clusters with masses between  $10^3 - 10^5 M_{\odot}$ . They suggested that a massive “bully” binary that formed during the initial collapse of the clusters is responsible for most ejections. The least massive clusters in their set of simulations ( $\sim 6000 M_{\odot}$ ) produce the largest fraction of RWs. However, the overall fraction is still low at only 1%. The fraction increases for stars with higher mass and reaches a maximum of around 10% for stars with 40-50  $M_{\odot}$ .

[Perets & Šubr \(2012\)](#) used  $N$ -body simulations of massive clusters with total masses of 5000  $M_{\odot}$  and  $\sim 7400$  stars to determine the properties of RWs formed via the DES. They defined escapers based on their distance (minimum 5 pc) to the birth cluster after 2.7 Myr and used  $20 \text{ km s}^{-1}$  as their boundary. They found a very low RW fraction

(below  $\sim 2\%$ ) for stars below a mass of  $4 M_{\odot}$ . This is possibly caused by the initially single star status of these stars in their simulations and the authors suggested that a large population of low-mass RWs should exist. The RW fraction increases with mass with around 22% of O-stars with a mass of around  $\sim 30 M_{\odot}$  being ejected from their birth clusters. This fraction is about 3 times greater than the percentage of ejected B-stars. The authors also showed that the DES can create non-compact RW binaries. They concluded that the rate of cluster stars that turn into RWs strongly correlates with higher stellar masses and increasing cluster density.

Oh & Kroupa (2016) studied the effects of different initial conditions in young moderately massive clusters with a mass of around  $10^{3.5} M_{\odot}$  ( $\sim 3200 M_{\odot}$ ) on ejected stars via the DES using  $N$ -body simulations. They varied the initial density, binary population and the initial mass segregation. Like Perets & Šubr (2012), they found that the ejection fractions increase with stellar mass. Oh & Kroupa (2016) found that the initial stellar density appears to be the dominant factor on the RW fraction and increases further when the primordial binary fraction are high as well. With very dense initial conditions (half-mass radius = 0.1 pc), up to 50% of all O-star systems can be ejected, which dramatically reduces to at most 4.5% in initially less dense clusters (half-mass radius = 0.8 pc). They found that the ejections show a peak at 1 Myr or earlier in their 3 Myr simulations, which possibly coincides with the minimum radius of the clusters as suggested by Poveda et al. (1967). In the Oh & Kroupa (2016) simulations, massive O-star binary systems are also ejected. However, the fraction is lower than the original binary fraction in the cluster, as was also found by Fujii & Portegies Zwart (2011).

### What velocities can RWs formed via the DES achieve?

In the seminal paper of Poveda et al. (1967), the 38 ejected RWs in their analysis reach upper velocities of  $185 \text{ km s}^{-1}$ . In the work of Leonard & Duncan (1990), RWs reach a maximum ejection velocity of  $\lesssim 200 \text{ km s}^{-1}$  and the authors also confirmed that lower-mass stars reach higher velocities. Shortly thereafter, further simulations by Leonard (1991) increased the possible maximum ejection velocity via the DES to  $\lesssim 400 \text{ km s}^{-1}$ . They also highlighted that in rare cases velocities up to  $\lesssim 1000 \text{ km s}^{-1}$  are possible, if the initial conditions are tailored to achieve this outcome.

Gvaramadze et al. (2009) performed three-body scattering experiments of a massive, tight binary with a massive star ( $> 50 M_{\odot}$ ) to determine the maximum achievable velocity and found that a small percentage of these encounters result in ejection velocities  $\geq 300 - 400 \text{ km s}^{-1}$ , while in rare cases even higher velocities can be reached with RWs

capable of escaping the Milky Way.

In their  $N$ -body simulation work, [Perets & Šubr \(2012\)](#) showed that the DES can produce RWs with velocities  $> 500 \text{ km s}^{-1}$ . However, to reach velocities above  $\sim 300 \text{ km s}^{-1}$  requires very close encounters with binaries only possible in the dense central cores of massive clusters. Even then, the average number of these very fast RWs is less than 0.5 stars per cluster. These higher ejection velocities are only reached for single stars, even though 99% of ejected single stars do not exceed  $200 \text{ km s}^{-1}$ . For binaries the maximum ejection velocities do not exceed  $150 \text{ km s}^{-1}$  in these simulations, which is still considered the upper velocity for DES RW binaries. The binary fraction of RWs created via the DES decreases with ejection velocity, falling from  $\sim 40\%$  at velocities of  $30 \text{ km s}^{-1}$  to  $\sim 10\%$  at  $150 \text{ km s}^{-1}$ .

Based on these studies, it was shown that the ejection velocities that can be reached via the DES lie in a similar range than the BSS with most stars reaching maximum velocities of  $300\text{-}400 \text{ km s}^{-1}$ . Velocities above this range are possible and they can exceed the escape velocity of the Milky Way. I discuss these very fast RW stars in Section 1.5.4.

### 1.5.3 Can we distinguish between the formation scenarios?

The previous sections show that the BSS and DES can accelerate young stars to very similar velocities making it difficult to distinguish between the scenarios using velocity information alone. The predicted velocity distributions for both scenarios are also very similar, albeit with an enhanced high-velocity tail in the DES ([Silva & Napiwotzki 2011](#)). It is thought that the velocities in the DES are generally higher than those that can be produced in the BSS, even though the fastest RWs are expected to be the result of the BSS (e.g. [Renzo et al. 2019](#); [Dorigo Jones et al. 2020](#); [Raddi et al. 2021](#)).

Unlike the DES, which can occur immediately after the formation of a young star-forming region, any contribution to the RW population from the BSS can only start after a few Myr, when the most massive stars reach the end of their MS life. When young stars can be traced back to their birth regions and these regions are young enough to not have been subjected to a SN, then the DES is the most likely mechanism responsible for these ejections. For RW stars that are no longer close to any star-forming region, it might be impossible to find their origin and other methods based on observational characteristics could be used to identify the formation mechanisms.

[Renzo et al. \(2019\)](#) argued that a majority of the massive companions, as well as a large fraction of lower-mass companions that are ejected via the BSS are not standard single MS stars. Even before the binary disruption occurs, they accrete mass from the

more massive primary star. This mass transfer can lead to a rejuvenation of the accreting star by causing a growth of the convective core (e.g. [Schneider et al. 2016](#)). This can possibly explain RWs that, when traced back, show flight times that are longer than their apparent stellar age (e.g. [Perets 2009](#); [Gvaramadze et al. 2019](#); [Gvaramadze 2020](#)).

[Blaauw \(1993\)](#) suggested that RWs via the BSS can be identified via their rapid rotational velocities and higher helium abundances. It is thought that high rotational velocities in the BSS originate from a correlation of the ejection velocity to the orbital velocity, which in turn correlates closely with the rotational velocity in close binaries (e.g. [McEvoy et al. 2017](#)). However, [Leonard \(1995\)](#) pointed out that both characteristics are also present in RWs from the DES. [Wolff et al. \(2007\)](#) showed that B-stars that form in dense environments rotate faster than those from less dense regions, therefore making DES RWs also fast rotators. [Bragança et al. \(2012\)](#) confirmed the results that higher rotation velocities in RWs are likely the result of the DES. [Dorigo Jones et al. \(2020\)](#) on the other hand suggested in their analysis of OB-RWs in the SMC that the DES corresponds to high transverse velocities but slower projected rotational velocities, whereas the opposite is true for RWs formed via the BSS.

The abundance argument stems from the idea that the BSS could provide some imprint of the SN ejecta in the surface composition of the RW. The DES is more likely to leave the star's surface unchanged and therefore would be identified as MS stars (e.g. [Eldridge et al. 2011](#)). [McEvoy et al. \(2017\)](#) used the lists of RWs by [Silva & Napiwotzki \(2011\)](#) and [Tetzlaff et al. \(2011\)](#) to investigate if the N, Mg, and Si abundances show any differences amongst known RWs that could be attributed to either the BSS or the DES. However, they found no anomalies in the abundances and concluded that neither formation mechanism is able to change the abundances. Therefore, the surface composition represents the chemistry found in the birthplaces of these RWs ([McEvoy et al. 2017](#)).

[Dorigo Jones et al. \(2020\)](#) compared the two ejection mechanisms (BSS and DES) for OB-RWs formed in the SMC. They used a sample of  $\sim 300$  field OB stars and found that approximately 70% of these field stars are RWs. For SMC metallicity, the binary population synthesis models used by [Renzo et al. \(2019\)](#) predicts a split of 80:20 between the DES and the BSS for their sample. The [Dorigo Jones et al. \(2020\)](#) result indicates a factor 2-3 between the dominating DES and the BSS. [Dorigo Jones et al. \(2020\)](#) pointed to a further complication in distinguishing between the scenarios. They suggested that two-step ejections are included in the RWs from the BSS and that it is likely that these ejections dominate the BSS population.

There is a subset of the BSS RW population, where the RWs appear to be the



result of two-step ejections (Pflamm-Altenburg & Kroupa 2010). The first step is the ejection of a binary via the DES. The primary star in the RW binary then evolves and explodes in a SN disrupting the binary, creating a single RW star. This single RW star can potentially bear observational imprints of the preceding binary evolution and SN explosion. However, its ejection origin and velocity will at least partly be due to the ejection via the DES. The authors suggested that these ejections make up a considerable fraction of the BSS RWs. They can possibly explain field OB-stars that are thought to have formed in isolation (de Wit et al. 2005) but are instead two-step RWs that cannot be traced back to their origin. Eldridge et al. (2011) pointed out that RWs especially those with velocities  $> 200 \text{ km s}^{-1}$  can possibly have formed via this two-step process.

This brief discussion highlights the difficulties in being able to distinguish between the two RW formation scenarios in observations. In the next section, I will discuss two additional populations of high-velocity stars, which further complicate the situation.

#### 1.5.4 Hyper-runaway and Hyper-velocity stars

For several years, there were speculations if stars could move even faster than the RW stars. Hills (1988) predicted that it is possible for stars to reach velocities that can unbind them from a galaxy in three-body exchange interactions involving a supermassive BH (SMBH). The mechanism involved became known as the Hills mechanism and these stars were termed hypervelocity stars. In this ejection scenario a binary approaches a SMBH and can be tidally disrupted. One star then becomes bound to the SMBH and the other is ejected with a velocity that can exceed the Galactic escape velocity (Hills 1988). The Galactic escape velocity decreases with distance from the Galactic centre. At the distance of the Sun ( $\sim 8 \text{ kpc}$  from the centre), the escape velocity is estimated to be between  $500\text{-}600 \text{ km s}^{-1}$ , depending on the model used for the Galactic potential. The escape velocity drops to values  $\leq 400 \text{ km s}^{-1}$  beyond a distance of  $50 \text{ kpc}$  (e.g. Piffi et al. 2014; Williams et al. 2017).

When it became accepted that there were SMBHs located at the centre of many spiral galaxies, it was only a matter of time until the first HVSs were discovered, as HVSs are an automatic outcome of the Hills mechanism. The first HVS was discovered by chance during a radial velocity survey of stars in the outer Milky Way halo (Brown et al. 2005), followed by two further discoveries in the same year (Hirsch et al. 2005; Edelmann et al. 2005). Additional HVSs were discovered following a systematic search, which also uncovered a population of HVSs that were possibly still bound to the Galaxy (Brown et al. 2007a,b).

As I described in the previous sections, it is possible for both the DES and BSS to accelerate stars to velocities that can unbind them from the Galaxy. However, these stars are not created via the Hills mechanism and are therefore technically not HVSs. Heber et al. (2008) discovered the first of these fast, non-HVSs by showing that the star in question do not originate from the Galactic centre but are ejected from the outer disc. The stars were then termed hyper-runaway stars in the same year by Przybilla et al. (2008).

Brown (2015) suggested that HRWs with velocities of  $\sim 600 \text{ km s}^{-1}$  can only be explained by the DES. However, Evans et al. (2020) concluded that the classical BSS can also contribute, albeit not significantly, to the HRW population. Recently, low-mass HRWs were discovered that could be the result of a thermonuclear type Ia SNe, instead of the core-collapse type II SNe associated with the BSS (Justham et al. 2009). Raddi et al. (2018) investigated the low-luminosity unbound HRW star LP 40–365, which has been proposed to be the remnant of a type Iax SNe with an estimated ejection velocity of  $\sim 600 \text{ km s}^{-1}$ . Neunteufel (2020) analysed the velocity limits possible in this type Ia SN - HRW scenario and showed that velocities in excess of  $1000 \text{ km s}^{-1}$  are possible providing an alternative formation mechanism for HRWs.

HRWs originating from the disc will achieve the highest velocities if they are ejected in the direction of the rotation of the Galaxy (e.g. Brown 2015). Predictions from simulations indicate that 90% of HRWs should be found at Galactic latitudes  $< 25^\circ$  (Bromley et al. 2009) and that only one HRW per 100 HVS should be found in the halo (e.g. Brown 2015, and references therein). Using *Gaia* EDR3, Irrgang et al. (2021) showed that for 30 extreme velocity disc-runaways located in the halo only three have an origin consistent with the Galactic centre and are located at high Galactic latitudes instead of smaller latitudes.

Lu & Naoz (2019) showed that a SN of one of the stars in the inner binary of a hierarchical triple system with a tertiary SMBH component could disrupt the triple system and create a hypervelocity binary via the BSS. Gao et al. (2019) followed up from this study and investigated if it is theoretically possible to produce very fast binary RWs via the BSS that can reach the Galactic escape speed. The authors analysed the velocity distributions of binary RWs that are ejected after the tertiary star in a triple system explodes as a core-collapse SN. They showed that it is possible to generate binaries travelling at velocities  $> 350 \text{ km s}^{-1}$ . Németh et al. (2016) reported the discovery of a potential binary companion to a hot subdwarf HVS candidate (PB3877). They suggested that the binary could have an origin outside of the Milky Way from a disrupted dwarf

satellite galaxy or from the Milky Way halo.

The *Gaia* mission has revolutionised the search for RWs, HRWs and HVSs especially since the second data release, and future data releases are expected to further enlarge the number of observed fast stars in our Galaxy (e.g. [Boubert et al. 2018](#); [Bromley et al. 2018](#); [Irrgang et al. 2018](#); [Lennon et al. 2018](#); [Marchetti et al. 2019](#); [Raddi et al. 2018](#); [Brown et al. 2018](#); [Rate & Crowther 2020](#); [Li et al. 2021](#)). This will allow us to better constrain the contributions from the different formation mechanisms and to better distinguish between single and binary RWs, HRWs and HVSs.

## 1.6 Work in this thesis

The work in this thesis focuses on whether young stars that have been ejected from their birth star-forming regions during their early dynamical evolution can be used to constrain the initial conditions of these regions. For this work, I use a combination of  $N$ -body simulations and observations from the *Gaia* telescope ([Gaia Collaboration et al. 2018a](#)). I establish the suitability of ejected stars in constraining the initial density, spatial and kinematic substructure.

In the next two chapters, I provide an overview of the methods that form the basis of the science presented in Chapters 4, 5, 6 and 7. First, I discuss the physics involved in  $N$ -body simulations, the software I use and how I set up the simulations. I then turn to the *Gaia* observations, where I briefly describe the telescope itself, followed by the method I use to trace back stars to young star-forming regions.

In Chapter 4, I use  $N$ -body simulations with differing initial conditions to investigate if the number and velocity distributions of unbound stars allow me to place constraints on the initial density and velocity structure in star-forming regions. I aim to make predictions for observations of fast unbound stars from young star-forming regions that can be probed with *Gaia* DR2. In Chapter 5, I use  $N$ -body simulations to predict the number of RW and WW stars for an ONC-like star-forming region. I then use *Gaia* DR2 observations to search for RW and WW stars around the ONC. In Chapter 6, I search for evidence of circumstellar discs around recently ejected stars from the ONC found in *Gaia* DR2, as well as discs around future and past visitors to determine whether protoplanetary systems can experience more than one dense stellar environment. In Chapter 7, I use *Gaia* DR2 and EDR3 observations to search for RW and WW stars around another young star-forming region NGC 2264. I then compare these numbers to  $N$ -body simulations with different initial conditions to constrain the initial conditions of NGC 2264.

## Chapter 2

### N-body simulations

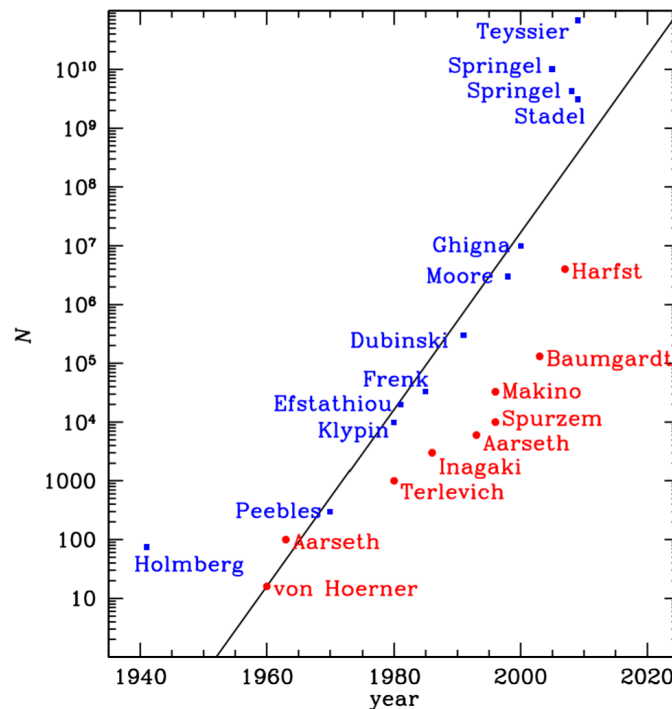


Figure 2.1: Evolution of particle numbers in collisional (red) and collisionless (blue)  $N$ -body simulations starting with von Hoerner (1963) in the 1960s. The black line is based on Moore’s Law (Moore 1965), if costs scale  $\propto N$ . Figure credit: Fig. 1 from Dehnen & Read (2011).

One of the key forces governing the interactions of stars in star-forming regions is gravity, which affects the position, velocity and acceleration. The difficulty of predicting the motion of  $N$  individual stars in these regions under the effect of gravity is described in the  $N$ -body problem. When up to three bodies are involved in these gravitational interactions, the problem can be solved analytically. However, when  $N$  increases, numerical methods such as  $N$ -body calculations/simulations are used (Aarseth 2003).

$N$ -body simulations have been used for decades for modelling smaller systems like star clusters up to the size of galaxies in cosmological simulations. With the increase in computing capabilities,  $N$ -body calculations quickly evolved into increasingly complex  $N$ -body simulations. In the 1960s, the first  $N$ -body calculations started with only  $N = 16$  particles (later increasing to  $N = 25$ , von Hoerner 1963) and nowadays the largest collisionless cosmological  $N$ -body simulations have reached particle numbers of  $> 4$  trillion particles with the numbers expected to increase further (Cheng et al. 2020). Fig. 2.1 shows the evolution of particle numbers  $N$  in simulations from the 1960s to  $\sim 2010$  (Dehnen & Read 2011) in collisional and collisionless simulations.

Two different general types of  $N$ -body codes are used, collisional and collisionless. Collisional codes are used for systems where the relaxation time (time it takes for a star to forget its initial conditions, e.g. energy) of the system is smaller than the time covered (duration) in the  $N$ -body simulations. These types of codes are used to model the evolution of older globular clusters and young star clusters, which is the type I use for my simulations. In these collisional simulations, the motion of each particle and interactions with other particles, such as those of binary or other multiple systems, are followed with direct summation techniques. Collisionless codes on the other hand are used for systems where the relaxation time of the system is much larger than the time covered in the  $N$ -body simulations (e.g. Binney & Tremaine 2008). These collisionless codes model systems with very large number of particles and examples are galaxy or dark matter halo simulations.

## 2.1 $N$ -body interactions in simulations

In  $N$ -body simulations stars, planets and other bodies are represented by point-like particles, which are assigned a mass. The Newtonian equation of motion is then used to determine the gravitational forces between all particles in the simulation. The force  $\mathbf{F}_i$  acting on particle  $i$  and its associated equation of motion, i.e. the acceleration  $\mathbf{a}_i$ , of a particle  $i$  due to another particle  $j$  for a total number of  $N$  particles are described by:

$$\mathbf{F}_i = m_i \mathbf{a}_i \tag{2.1}$$

$$\mathbf{a}_i = -G \sum_{j=1, j \neq i}^N m_j \frac{\mathbf{x}_i - \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^3}. \tag{2.2}$$

In these equations,  $m_i$  and  $m_j$  are the masses of particle  $i$  and  $j$ ,  $G$  is the gravitational constant and  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are the position vectors of the two particles  $i$  and  $j$ . The first time derivative  $\dot{\mathbf{a}}_i$ , which is also called ‘jerk’  $\mathbf{j}$  is described by:

$$\dot{\mathbf{a}}_i = -G \sum_{j=1, j \neq i}^N m_j \frac{(\mathbf{x}_i - \mathbf{x}_j)^2 (\mathbf{v}_i - \mathbf{v}_j) - 3(\mathbf{x}_i - \mathbf{x}_j) ((\mathbf{x}_i - \mathbf{x}_j) \cdot (\mathbf{v}_i - \mathbf{v}_j))}{|\mathbf{x}_i - \mathbf{x}_j|^5}, \tag{2.3}$$

where  $\mathbf{v}_i$  and  $\mathbf{v}_j$  are the velocity vectors of the two particles  $i$  and  $j$ , and all other terms are as in equation 2.2.

### 2.1.1 Different $N$ -body techniques

## Direct summation

In collisional simulations, a system with a number of particles  $N_*$  is evolved where  $N_*$  is equal to the number of particles  $N$  for which the equation of motion is calculated (see Eq. 2.2). In these systems, the gravitational interactions between the particles are integral to the dynamical evolution and need to be fully calculated. Direct summation is used and in this approach the forces acting on particles in the simulation are calculated by summing up the contributions from all other particles (e.g. [Aarseth 1999](#); [Spurzem 1999](#); [Portegies Zwart et al. 2001](#); [Dehnen & Read 2011](#)).

While direct summation is a highly accurate technique, it comes at a high computational cost (and time). The number of calculations required for each particle is  $N(N - 1)/2$ , so it has a quadratic time complexity  $O$  and scales as  $O(N^2)$ . This means that the time to execute calculations is proportional to the square of the total number of particles  $N$  in the simulations. This applies to every single timestep in the simulations, therefore further increasing the complexity to scale with  $\sim O(N^3)$  (for an overview of the impact of different timestep approaches on complexity, see [Wang et al. 2015](#)). For the time integration of interactions in the simulations different integrators can be used. One of the common methods used is the Hermite scheme, which I explain in more detail in Section 2.2 (e.g. [Aarseth 1999](#); [Spurzem 1999](#); [Portegies Zwart et al. 2001](#); [Dehnen & Read 2011](#)).

For many decades, it appeared difficult to reach a million particles with direct N-body codes (“gravitational million-body problem”, [Heggie & Hut 2003](#)). However, direct  $N$ -body simulations have now reached particle numbers of  $\sim 10^6$  using special purpose hardware, graphical processing units and highly efficient parallelised software (e.g. [Makino et al. 1997](#); [Harfst et al. 2007](#); [Gaburov et al. 2009](#); [Wang et al. 2015, 2016](#)).

While these particle numbers are sufficient for many collisional simulations, they are not sufficient for simulations of large-scale cosmological systems, where the required particle numbers are much larger. In these larger systems, the stellar densities are lower and interactions between two individual particles are rarer. The interactions do not contribute to the evolution of the large-scale systems that these simulations are tasked to solve (e.g. [Athanassoula et al. 2000](#)). These systems can therefore be simulated using collisionless simulations. In these simulations, the size of the system  $N_*$  that is evolved in the simulation is larger than the number of particles  $N$  for which the equation of motion is calculated. They do not require the full calculation of the forces acting on all particles but use approximations, thereby providing an estimate of the forces in the system (e.g. [Binney & Tremaine 2008](#); [Dehnen & Read 2011](#)). In these collisionless



simulations, softening is commonly used, which I explain briefly following this section.

I now briefly describe three methods that have been used for these collisionless systems and comment on their complexity.

### Tree code

The tree code is a hierarchical approach, where particles are grouped together based on their spatial distribution and the force of each group is then evaluated. In the method first suggested by [Barnes & Hut \(1986\)](#), this grouping occurs by subdividing space again and again following the placement of a root cube that encompasses the whole system. The forces of these smaller cubic cells are then calculated using a truncated multipole expansion related to the centre of mass of the cells. This method requires fewer calculations for the force of one particle, however, it is also less accurate than direct summation. The time complexity of this method scales as  $O(N \log N)$ .

### Fast Multipole Method

The Fast Multipole Method (FMM, [Greengard & Rokhlin 1987, 1997](#)) is also a hierarchical approach as is the previously described [Barnes & Hut \(1986\)](#) tree code. In contrast to this approach, FMM makes use of the fact that particles that are located close to each other are affected in a similar way by more distant particle groups. The forces are calculated using multiple expansion, where groups in close proximity are considered to be one group. It is suggested that the complexity of this method scales as  $O(N)$ . While this method is more commonly used for collisionless simulations, it has recently been used as an approach to solve collisional systems with  $N > 10^6$  ([Mukherjee et al. 2021](#)).

### Particle-mesh codes

Particle-mesh (PM) codes are yet another method to compute the forces in collisional systems. Unlike the previous two methods, it is a grid-based method using the fast Fourier transform. In these codes, the particles in the system do not interact with each other directly, so no forces between these particles are calculated. Instead a grid with density values is applied to the system and the forces are calculated based on where a particle is located within this grid. It is the fastest of the methods presented here and scales as  $O(n_{\text{grid}} \ln n_{\text{grid}})$  (e.g. [Trenti & Hut 2008](#); [Dehnen & Read 2011](#)). [Cheng et al. \(2020\)](#) created a two-level PM code using compression techniques to run a collisionless, cosmological simulation with  $> 10^{12}$  particles, which the authors claim was the largest one run at that time.

## Monte-Carlo and Fokker-Planck methods

Two alternatives to  $N$ -body simulations are Monte-Carlo (MC) and Fokker-Planck methods. Unlike direct  $N$ -body simulations, these two methods do not integrate the interactions/orbits of particles directly but apply a statistical approach. In the MC methods, the particles are replaced by spherical shells. It is assumed that they are in a dynamical equilibrium in a smooth, spherical potential and that the large-scale evolution of the system is driven, e.g. by stellar encounters, moving the system from one equilibrium to another (e.g. Hénon 1971, 1975; Giersz 1998). Fokker-Planck methods are very similar to the MC method in concept. However, instead of using spherical shells, this approach uses distribution functions (Cohn 1979, 1980; Spurzem & Takahashi 1995).

While these methods still use approximations, they can be used for collisional simulations, when larger particle numbers are required. Rodriguez et al. (2022) recently used a Cluster MC method to simulate the evolution of a Plummer (1911) sphere with  $10^8$  particles up to core collapse. In comparison to direct summation, MC methods scale linearly with the number of particles for the calculation of orbits and dynamical evolution (e.g. Trenti & Hut 2008; Rodriguez et al. 2016, 2018). However, Rodriguez et al. (2018); Plummer (1911) suggested that some aspects inherent in their MC method made it scale with  $O(N \log N)$ , which is still much faster than the  $O(N^3)$  scaling of direct summation.

The methods described above are examples of the different types that are currently used for collisional and collisionless simulations. There are also many combinations and hybrids of these methods being explored, with the aim of increasing the particle numbers, while keeping computational costs at an acceptable level (for examples of these methods, see Trenti & Hut 2008).

### 2.1.2 Softening in $N$ -body simulations

The force acting between particles depends on the value of their separation. The smaller their separation, the larger the force. When point-like particles (no physical radius) in an  $N$ -body simulation come too close to each other, their small separation can cause the resulting force to become increasingly large (goes towards infinity). This can lead to unwanted effects, such as very high velocities and spurious scattering.

Softening is used to prevent this numerical divergence ( $1/|x_i - x_j| \rightarrow \infty$ ) from happening by adding a softening length value  $\epsilon$ . This is essentially like introducing a physical radius to the particles in the simulations (Trenti & Hut 2008). The denominator in the

equation of motion in 2.2 is adjusted by adding softening to turn into:

$$\mathbf{a}_i = - \sum_{j=1, j \neq i}^N G m_j \frac{\mathbf{x}_i - \mathbf{x}_j}{(|\mathbf{x}_i - \mathbf{x}_j|^2 + \epsilon^2)^{3/2}}. \quad (2.4)$$

It was first introduced by [Aarseth \(1963\)](#) using a [Plummer \(1911\)](#) potential, where each particle (or its potential) was replaced by a small Plummer model (or Plummer potential). The optimal value of the softening length has been debated for decades since and should in general be kept as small as possible to reduce systematic biases ([Dyer & Ip 1993](#); [Merritt 1996](#); [Athanassoula et al. 2000](#); [Dehnen 2001](#); [Das et al. 2021](#)).

Softening is used in collisionless simulations. In addition to preventing divergence, it reduces the error introduced to these simulations by the use of approximate solutions. In contrast, my simulations are collisional, where the forces in close interactions have to be calculated without approximations using direct summation (e.g. [Dehnen & Read 2011](#)).

## 2.2 The Starlab software

I use the **Starlab** software environment in my work for simulations of the evolution of star-forming regions with different initial conditions ([Portegies Zwart et al. 2001](#)). **Starlab** consists of different programmes that are linked together to generate, evolve and finally analyse the interactions in collisional systems. The main components are the *N*-body integrator **kira** and a package called **SeBa** for stellar and binary evolution.

Both components run for any simulation started within the **Starlab** environment. However, data output via the **SeBa** module is only produced when the user specifies that the stars in the simulation undergo stellar or binary evolution ([Hut 2003](#)). It is also possible to add an external tidal field to simulate the effect of the Galactic gravitational potential. However, I have not used this parameter in the simulations described in this work.

### 2.2.1 *N*-body integrator kira

The *N*-body integrator **kira** uses an input *N*-body system defined by initial conditions for a star-forming region (e.g. initial radius, number of stars, virial ratio) and evolves it over a certain time period giving output at different snapshots along the way. The motion of the particles, i.e. stars, are followed using a fourth-order, block-timestep Hermite scheme, which is used for the time integration in collisional stellar systems (e.g. [Makino & Aarseth 1992](#); [Portegies Zwart et al. 2001](#); [Dehnen & Read 2011](#)). The Hermite

scheme is a predictor-corrector scheme using interpolation (Makino & Aarseth 1992). It first predicts position  $\mathbf{x}_p$  and velocities  $\mathbf{v}_p$  of a particle in the system during a timestep  $\Delta t = t - t_0$  with the known values of acceleration  $\mathbf{a}$ , current position  $\mathbf{x}$ , velocity  $\mathbf{v}$  and ‘jerk’  $\mathbf{j}$  at the beginning of the respective timestep (see Eq. B1 and B2 in Portegies Zwart et al. 2001):

$$\mathbf{x}_p = \mathbf{x} + \mathbf{v}\Delta t + \frac{1}{2}\mathbf{a}\Delta t^2 + \frac{1}{6}\mathbf{j}\Delta t^3, \quad (2.5)$$

$$\mathbf{v}_p = \mathbf{v} + \mathbf{a}\Delta t + \frac{1}{2}\mathbf{j}\Delta t^2. \quad (2.6)$$

These equations are third-order Taylor series expansions for the timestep  $\Delta t$ . For the predicted time, the new acceleration  $\mathbf{a}_p$  and ‘jerk’  $\mathbf{j}_p$  are estimated using the predicted position  $\mathbf{x}_p$  and velocity  $\mathbf{v}_p$  in Eq. 2.2 and 2.3. The motion of each particle is then corrected by information gained from additional derivatives  $\mathbf{k}$  and  $\mathbf{l}$  (see Eq. B3 and B4 in Portegies Zwart et al. 2001):

$$\mathbf{k} \equiv \frac{1}{2}\mathbf{a}''\Delta t^2 = 2(\mathbf{a} - \mathbf{a}_p) + \Delta t(\mathbf{j} - \mathbf{j}_p), \quad (2.7)$$

$$\mathbf{l} \equiv \frac{1}{6}\mathbf{a}'''\Delta t^3 = -3(\mathbf{a} - \mathbf{a}_p) - \Delta t(2\mathbf{j} - \mathbf{j}_p). \quad (2.8)$$

Finally, the corrected position  $\mathbf{x}_c$  and velocity  $\mathbf{v}_c$  for this timestep are calculated using the corrections  $\mathbf{k}$  and  $\mathbf{l}$  (see Eq. B5 and B6 in Portegies Zwart et al. 2001):

$$\mathbf{x}_c = \mathbf{x}_p + \left(\frac{1}{20}\mathbf{l} + \frac{1}{12}\mathbf{k}\right)\Delta t^2, \quad (2.9)$$

$$\mathbf{v}_c = \mathbf{v}_p + \left(\frac{1}{4}\mathbf{l} + \frac{1}{3}\mathbf{k}\right)\Delta t. \quad (2.10)$$

### Timestepping in *kira*

When calculating the motion of stars in the  $N$ -body simulations, the timestep  $\Delta t$  plays a very important role for the accuracy and efficiency of the  $N$ -body integration. Smaller timesteps provide more accurate calculations but increase the time required for a calculation step, thereby affecting the length of the  $N$ -body simulations. To balance the accuracy with an efficient performance of the simulations, the Hermite scheme used in the *kira* integrator employs a block-timestep, which essentially groups stars (or particles) together in the calculations and advances them to the next timestep as a group (Aarseth 2003; Dehnen & Read 2011).

The block-timestep approach in the Hermite method uses a base timestep  $\Delta t_0$ , which

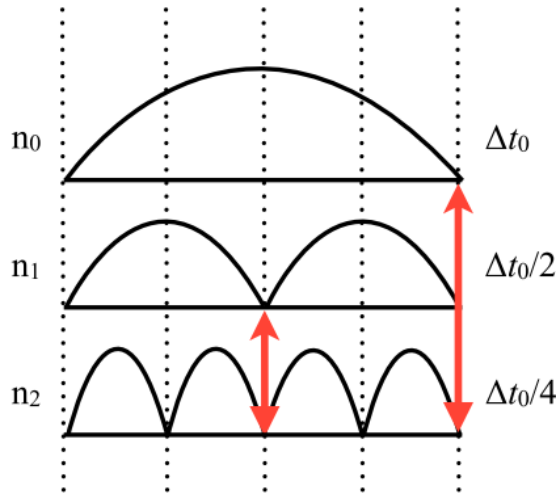


Figure 2.2: Block timesteps in the *kira*  $N$ -body integrator. The particles during each timestep in the simulations are organised in a hierarchy for defined levels  $n$  based on powers of two. The base step  $\Delta t_0$  is the maximum timestep. Shorter timesteps are possible for more accurate calculations. Particles can be synchronised between the levels at defined points, indicated by the “red” arrows. All particles are re-synchronised at the end of each timestep. Figure credit: [Dehnen & Read \(2011\)](#).

is organised into a hierarchy based on powers of two for defined levels of  $n$  ([Makino 1991](#)):

$$\Delta t_n = \frac{\Delta t_0}{2^n} \tag{2.11}$$

This timestep  $\Delta t_0$  represents the maximum timestep for a set of particles with shorter timesteps possible for higher values of  $n$ . Each particle in the calculations is assigned a timestep based on the predicted position of the particle at the end of the timestep. The smaller the assigned timestep, the more accurately the positions, velocities and accelerations will be calculated during each base step  $\Delta t_0$ . Fig. 2.2 from [Dehnen & Read \(2011\)](#) illustrates this approach. The base step for  $n = 0$  is the longest or maximum timestep and provides the calculations for particles where less precision is required, e.g. particles on a straight trajectory, such as already ejected stars. If more accuracy is required, the number of calculations per base timestep doubles for each higher level. The levels are synchronised with each other at specific times indicated by the “red” arrows in the figure and particles can change levels (upwards/downwards) at these points as well. At the end of the base timestep all particles are re-synchronised, and the process is repeated until the simulation has concluded ([Dehnen & Read 2011](#)).

### N-body unit scaling

The input parameters in the initial conditions for the simulations are given in SI-units (International System of Units). The simulations itself use  $N$ -body units, so the first step is the scaling of the SI-units to  $N$ -body units.

Within the  $N$ -body model, the units are then standardised to reduce computational errors and resources. The standardisation follows [Heggie & Mathieu \(1986\)](#), where  $G = M = -4E = 1$ .  $G$  is the gravitational constant,  $M$  is the initial total mass and  $E$  is the initial energy of the system. The value of  $E = -1/4$  is due to the choice of the value for the virial radius  $R = 1$  in a bound system in virial equilibrium. The scale-free  $N$ -body units are used within the simulations to increase accuracy and improve efficiency ([Heggie & Mathieu 1986](#)). For example, setting  $G = 1$  reduces the amount of more detailed calculations. This standardisation also turns the order of magnitude of other values to 1. Instead of storing exponents, the numbers calculated can be stored to higher accuracy using the same amount of storage memory, thereby using it more efficiently. At the end of the simulations, the standardised  $N$ -body units are scaled back into physical SI-units for analysis.

### Treatment of binaries in kira

The motion of single stars is only affected by the interactions with the rest of the stellar systems in the simulations. For binary (or multiple) systems, the motion of each component in a binary is strongly influenced by an additional factor, which is the bound binary companion. The binding energy in a close/hard binary can be much higher than the energy of an interaction with a perturbing, passing star ([Hills 1975a](#); [Heggie 1977](#)). The further a perturbing star is away from a binary the lower its chance to disrupt the binary. In *kira*, this is handled by only keeping a list of the closest perturbers, which are recalculated every time the binary's centre of mass is updated ([Portegies Zwart et al. 2001](#)).

*kira* uses a perturbation ratio to determine whether the motion of each binary is purely driven by two-body motion and its motion is calculated using Newtonian gravity or if the influence of the perturber needs to be considered. This perturbation ratio is the ratio of external perturbation (by another body in the simulation) to the internal binary binding energy. This dimensionless ratio is approximately  $10^{-6}$  for unperturbed binaries and  $10^{-5}$  to  $10^{-6}$  for lightly perturbed binaries. For these lightly perturbed binaries, an approach based on the method by [Mikkola & Aarseth \(1998\)](#) is followed, which slows down the internal binary motion and increases the perturbing influence.

Strongly perturbed binaries are fully resolved, and the motion of their individual components is directly computed. During a perturbation, they are treated as two single stars, whereas they are treated (and evolved) as a binary using the **SeBa** package when they are unperturbed (Portegies Zwart et al. 2001).

### 2.2.2 Stellar and binary evolution package SeBa

The evolution of stars and binaries in my  $N$ -body simulations are calculated by the **SeBa** package, which is described in Appendix B2 in Portegies Zwart et al. (2001). The stars in the  $N$ -body simulations are subdivided into different types, which represent different evolutionary stages in the star (and planet) formation process. These types include planets (not included in my simulations), MS stars, hypergiants, helium stars, Hertzsprung gap stars, subgiants, horizontal branch stars, supergiants, Thorne-Żytkov objects, BHs, NSs, white dwarfs and disintegrated stars. The **SeBa** package evolves stars using a mass-radius relation that is time-dependent and it uses solar metallicities (for the relations see Eggleton et al. 1989, 1990; Tout et al. 1997; Hurley et al. 2000). The stars in the simulations can undergo mass-transfer, collide, merge and also explode as a SN to become compact objects. The mass of the core of the star as well as the mass-loss rate of higher-mass stars are not described in the above relations but are included based on the method described in Portegies Zwart & Verbunt (1996), which provides additional prescriptions for the evolution of single stars as well as binaries.

The  $N$ -body simulations with stellar/binary evolution analysed in this work only cover the first 5 Myr of the evolution. For the highest-mass stars, this can cover the complete lifetime from formation to death. Star with an initial mass of  $>8 M_{\odot}$  in my simulations will eventually undergo a SN to end form a compact object (NS or BH) or they can even collapse directly into a BH.

In my simulations, only stars above  $25 M_{\odot}$  ZAMS lose mass via stellar winds and will eventually collapse into a BH following a SN explosion. Only for stars with these high ZAMS masses do I see this evolutionary step in my 5 Myr time frame. However, it is possible for massive stars to also collapse directly into a BH without a SN ( $\sim 40$ - $50 M_{\odot}$  ZAMS, e.g. Heger et al. 2003; Spera et al. 2015; Smartt 2015). This is not included in the **SeBa** protocol used in my simulations and all stars  $>25 M_{\odot}$  will form a BH after a SN instead.

Stars with a ZAMS mass of  $8$ - $25 M_{\odot}$  become NSs or BHs at the end of their lives following SNe. These BHs are the consequence of NSs that collapse further (e.g. when they accrete mass from a binary companion and become too massive). However, it takes

stars in this mass range longer than the 5 Myr period covered in my simulations with stellar evolution to reach these evolutionary steps (at least  $\sim 7$  Myr for stars with masses of  $25 M_{\odot}$  to form a NS). For these stars, the last evolutionary step reached during my 5 Myr simulations is the main-sequence.

When these high-mass stars are in binaries, they can eject stars as RWs or WWs as a consequence of the BSS. As discussed in Chapter 1.5.1, the mass loss of the primary star during a symmetric SN can lead to a recoil causing the binary to unbind and one or both of the binary stars to get ejected, if enough mass is lost (Blaauw 1961; Boersma 1961). In addition, asymmetric SNe can provide an additional natal velocity kick, which can accelerate the SN remnant to very high velocities. However, this additional natal kick will not occur, if the primary collapses directly into a BH without a SN. Although, it is still possible that a RW/WW star is created at this point due to a “Blaauw-Boersma”-recoil. This is normally connected to SNe, but is actually based on the reaction to the instantaneous loss of more than half of the system’s mass, which can also occur during a direct collapse of a massive star into a BH.

Stars with a ZAMS mass of below  $8 M_{\odot}$  conclude their evolution as white dwarfs. However, this happens far outside the 5 Myr window (on the scale of Gyr). After 5 Myr, most of these stars would still be within their pre-MS phase as Class II/III stars (see Chapter 1.3.3) or very young MS-stars. In the SeBa version used in my simulations, I do not identify any evolutionary stage before the MS-phase, so these stars would be shown as MS-stars during the simulations.

### Evolution of binaries

Binaries are only evolved in the SeBa package when they are either isolated or unperturbed. This evolution happens when one of the stars in the system requires an update due to a change in its individual stellar evolution or while a binary undergoes mass transfer, which occurs on smaller timesteps. A binary evolution loop starts with a check for angular momentum loss, e.g. due to stellar winds from the primary leading to mass loss, followed by a check if this angular momentum loss has led to a binary merger. If that is not the case, the primary star is evolved, and binary parameters are adjusted. If no SN has occurred, the secondary star is also evolved (with adjustment of binary parameters and another SN check). Further checks and adjustments to the orbits are applied, if needed. Finally, this loop concludes with checks related to a possible mass transfer via Roche-lobe overflow and adjustments to the parameters of the stars (e.g. changes in mass) and the binary (Portegies Zwart et al. 2001).



### 2.2.3 Other $N$ -body software

In the following two subsections, I will briefly describe two other widely used  $N$ -body software codes/environments.

#### NBODY# code series

The development of the NBODY# code series started in the early 1960s using force polynomials and the most recent version of this code is NBODY7 (Aarseth 1999, 2012). NBODY7 was developed specifically to deal with difficulties of handling compact subsystems in  $N$ -body simulations, such as interactions in BH-binaries, which involves general relativistic effects and post-Newtonian terms (Aarseth 2012).

NBODY6, as well as the NBODY4 code use a Hermite fourth-order integration scheme similar to the one used in the `Starlab kira` integrator. The NBODY4 code is optimised for use with the special-purpose HARP (“Hermite Accelerator Pipeline”) computers, whereas the NBODY6 is the version for normal workstations and supercomputers (see review by Aarseth 1999). The NBODY6 code has also been extended for use with general parallel supercomputers in the NBODY6++ code and has been recently adapted to be used with GPUs (see Wang et al. 2015, and references therein).

NBODY6 and `Starlab` provide comparable results in most areas. Anders et al. (2012) compared these two codes and found differences in the energy conservation (total energy change during one unit of  $N$ -body time) in that the energy conservation in NBODY6 is worse compared to `Starlab`. The results in Anders et al. (2012) for NBODY6 are in agreement with comparisons of `Starlab` with the NBODY4 version of the code in Anders et al. (2009). Other differences appear in the distances between the furthest star and the centre of the simulated cluster and most importantly in the treatment of binaries.

#### AMUSE software environment

While `Starlab` and the NBODY# code series have been in use for decades, the Astrophysical Multi-purpose Software Environment (AMUSE) is one of the more recent additions (e.g. Portegies Zwart et al. 2013; Pelupessy et al. 2013; Portegies Zwart & McMillan 2018). It is essentially a framework for using existing codes with a Python-based interface. For solving the gravitational dynamics, solvers are available for differently sized  $N$ -body systems, e.g. direct  $N$ -body integrators with Hermite timestepping schemes and hardware accelerated force calculations. AMUSE includes modules for mixing gravitational  $N$ -body simulations with hydrodynamical SPH simulations. Single and binary

stellar evolution modules (SSE, BSE and also the **SeBa** package) are available and there are also options to include radiative transfer. However, these are limited due to the high level of computational cost (Pelupessy et al. 2013).

## 2.3 Setting up $N$ -body simulations to search for RW/WW stars

To set up the  $N$ -body simulations, I define a set of initial conditions that are used to create the input data for the simulations at time 0. Below, I describe the set-up of the spatial and kinematic substructure using the box-fractal method, the sampling of stellar masses from the IMF, the set-up of primordial binaries, the stellar evolution protocol and other parameters that I define in the initial conditions. Project-specific details about the  $N$ -body simulations are discussed in the Methods section in Chapters 4, 5 and 7.

### 2.3.1 The box-fractal method

In Chapter 1.2, I describe how many star-forming regions show evidence of past/initial or current spatial and kinematic substructure. In  $N$ -body simulations, this initial substructure can be created using the box-fractal method as described in Goodwin & Whitworth (2004).

To define the observed spatial substructure in young star-forming regions, this method uses fractal distributions defined by a single parameter, the fractal dimension  $D$ . Starting with a cube with side  $N_{\text{div}} = 2$ , a parent particle is placed at the centre. This first parent cube is subdivided into equal-sized  $N_{\text{div}}^3$  sub-cubes with a first-generation descendant in each centre. Depending on the survival probability  $N_{\text{div}}^{(D-3)}$  that is set by the fractal dimension  $D$ , these descendants can become parents themselves. For a low fractal dimension, fewer descendants become parents, whereas more descendants survive when using a high fractal dimension. Descendants that do not survive are deleted along with their parent. The positions of the surviving particles are adjusted by adding a small amount of noise. This process continues until more stars than required are generated within the original cube. A sphere is then cut from this cube and the remaining stars are reduced down to the required number by random deletion (Goodwin & Whitworth 2004).

Fig. 2.3 shows two different levels of initial spatial substructure as a result of applying the box-fractal method. On the left is a star-forming region with a fractal dimension of

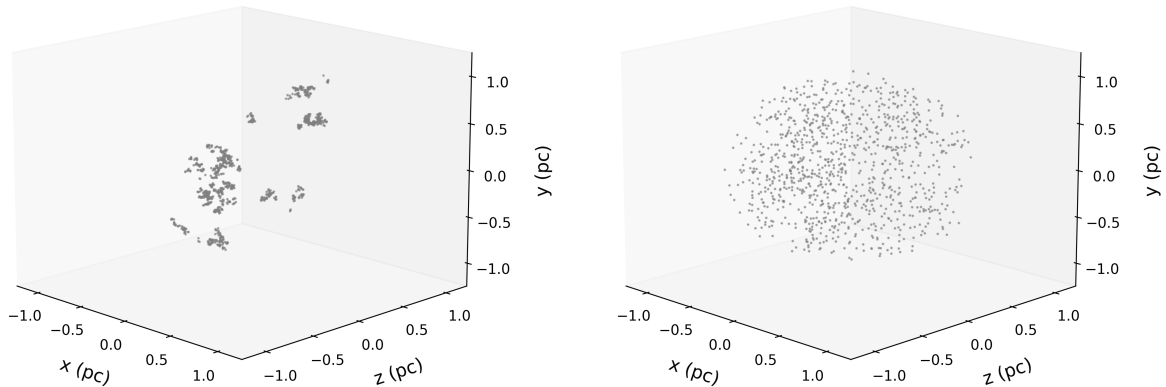


Figure 2.3: Different initial levels of spatial substructure for star-forming regions with 1000 single stars and an initial radius of 1 pc at 0 Myr. Left: A region with a fractal dimension of  $D = 1.6$ , i.e. a high amount of initial spatial substructure (“clumpy”). Right: A region with a fractal dimension of  $D = 3.0$ , i.e. an almost sphere-like initial spatial distribution (“smooth”).

$D = 1.6$ , on the right is a region with an initial fractal dimension of  $D = 3.0$ . This figure clearly shows the difference between a highly spatially substructured region compared to a smooth, sphere-like distribution.

Like the spatial structure, the velocity structure in my simulations is also set up to mimic observed star-formation environments. Molecular gas clouds show evidence of turbulence, and the spatial and kinematic distributions of stars suggest that they formed from these turbulent clouds. The velocity dispersion increases with the size of the clouds. In molecular clouds, large velocity dispersions can occur on large scales whereas on small scales there are smaller dispersions, i.e. similar velocities (Larson 1981). Star formation occurs in filamentary structures within these gas clouds, where the velocity dispersion is low (e.g. André et al. 2014). To represent this velocity structure in the simulations I follow Goodwin & Whitworth (2004), which results in close stars with similar velocities and distant stars with different velocities.

The observation-based Larson (1981) relations show that velocities scale as  $v(L) \propto L^{0.38}$  on local scales up to a size  $L$  of 100 pc. In the box-fractal method used here, the velocities scale as  $v(L) \propto L^{3-D}$ , so when  $D = 2.6$  then  $v(L) \propto L^{0.4}$ , whereas velocities scale as  $v(L) \propto L^{1.4}$ , when  $D = 1.6$ . So the initially highly spatially substructured box-fractals ( $D = 1.6$ ) are less consistent with the Larson (1981) relations than those with a more moderate amount of initial spatial substructure ( $D = 2.0$  and  $2.6$ ). The

substructure will be erased on timescales in the order of  $t(L) \propto L/v(L)$ . For  $D = 1.6$  this is  $t(L) \propto L^{-0.4}$ , which means substructure is erased faster on large scales; whereas for  $D = 2.6$ , the timescale is  $t(L) \propto L^{0.6}$ , which means the erasure of the substructure here is faster on small scales (Parker & Wright 2018).

The process of setting up the velocity structure starts by assigning a random velocity to the parents. The next generation inherits this velocity, which is in turn adjusted by a random component that gets smaller with every following generation. The velocities of the stars are finally scaled to different global virial ratios (subvirial, virialized or supervirial; see Chapter 1.2).

### 2.3.2 Initial mass function

The masses for single stars and, if present, primary stars in primordial binaries are sampled randomly for each simulation, so no two simulations share exactly the same stellar masses for individual stars. I sample my stellar masses from the Maschberger IMF (Maschberger 2013) and refer to Chapter 1.3.1 for the history and theory of different IMFs. The Maschberger IMF is a combination of a Chabrier (2003a,b, 2005) lognormal IMF approximation for low-mass stars combined with the power-law slope of Salpeter (1955) for stars above  $1 M_{\odot}$ :

$$p(m) \propto \frac{\left(\frac{m}{\mu}\right)^{-\alpha}}{\left(1 + \left(\frac{m}{\mu}\right)^{1-\alpha}\right)^{\beta}} \quad (2.12)$$

The IMF is described by the above probability density function, where  $\alpha = 2.3$  (power-law exponent for higher mass stars),  $\beta = 1.4$  (describing the IMF slope for lower-mass stars) and  $\mu = 0.2$  (average stellar mass) (Maschberger 2013).

The integral of the probability density function in Eq. 2.12 is represented by the following auxiliary function:

$$G(m) = \left(1 + \left(\frac{m}{\mu}\right)^{1-\alpha}\right)^{1-\beta}, \quad (2.13)$$

which can be turned into the cumulative distribution function, where  $m_u$  and  $m_l$  are the

upper and lower mass limits:

$$P(m) = \frac{G(m) - G(m_l)}{G(m_u) - G(m_l)}. \quad (2.14)$$

The inversion of this cumulative distribution function then provides the quantile function, which allows to easily draw a random mass  $m(u)$ , where  $u$  is a random number, within the mass range given by the lower and upper limit (Maschberger 2013):

$$m(u) = \mu \left( \left[ u \left( G(m_u) - G(m_l) \right) + G(m_l) \right]^{\frac{1}{1-\beta}} - 1 \right)^{\frac{1}{1-\alpha}}. \quad (2.15)$$

The total masses of the star-forming regions differ between the simulations described in this work. In the initial conditions, I define the number of systems present in the  $N$ -body simulations. The total masses of the simulated regions are then a consequence of the sampling from the IMF. In Chapters 4, 5 and 7, the average masses for the star-forming regions depend not only on the total number of systems but also on the presence/absence of primordial binaries. For the simulations in Chapters 4 and 7, I have a similar number of stars per simulation, even though the number of systems in Chapter 7 is smaller due to the presence of primordial binaries. This presence of primordial binaries is also responsible for a difference in the average total mass of the star-forming regions.

### 2.3.3 Primordial binaries

The number of systems in a simulation is equal to the number of stars if no primordial binaries are included in the initial conditions. If primordial binaries are present at the beginning of the simulations, then the total number of stars will not be equal to the number of systems and will instead depend on the binary fraction applied to the simulations.

For two of my projects, I incorporate primordial binaries in my simulations. These projects are described in Chapters 5 and 7. The fraction of the stars that begin as primordial binaries is defined in the binary fraction  $f_{\text{bin}}$  and it depends on the mass of the primary star  $m_p$ . This fraction is defined as:

$$f_{\text{bin}} = \frac{B}{S + B}, \quad (2.16)$$

where  $S$  and  $B$  represent the number of single or binary systems, respectively. I do not

Table 2.1: Binary fractions in the  $N$ -body simulations. Column 1 shows the mass range based on the mass of the primary star; Column 2 shows the binary fraction  $f_{\text{bin}}$ .

$m_p$ [ $M_\odot$ ]	$f_{\text{bin}}$	Source
$0.10 \leq m_p < 0.45$	0.34	Janson et al. (2012b)
$0.45 \leq m_p < 0.84$	0.45	Mayor et al. (1992)
$0.84 \leq m_p < 1.20$	0.46	Raghavan et al. (2012)
$1.20 \leq m_p \leq 3.00$	0.48	de Rosa et al. (2012, 2014)
$m_p > 3.00$	1.00	Mason et al. (1998) Kouwenhoven et al. (2007)

include any primordial higher-order multiple systems (triples or quadruples). For all simulations with primordial binaries, I impose  $f_{\text{bin}}$  as shown in Table 2.1 depending on the primary star’s mass with binary separations as shown in Table 2.2.

These binary fractions and separations are similar to those observed in the Galactic field (e.g. Köhler et al. 2006; Reipurth et al. 2007; Parker 2014; Duchêne et al. 2018). For the two star-forming regions I investigate in this work (ONC and NGC 2264), Kounkel et al. (2016) suggested that the binary fractions in NGC 2264 are similar to those in the ONC and that both are consistent with how binaries are distributed in the field.

For stars in primordial binaries, the secondary star is assigned a mass  $m_s$  based on a flat mass ratio distribution, which is observed in the field and many star-forming regions (e.g. Reggiani & Meyer 2011, 2013; Parker & Reggiani 2013). The binary mass ratio  $q$  is:

$$q = \frac{m_s}{m_p}. \tag{2.17}$$

The secondary stars in all my simulations with primordial binaries are allowed to have a mass lower than the primary stars, which are limited to masses  $\geq 0.1 M_\odot$ . The secondary stars can have a mass of  $m_p > m_s \geq 0.01 M_\odot$ , so BDs are possible as secondaries in the primordial binaries; however, not as primary or single stars.

Assigning the mass of the secondary star from a flat mass ratio distribution leads to higher overall total masses for the star-forming regions compared to simulations where only single stars are present. Fig. 2.4 from Duchêne & Kraus (2013) shows the difference between a flat mass ratio distribution, and one based on following a single-star mass function.  $\gamma$  on the y-axis represents the index of the power-law fit to the mass ratio distribution, with a flat mass ratio being given the index 0. In Fig. 2.4, the dotted curve represents the flat mass ratio distribution, i.e. equal likelihood for the companion to have any mass within the given mass range. The dashed curve represents an index based on the assumption that the binary companions follow a single-star IMF by Chabrier (2003a)

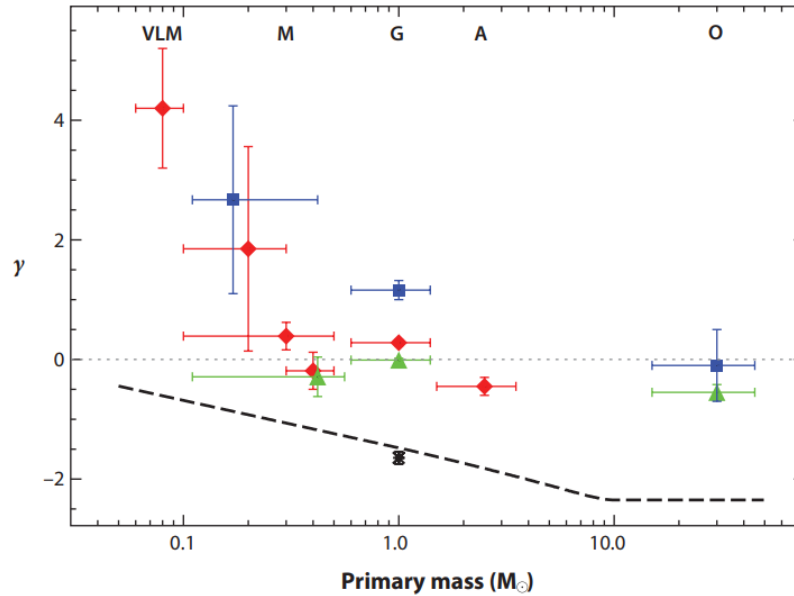


Figure 2.4: Flat mass ratio sampling compared to IMF sampling. The dotted curve represents the flat mass ratio distribution. The dashed curve represents an index based on the assumption that the binary companions follow a single-star IMF by [Chabrier \(2003a\)](#) with the fitting of a simple power law to the resulting companion mass distribution for a mass ratio range with  $0.1 \leq q \leq 1$  Figure credit: Fig. 3 from [Duchêne & Kraus \(2013\)](#).

with the fitting of a simple power-law to the resulting companion mass distribution for a binary mass ratio range with  $0.1 \leq q \leq 1$  ([Duchêne & Kraus 2013](#)). This figure shows that for an IMF-based mass for the binary companion, the index  $\gamma$  changes with higher primary masses, i.e. lower companion masses would be assigned. Using the flat mass ratio distribution for the secondary stars in the binaries lead to higher masses overall, due to a higher likelihood of equal-mass binaries for stars with higher primary masses than in an IMF-sampled scenario.

The initial binary separation in my simulations, i.e. semi-major axis, is based on a log-normal distribution with mean values for the separation  $\bar{a}$  in au and the variance shown in Table 2.2 for different primary masses. For primary masses  $> 3 M_{\odot}$ , I use the [Öpik \(1924\)](#) law, which prescribes a flat distribution in the logarithms of the binary separations. These values also follow recent observations of binaries in the field (e.g. [Duchêne & Kraus 2013](#); [Parker & Meyer 2014](#)). Binaries with orbital periods  $< 0.1$  au have circular orbits, which is in line with observations. Binaries with larger orbital periods have eccentricities drawn randomly from a flat distribution (e.g. [Duchêne & Kraus 2013](#); [Parker & Meyer 2014](#); [Wootton & Parker 2019](#)).

Table 2.2: Mean binary separations in the  $N$ -body simulations. Column 1 shows the mass range based on the mass of the primary star; Column 2 shows the mean binary separation  $\bar{a}$ ; Column 3 represents the variance  $\sigma_{\log \bar{a}}$  of the log-normal fit to the binary separation distributions.

$m_p [M_\odot]$	$\bar{a}$ [au]	$\sigma_{\log \bar{a}}$	Source
$0.10 \leq m_p < 0.45$	16	0.80	Janson et al. (2012b)
$0.45 \leq m_p < 1.20$	50	1.68	Raghavan et al. (2012)
$1.20 \leq m_p \leq 3.00$	389	0.79	de Rosa et al. (2014)
$m_p > 3.00$	Öpik law (0-50)	-	Öpik (1924), Sana et al. (2013)

### 2.3.4 Stellar and binary evolution

The stellar and binary evolution package **SeBa** is described in Chapter 2.2.2. For the simulations in Chapters 5 and 7, this protocol is “turned on”, which means that the stars and binaries evolve over the time period covered in the simulations. In Chapter 4, this option is not active, which means the stars (and dynamical binaries) do not evolve and will have the same mass (and evolutionary stage) at the end of the simulations as they do at the beginning.

The simulations with stellar/binary evolution run for a maximum of 5 Myr in the projects described in this work. There are no SNe during the 4 Myr simulations for the ONC. However, there are SNe for the highest-mass stars in the NGC 2264 simulations at an age of 4-5 Myr. The furthest evolutionary step for the highest-mass stars in the ONC simulations is that of a helium star, whereas there is the formation of several BHs in the NGC 2264 simulations. In all my simulations with binary evolution, I also find several binary mergers as a result of the evolution.

### 2.3.5 Other parameters defined in the initial conditions

The total simulation time in the simulations presented here is 10 Myr or less. For the simulations in Chapters 4, I evolve the star-forming regions to this upper value, whereas I stop the simulations earlier for the ONC (4 Myr) and NGC 2264 (5 Myr). I take snapshots at different intervals (0.01 Myr and 0.1 Myr, depending on the region), this provides me with information about the position and motion of my stars during the simulations. I also define the initial radius of the star-forming region, which greatly influences the initial density of the regions.



## Chapter 3

### Observations with *Gaia*

In this chapter, I will give a brief introduction to the *Gaia* telescope and how it operates. I will then continue by describing how I use *Gaia* data to find stars that have been ejected from young star-forming regions. I start with a description of the applied quality filters, followed by the method I use to trace back stars using the astrometric data and how I construct Colour Absolute Magnitude Diagrams (CAMDs) using the photometric data. I conclude this chapter by explaining the use of isochrones to obtain age estimates.

### 3.1 Astrometry and the *Gaia* telescope

Astrometry deals with the measurement of the position and motion of objects in the night sky. This astronomical discipline has a long history dating back over 2000 years and has made large advancements since the early 18th century when telescopes became more and more powerful.

The *Gaia* telescope is the astrometric successor to the HIPPARCOS satellite, which was operational during the early 1990s (Perryman et al. 1997) and both of these telescopes are projects of the European Space Agency (ESA). HIPPARCOS was the first mission to measure the absolute parallax of over 100,000 stars with then unrivalled accuracy. The *Gaia* mission is mainly an astrometric endeavour to measure the spatial and velocity distributions of stars in our Galaxy. To fully understand the formation, structure and evolution of the Milky Way, it also aims to determine a wide range of non-astrometric astrophysical properties, such as the effective temperature, surface gravity and stellar radii using photometry (Gaia Collaboration et al. 2016a).

*Gaia* was first proposed in 1993 and the science case followed in 2001. The name of the telescope originated as an acronym for “Global Astrometric Interferometer for Astrophysics”, even though the telescope ended up not using an interferometer but is based on direct imaging instead. The construction phase started in 2006 and the mission was launched in December 2013 with an initial plan for operations covering a five-year period starting in mid-2014 (Gaia Collaboration et al. 2016a). The *Gaia* mission is still operational in 2021, having been extended to at least 2022<sup>1</sup> with an indicative extension to 2025<sup>2</sup>. In mid-2019, it successfully completed a major manoeuvre to change its orbit allowing it to continue observing. During this manoeuvre, the orbit of the telescope was changed so that it will not move through the Earth’s shadow in the coming years. This would have negatively affected the telescope in several ways, e.g. shutdown of

---

<sup>1</sup><https://sci.esa.int/s/AjogJnw>

<sup>2</sup><https://sci.esa.int/s/80JDymW>

the telescope when it would not receive solar power to charge its battery and thermal disturbances during the dramatic change in temperature, which would affect operations for weeks at a time<sup>3</sup>.

## 3.2 Observations with *Gaia*

The *Gaia* satellite features two identical (three-mirror anastigmatic) telescopes, each with an aperture of 1.45 metre by 0.50 metre pointing in directions separated by a baseline angle of  $\sim 105^\circ$ . Its observations are based on the principle of scanning space astrometry, which consists of a satellite slowly spinning around its axis measuring the time it takes for stars and other objects to transition across the focal plane of the telescope (Lindegren & Bastian 2010). *Gaia* co-rotates with the Earth around the Sun and is located around the second Lagrange point L2. Fig. 3.1 illustrates this scanning procedure for *Gaia* with the spin axis and the path of the viewing direction of one of the two telescopes during a period of four days (Gaia Collaboration et al. 2016a).

The *Gaia* astrometric catalogue consists of a large number of repeat observations with data in Data Release 1 (DR1) covering an observation period of 14 months, increasing to 22 months in Data Release 2 (DR2) and 34 months with Early Data Release 3 (EDR3) (Gaia Collaboration et al. 2016a, 2018a, 2021). *Gaia* does not observe new parts of the sky with each data release. It observes the whole sky (northern and southern hemisphere) repeatedly and the longer its observation period, the higher the expected accuracy of its measurements even at fainter magnitudes.

Unlike other astronomy missions, *Gaia* does not have a predefined input catalogue of observation sources but instead is an unbiased (but flux-limited) survey of the sky. It has a magnitude limit on the faint end (apparent magnitude  $G \approx 21$  mag), but also on the bright end (apparent  $G$ -magnitude  $G \approx 3$  mag). While it can observe some of the stars brighter than this magnitude, the data processing is not automated making data for these stars currently unavailable. Its wavelength coverage is 330–1050 nm, which is the broadband white-light photometric  $G$ -band. In addition to providing astrometry, it also produces detailed photometry data using two photometers, with the blue one operating in the range of 330–680 nm ( $G_{BP}$  passband) and the red one in the range of 640–1050 nm ( $G_{RP}$  passband) (Gaia Collaboration et al. 2016a; Carrasco et al. 2016; van Leeuwen et al. 2017).

---

<sup>3</sup>[http://www.esa.int/Enabling\\_Support/Operations/Gaia\\_s\\_biggest\\_operation\\_since\\_launch](http://www.esa.int/Enabling_Support/Operations/Gaia_s_biggest_operation_since_launch)

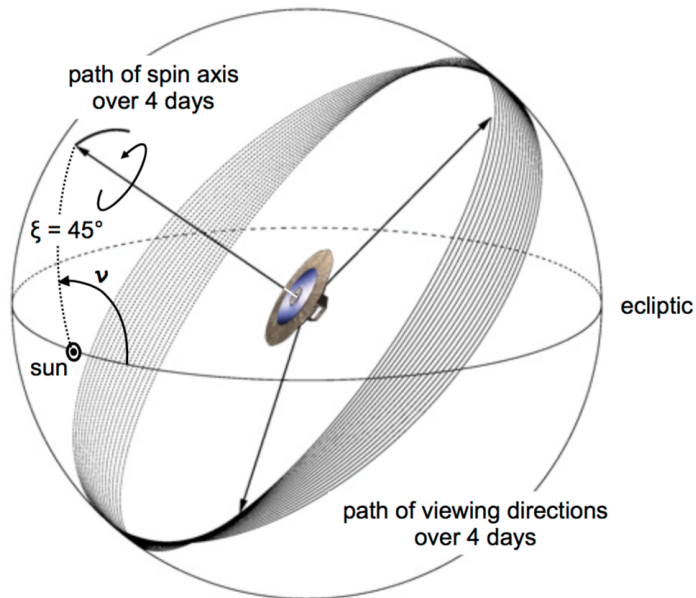


Figure 3.1: Illustration showing the scanning procedure of *Gaia* during a time period of four days. The spin axis ( $z$ ) and its path are shown as well as the corresponding path of the preceding viewing direction. Figure credit: [Gaia Collaboration et al. \(2016a\)](#).

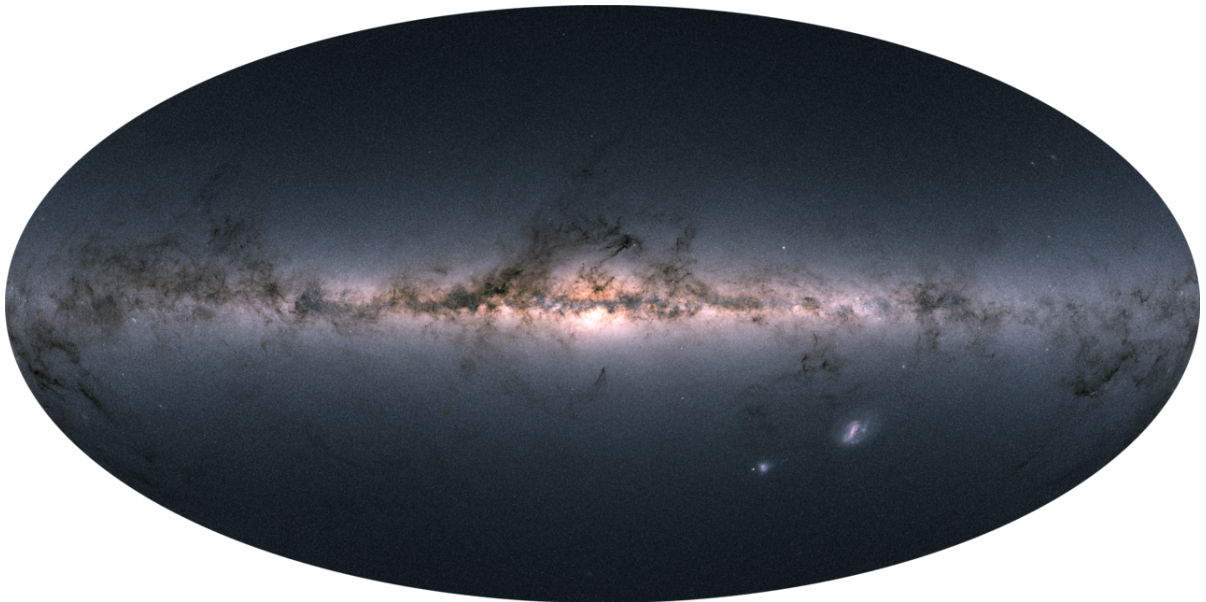


Figure 3.2: Map of the Milky Way using flux (measured brightness) of  $\sim 1.7$  billion sources in *Gaia* DR2. It was composed by separately compiling fluxes measured in the three photometric bands  $G$ ,  $G_{BP}$  and  $G_{RP}$ . Image credits: ESA/Gaia/DPAC, A. Moitinho/A. F. Silva/M. Barros/C. Barata, University of Lisbon, Portugal; H. Savietto, Fork Research, Portugal.

*Gaia* DR1 was released in September 2016 and provided a first glimpse at the capabilities of this mission (Gaia Collaboration et al. 2016a). However, it was limited in the availability of accurate astrometry and photometry, i.e. five-parameter astrometry with position, parallax and proper motion information was only available for a small set of stars ( $\sim 2$  million sources). This changed with the second data release in April 2018. *Gaia* DR2 provided five-parameter astrometry for over 1.3 billion sources down to an apparent G-magnitude limit of  $G \approx 21$  mag (Gaia Collaboration et al. 2018a). In addition to the highly accurate astrometry, photometry information for roughly the same number of sources was also made available in the three passbands ( $G$ ,  $G_{BP}$  and  $G_{RP}$ ). Fig 3.2 uses the measured brightness of  $\sim 1.7$  billion of the DR2 sources to construct a map of the Milky Way and even includes satellite galaxies, such as the LMC and SMC (Gaia Collaboration et al. 2018a). *Gaia* EDR3 was released in December 2020 and covers a longer observation period than the previous releases, which increases the number of sources available with five-parameter astrometry and photometry to  $\sim 1.5$  billion (Gaia Collaboration et al. 2021). In this latest data release, no further RVs, extinction or reddening information were added.

### 3.3 Using *Gaia* DR2 to find ejected stars

In this section, I describe the analysis process I use for the search for ejected stars around two young star-forming regions, the ONC and NGC 2264. The results of these analyses are presented in Chapters 5 and 7 for these two regions and were published in Schoettler et al. (2020) and Schoettler et al. (2022).

For the data selection and download from the public *Gaia* archive<sup>4</sup>, I start by selecting centre values for the regions. For both young star-forming regions, I adopt the centres as described in Kuhn et al. (2019). The values for these centres are shown in Tables 5.1 and 7.1 in the respective chapters describing these two projects. From these central locations, I select sources within a radius of 100 pc on the sky, which translates to different angular sizes in degrees, due to the different distances to these regions. For the ONC, which is closer to Earth at a distance of  $\sim 400$  pc, 100 pc around the centre translates to  $\sim 14^\circ$ , whereas to only  $\sim 8^\circ$  for NGC 2264, which is over 300 pc further away than the ONC. The choice of the 100 pc radius is arbitrary and driven by the need to keep the data files from the *Gaia* archive at a manageable size. RWs travelling at a velocity of  $30 \text{ km s}^{-1}$  from the centre will exit the region within  $\sim 3.3$  Myr, as a velocity of  $1 \text{ km s}^{-1}$  leads to

<sup>4</sup><https://gea.esac.esa.int/archive/>

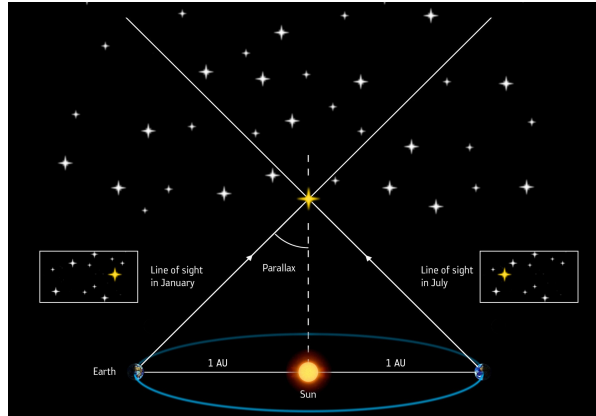


Figure 3.3: Measuring parallax with *Gaia*. While *Gaia* co-rotates with the Earth around the Sun, relatively nearby stars appear to move against the apparently fixed background stars that are much further away. The distance between the Sun and Earth is known allowing us to use the parallax angle to calculate the distance to each star targeted by *Gaia*. Image Credits: ESA/ATG medialab.

a distance of  $\sim 1$  pc being travelled in  $\sim 1$  Myr.

*Gaia* provides the distances to the stars in parallax, which is the difference in the apparent position of objects when viewed from two different lines of sight. In the case of *Gaia*, these two positions are achieved by the different positions of the Earth (and *Gaia* co-rotating with Earth) around the Sun during the year, as illustrated in Fig. 3.3. Parallax is measured in milli-arcseconds (mas) in *Gaia* and the smaller the parallax the further away an object is located from Earth. A simple way to convert from parallax  $\varpi$  to distance  $r$  in pc is based on the following relationship:  $r=1/\varpi$  with the parallax measured in arcseconds (as) and distances in parsec. Luri et al. (2018) highlighted issues with this simple approach and suggested instead the use of a Bayesian approach to convert from parallax to distances. Bailer-Jones et al. (2018) provided such a Bayesian analysis resulting in converted distances for the 1.3 billion sources with parallax in the *Gaia* DR2 catalogue and I use the distances in their database for my analysis. After I apply the distances from this catalogue to all sources within the converted 100 pc on-the-sky radius, I reduce the sample by selecting only sources within  $\pm 100$  pc of the centres in distance  $r$ .

### 3.3.1 Filtering data to achieve high quality

The data provided in *Gaia* DR2 is generally of very high quality and can often be used without applying any further quality cuts. However, in my analysis, I aim to identify very solid RW/WW candidates and therefore decided to apply further cuts to my data

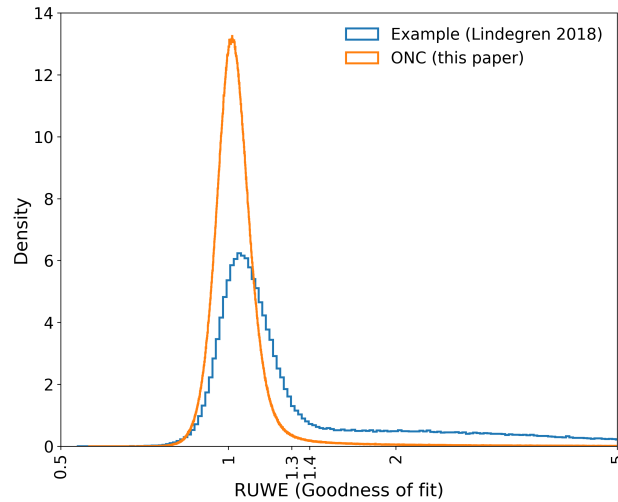


Figure 3.4: Distribution of my ONC data (orange line) according to RUWE value following the example approach in Lindegren (2018). The data from this example (blue line) result in a defined RUWE boundary of 1.4. The “knee” of my ONC distribution is located at  $\text{RUWE} = 1.3$ , which I use as my upper limit for the goodness of fit of observations with a single-star model. Figure credit: Schoettler et al. (2020).

to increase both the astrometric and photometric quality.

For the astrometric quality cut, I use the re-normalised unit weight error (RUWE) (Lindegren 2018, “Gaia known issues” website<sup>5</sup>), which is an indicator of how well the single-star model used for the data processing in DR2 fits the observations. Higher RUWE values indicate problems with the astrometry or the presence of non-single stars. The technical note GAIA-C3-TN-LU-LL-124-01 (Lindegren 2018) provided an example for calculating an upper boundary value for a good fit. This example derived a RUWE boundary of 1.4. Plotting my data for the ONC in the same way as described in the technical note results in a RUWE distribution as shown in Fig. 3.4, where I also include the example distribution from Lindegren (2018). The resulting graph shows that excluding stars with  $\text{RUWE} > 1.4$  leaves stars located in the upper tail of the distribution. Plotting this graph for my second young star-forming region NGC 2264 produces very similar results and for both regions analysed, I consider  $\text{RUWE} = 1.3$  a better (more conservative) upper boundary for my data set and I exclude all data above this value.

For the photometric quality cut, I use the photometric excess noise (flux excess factor

<sup>5</sup><https://www.cosmos.esa.int/web/gaia/dr2-known-issues>

$E$ ) filter (Lindgren et al. 2018; Gaia Collaboration et al. 2018b):

$$1.0 + 0.015(G_{\text{BP}} - G_{\text{RP}})^2 < E < 1.3 + 0.06(G_{\text{BP}} - G_{\text{RP}})^2 \quad (3.1)$$

This filter provides me with high-quality photometric data as well as further cleaning up the astrometry (Arenou et al. 2018). This filter removes sources with spurious photometry in dense, crowded areas (such as the centre of star-forming regions, Lindgren et al. 2018; Gaia Collaboration et al. 2018b). In my case, it filters a large number of stars located in the central region of the star-forming regions, which does not greatly affect the analysis as I am mainly interested in ejected stars that are no longer in these regions. The two panels in the top row of Fig. 3.5 show the difference between the unfiltered raw data on the top left and the data after applying the two quality filters on the top right. The quality filters remove a significant number of stars towards the fainter end on the CAMD.

### 3.3.2 Tracing back stars using their astrometric data

Since the published version of the search for ejected stars from the ONC (Schoettler et al. 2020), I have adjusted my approach of converting ICRS coordinates and velocities into a Cartesian coordinate system. The updated results from this approach are now presented in Chapter 5. This improved conversion approach was used for the search for ejected stars from NGC 2264 published in (Schoettler et al. 2022) and the analysis and results in Chapter 7 are identical with the published version.

In the following, I describe this updated approach for converting coordinates and follow up with the approach for tracing back stars to their birth regions. The original approach shown in Schoettler et al. (2020) follows Kuhn et al. (2019) and Getman et al. (2019). I incorporate the changes described in Vaher (2020), which takes into account the effects of individual stars' RVs on their proper motion when converting ICRS coordinates and velocities into a Cartesian coordinate system.

For stars without RV, I use the RV of the cluster centre. I then use the transformation matrix described in Vaher (2020) and shown in Eq. 3.2 to convert both the position and velocity to be cluster-aligned and then shift the origin of the coordinate system to the cluster centre.

$$\begin{pmatrix} \cos(\alpha - \alpha_c) & -\sin \delta \sin(\alpha - \alpha_c) & \cos \delta \sin(\alpha - \alpha_c) \\ \sin \delta_c \sin(\alpha - \alpha_c) & \cos \delta_c \cos \delta + \sin \delta_c \sin \delta \cos(\alpha - \alpha_c) & \cos \delta_c \sin \delta - \sin \delta_c \cos \delta \cos(\alpha - \alpha_c) \\ -\cos \delta_c \sin(\alpha - \alpha_c) & \sin \delta_c \cos \delta - \cos \delta_c \sin \delta \cos(\alpha - \alpha_c) & \sin \delta_c \sin \delta + \cos \delta_c \cos \delta \cos(\alpha - \alpha_c) \end{pmatrix} \quad (3.2)$$



Before I convert the velocities into cluster-centred velocities, I remove the Sun’s peculiar motion relative to the Local Standard of Rest (LSR, using velocity values from Schönrich et al. 2010) from the velocities of all stars in the data set as well as from the velocity parameters of the respective region’s centres (see Tables 5.1 and 7.1). I then apply a rest frame that is centred on these regions to my data set by subtracting the central velocity parameters. The xy-plane is defined as a projected representation of the positions on the sky and the z-direction represents the radial direction.

In my search for RW and WW candidates, I first trace back the positions of all stars in my data set in the xy-plane for a time period depending on the upper age estimate of the star-forming region using the converted proper motions. I define the cluster boundary based on existing member lists for each of the regions and once the backwards path of a star crosses this boundary, a star becomes a 2D-candidate. I use the time at which this path intersects the boundary as the minimum time since ejection, i.e. the flight time.

For tracing back stars in three dimensions an estimate of the depth (extent in the radial direction) of the star-forming region is required. This boundary is not well constrained or available at all for many young star-forming regions. For the two star-forming regions analysed in this work, I describe the approach to constrain the depth of the regions in the respective chapters. For both regions, the depth is larger than the radius of the region on the sky, which turns the search region from a circle (on the sky) to an ellipse in yz and xz-direction. The semi-minor axis has the value of the on-sky radius (in the x and y axes) and the semi-major axis has the value of the defined depth (in the z axis).

To trace back ejected stars in three dimensions, RVs are required in addition to the proper motions. *Gaia* DR2 provides RVs for a small subset of the sources. This subset is magnitude limited to sources with a mean apparent *G* magnitude of  $\sim 4$ –13 mag and an effective temperature in the range of  $\sim 3550$  to 6900 K (Katz et al. 2019). As a result, the number of sources with RV in *Gaia* DR2 is only  $\sim 7.2$  million, which is about 0.5% of all sources with five-parameter astrometry. The magnitude limit on the RVs reduces the number of available RVs for more distant star-forming regions, as only the brighter members will fall into the magnitude range.

The spectral type of the observed stars also affects the stars for which RVs can be measured accurately. “Early-type” stars are young, hot and massive stars that have O, B and A-type spectral classes, whereas cooler, older stars, such as G, K and M-type stars are considered to be “late-type” stars. Younger, early-type stars usually show faster rotation than older later-type stars. This rotation can lead to broader, shallower spectral

lines for these faster rotators, which makes it more difficult to measure RVs accurately (e.g. [Carroll 1928, 1933](#); [Struve 1930](#); [Collins & Truax 1995](#)).

In *Gaia* DR2, the broadness of the spectral lines for early-type stars were expected to result in higher uncertainties for any measured RVs and therefore excluded. DR2 also excluded RVs for cool M-type stars. However, that was not due to broadness but the type of spectral line (TiO). While RVs for many of these stars were still processed, only those that show high precision/accuracy were published. DR2 contains RVs only for late-type stars of medium temperature (FGK-types) with increasing accuracy towards cooler stars in this range ([Sartoretti et al. 2018](#)). Future *Gaia* data releases are expected to include RVs that can currently only be measured with lower accuracy/precision. Rotational velocities are expected to be derived allowing RVs of hotter stars to be better estimated ([Katz et al. 2019](#)).

For the 2D-candidates without RVs, I search through the Simbad/VizieR databases ([Wenger et al. 2000](#); [Ochsenbein et al. 2000](#)) and find several additional RVs in secondary literature sources. I then trace back any 2D-candidates with RV also in yz- and xz-direction and if they positively trace back in these two additional planes, the 2D-candidate becomes a 3D-candidate. For the trace-back, I consider errors in velocity and distance, due to their larger values. Any star that traces back when considering these errors becomes a candidate in 2D or 3D. However, I do not consider the errors in *Gaia* DR2 for the on-sky positions as these are considerably smaller. In practice, this approach leads to any star becoming a candidate if it traces back using a combination of highest, average, lowest velocities and distances (e.g. a star will become a candidate if it traces back using its proper motion plus related uncertainty, its RV minus related uncertainty and its average distance). For all my 2D-candidates, I also search for mass estimates, ages and spectral types from literature sources.

### Handling astrometric errors

The errors in the velocity of my RW and WW candidates are calculated using an approach based on [Kuhn et al. \(2019\)](#). I change how the astrometric errors are calculated compared to [Schoettler et al. \(2020\)](#) due to the inclusion of RVs in the transformation from ICRS to the Cartesian coordinate system.

The basis is a covariance matrix as used in *Gaia* DR2 for the astrometric solution (Equation B.3 in [Lindegren et al. 2018](#)). This covariance matrix considers only the proper motion errors and their correlation. I extend this matrix with the RV errors but do not add any correlations as the RVs and proper motions are not correlated. I then

convert them into errors in the Cartesian coordinate system. I multiply these internal *Gaia* DR2 uncertainties (proper motion) with a correction factor of 1.1 (Lindegren et al. 2018) and also use this correction factor on the RVs. However, I do not consider any corrections for systematic errors in the proper motions. I also include the errors in the motion of the centre of the regions, before finally converting into  $\text{km s}^{-1}$  velocities using distances  $r$  and  $\kappa=4.74$  (conversion factor from  $\text{mas yr}^{-1}$  to  $\text{km s}^{-1}$ ).

### 3.3.3 Constructing CAMDs using photometry

Not all stars that I can trace back to the star-forming regions have originated in those regions. To differentiate these stars, I use CAMDs and PARSEC isochrones (version 1.2S, Bressan et al. 2012) to differentiate older stars from those young enough to have originated in the star-forming regions.

The data sets I use cover regions with a radius of 100 pc around the respective centres, so I convert the apparent  $G$ -band magnitude of each star  $G$  to its absolute magnitude  $M_{G,0}$  using its distance  $r$  (Bailer-Jones et al. 2018) in Eq. 2 in Andrae et al. (2018):

$$M_{G,0} = G + 5 - 5 \log_{10} r - A_G . \quad (3.3)$$

This equation also includes a correction for the extinction in the  $G$ -band,  $A_G$ . In the CAMDs, I then plot this extinction-corrected  $M_{G,0}$  against  $G_{\text{BP}} - G_{\text{RP}}$ , which I correct for reddening,  $E(G_{\text{BP}} - G_{\text{RP}})$  and denote this dereddened intrinsic colour as  $(G_{\text{BP}} - G_{\text{RP}})_0$ .

#### Extinction and reddening correction

The *Gaia* DR2 catalogue includes values for extinction and reddening, but only for a small subset ( $\sim 90$  million) of all sources. For the stars with missing values, I estimate these values following an approach based on Zari et al. (2018). These authors assigned missing values based on the 3D-position of a star using averages of *Gaia* DR2 extinction and reddening values from surrounding stars. Using the transformed Cartesian coordinates for position and distance, I draw a 10 pc sphere around each star with missing values, calculate the average  $A_G$  and  $E(G_{\text{BP}} - G_{\text{RP}})$  values in this sphere from sources with *Gaia* DR2 values and assign these averages to the star in the centre.

The effect of this correction method on my complete data set is shown in Fig. 3.5 for the ONC data set. On the top left, I show the raw data for all stars within 100 pc of the ONC before applying the astrometric and photometric filters. On the top right, I show the data before extinction and reddening correction, but after already having applied

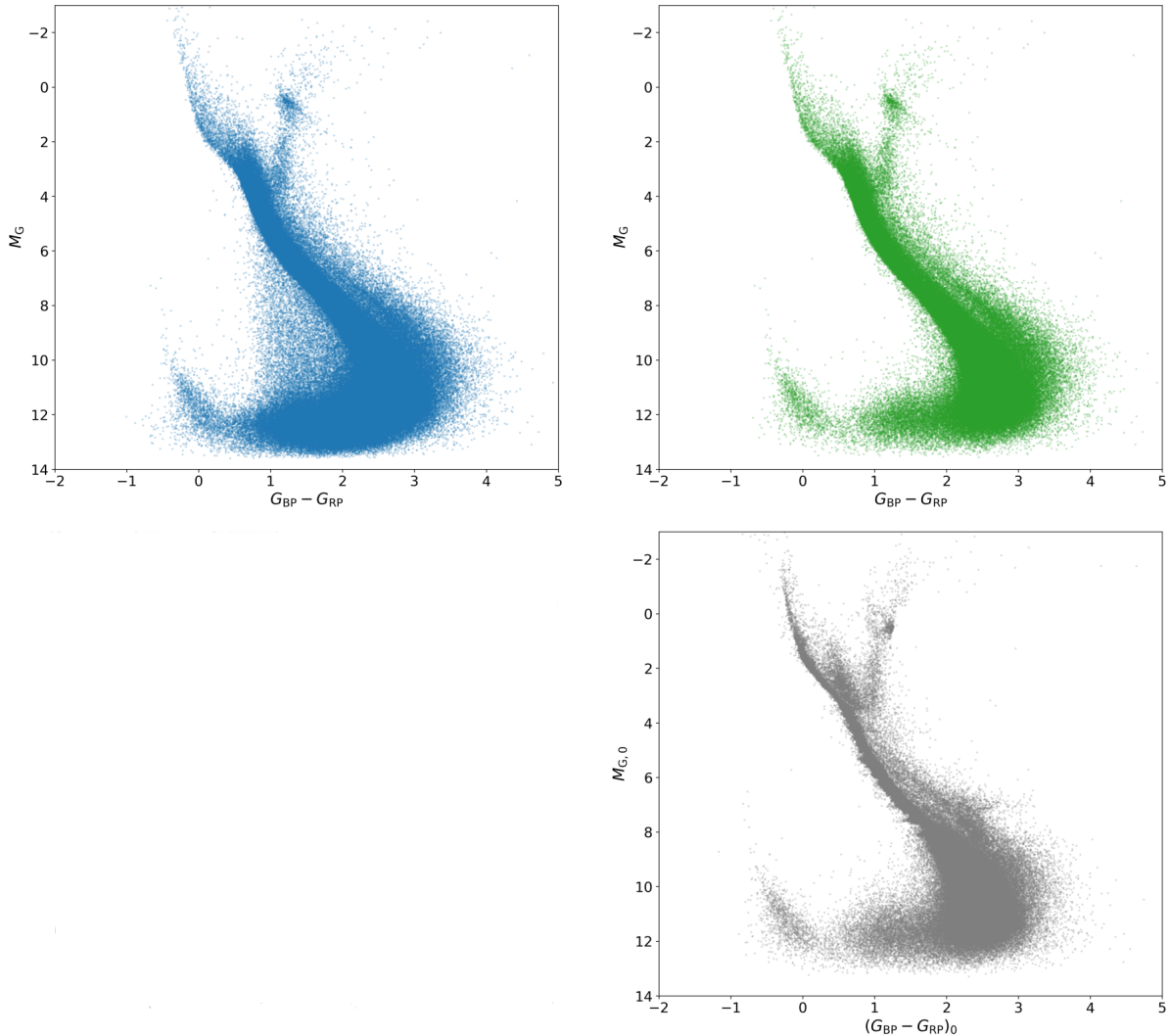


Figure 3.5: CAMD plots showing data before/after applying the astrometric/photometric filters and before/after the extinction and reddening correction, based on the data for the ONC as an example. Top left: Raw data showing all stars within 100 pc of the ONC, before applying the astrometric and photometric filters. Top right: Data before extinction and reddening correction, after applying the filters to the data in the top plot. This CAMD clearly shows at least two differently aged populations of stars with an obvious main-sequence and at least one younger population of stars above it. Bottom right: ONC data set after applying extinction and reddening correction. The correction sharpens the MS with a younger population of stars visible above it. Note: the horizontal stripes along the MS (in particular for  $M_{G,0}$  between 6 and 9 mag) are artefacts and present only for stars with *Gaia* DR2 extinction and reddening (Andrae et al. 2018). Figure adapted from Schoettler et al. (2020).

the astrometric and photometric filters described in section 3.3.1. This CAMD clearly shows at least two differently aged populations of stars with an obvious main-sequence and at least one younger population of stars above it.

On the bottom right, I show the data after correction. The MS emerges more clearly compared to the top left panel and a secondary brighter track is visible above it. This secondary track will include stars younger than the MS; however, it can also include older binaries. My choice of a low RUWE-value should reduce the likelihood of older binaries contaminating this region. I cannot fully remove the risk as the RUWE filter will only remove binaries, where the astrometric quality is compromised by the binary status, e.g. systems with larger orbital periods (e.g. [Penoyre et al. 2020](#); [Belokurov et al. 2020](#)).

In this bottom right panel, there are horizontal stripes along the MS, most obvious at an absolute magnitude of 6 to 9 mag. These are a result of the way the PARSEC evolutionary tracks were sampled in *Gaia* DR2 and are artefacts ([Andrae et al. 2018](#)). This affects only stars with *Gaia* DR2-provided  $A_G$  and  $E(G_{BP} - G_{RP})$  values and not those calculated by my method of averaging over neighbouring stars.

### Handling photometric errors

I calculate photometric errors in the G-magnitude,  $G_{BP}$  and  $G_{RP}$ . No errors are provided for these quantities in *Gaia* DR2 as the error distribution is only symmetric in flux space ([Hambly et al. 2018](#)). This converts to an asymmetric error distribution in magnitude space which cannot be represented by a single error value. The G-magnitude in *Gaia* DR2 is calculated following equation 5.20 in the Gaia DR documentation ([Busso et al. 2018](#)) adding a zero point,  $G_0$  to the instrumental  $G$ -magnitude value:

$$G = G_{\text{instr}} + G_0 \quad (3.4)$$

Using equation 5.26 in the Gaia DR documentation ([Busso et al. 2018](#)) allows me to calculate the error in  $G$ :

$$\sigma_G = \sqrt{\left(1.0857 \frac{\sigma_{\bar{I}}}{\bar{I}}\right)^2 + (\sigma_{G_0})^2} \quad (3.5)$$

In this equation  $\sigma_{\bar{I}}$  represents the error from internal calibration, labelled in the data as `phot_g_mean_flux_error` and  $\bar{I}$  represents the weighted mean flux, labelled in the data as `phot_g_mean_flux` in *Gaia* DR2.  $\sigma_{G_0}$  represents the passband error in the zero

point.

I calculate the errors for  $G$ -magnitude,  $G_{\text{BP}}$  and  $G_{\text{RP}}$  using the passband errors in the zero points in the VEGAMAG system (Evans et al. 2018). I then transform the apparent  $G$ -magnitude errors into absolute  $M_G$ -errors using distances from Bailer-Jones et al. (2018). I also consider errors in these distances and calculate the errors for  $G_{\text{BP}} - G_{\text{RP}}$  and then also include the errors in extinction and reddening.

For stars with *Gaia* DR2 data for extinction  $A_G$  and reddening  $E(G_{\text{BP}} - G_{\text{RP}})$ , the data also provide upper and lower percentile values, which I use to calculate upper and lower errors. For stars with averaged correction values, I take the standard deviation of the values that I averaged over to calculate extinction and reddening errors. The final photometric errors are dominated by errors in extinction,  $A_G$  and reddening,  $E(G_{\text{BP}} - G_{\text{RP}})$  with smaller contributions from errors in distance, apparent magnitude and colour.

### 3.3.4 Age estimates using PARSEC isochrones

I use upper age limits for the star-forming regions analysed in this work, all of which are 5 Myr old or below. I use PARSEC isochrones (version 1.2S, Bressan et al. 2012) to separate the stars into two age brackets. Younger stars are either fully located above the isochrone or when located below it they have error bars crossing the age boundary. Older (MS) stars are located fully below the isochrone without their error bars crossing it. I download data<sup>6</sup> to produce an isochrone using a linear age equal to the upper age limit and select an Initial Mass Function (IMF) option similar to the one I use in my simulations, i.e. a combination of Chabrier (2001) and Salpeter (1955).

I do not adjust for extinction in the isochrones, as I consider this in the stellar data and use the Maíz Apellániz & Weiler (2018) passbands. Other passband options available for *Gaia* DR2 are Weiler (2018) and Evans et al. (2018). Regardless of which of these three passbands options I apply to my data, the results do not change. The isochrones chosen for this work (with passbands from Maíz Apellániz & Weiler 2018) result in a better fit to the higher-mass end of the CAMD, where the stars have already reached the MS. In addition to a choice of IMF and passbands, the isochrones can also include adjustments to the metallicity of each of the regions. I chose metallicities specific to each region and comment on the values in the respective chapters.

Fig. 3.6 shows the CAMD for all stars within 100 pc of the ONC. PARSEC isochrones as described above with different linear ages are plotted. This figure shows that at young ages only the brighter stars would have already reached the MS. For isochrones repre-

<sup>6</sup><http://stev.oapd.inaf.it/cgi-bin/cmd>

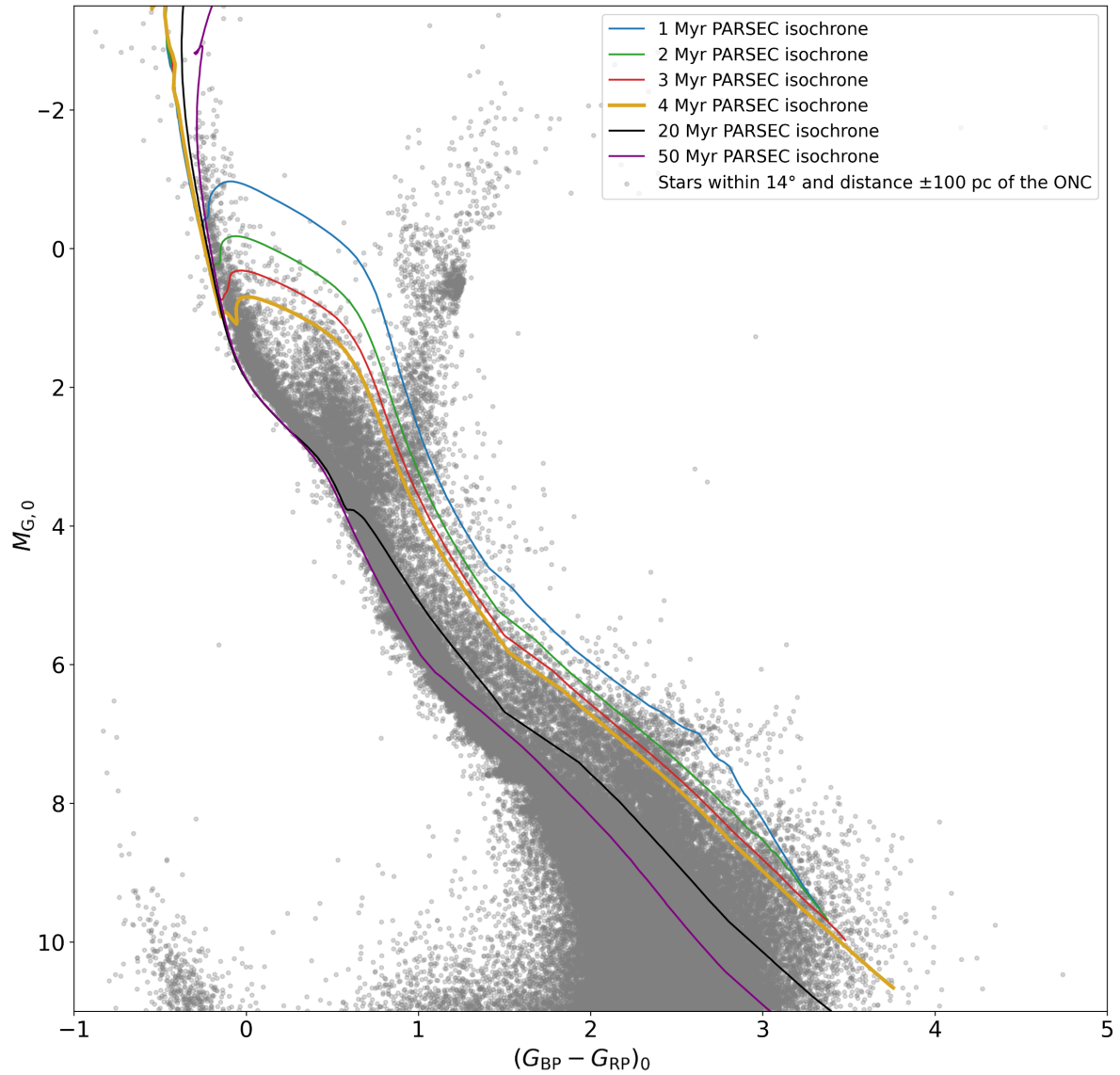


Figure 3.6: CAMD plot of the ONC data after filtering and correction for data analysed in section 5. The plot includes PARSEC isochrones (Bressan et al. 2012) for different linear ages with the metallicity of the ONC and the Maíz Apellániz & Weiler (2018) passbands.

sentative of older ages, the isochrone starts to trace the MS for increasingly fainter stars, i.e. lower-mass stars that take longer to reach the MS. This illustrates that isochrones assist in estimating ages for stars in my data set.



## Chapter 4

# Unbound stars in N-body simulations

## 4.1 Introduction

The project in this chapter is based on earlier work that showed that information from the spatial distribution of star-forming regions can be used to distinguish the initial bound/unbound state (initial virial ratio) (Parker et al. 2014, Paper I). Parker & Wright (2016, Paper II) showed that using the radial velocity dispersion in combination with a spatial structure diagnostic ( $Q$ -parameter, Cartwright & Whitworth 2004) can help constrain initial conditions in star-forming regions with high local densities. In this chapter, which has been published in Schoettler et al. (2019, Paper III of the above series), I use  $N$ -body simulations with differing initial conditions to investigate if the number and velocity distributions of unbound stars can allow me to place constraints on the initial density and velocity structure in star-forming regions. I aim to make predictions for observations of fast unbound stars from young star-forming regions that can be probed with *Gaia* DR2.

My simulations provide me with 6D-parameter space results (position and velocity), but I focus on the 2D-plane and 2D-velocity, i.e. the tangential velocity, which is calculated from proper motion and distance (or parallax) in observations. I follow this approach, as *Gaia* DR2 provides five-parameter astrometry for most of their sources and only a small percentage has six-parameter astrometry (Gaia Collaboration et al. 2018a).

This chapter is organised as follows. First, I present the initial conditions used for the  $N$ -body simulations and my definition of unbound stars. The  $N$ -body simulation set-up is mainly described in Chapter 2, and I only comment on parameters here that are unique to this analysis. After the results section, I follow this with a short discussion and conclude with some final remarks related to this work.

## 4.2 Methods

### 4.2.1 Initial conditions

My simulated star-forming regions are set up with 1000 systems per simulation distributed across an initial radius of 1 pc. All systems are initially single stars (no primordial binaries), and their masses are randomly sampled for every single simulation from the Maschberger (2013) IMF. I sample stellar masses  $m$  between  $0.1 M_{\odot}$  (I do not include BDs) and  $50 M_{\odot}$ , resulting in average total masses of  $\sim 600 M_{\odot}$  for each of my star-forming regions.

I use a set of four different fractal dimensions  $D$  for my simulations to investigate

a wide parameter space. Starting with highly substructured star-forming regions ( $D = 1.6$ ), I then gradually reduce the level of substructure ( $D = 2.0$  and  $D = 2.6$ ), finishing with a roughly uniform, smooth sphere ( $D = 3.0$ ).

In my parameter space, I investigate star-forming regions initially in virial equilibrium as well as two regions that are initially subvirial ( $\alpha_{\text{vir}} = 0.1$  and  $\alpha_{\text{vir}} = 0.3$ ) and two supervirial ( $\alpha_{\text{vir}} = 1.0$  and  $\alpha_{\text{vir}} = 1.5$ ) initial settings. These global virial ratios describe the bulk motion of the stars as a whole. On local scales stars have similar, correlated velocities, meaning star-forming regions can be locally subvirial even if they are not subvirial on a global scale. This can lead to local, but not global collapse during the early dynamical evolution of the star-forming region (e.g. Allison et al. 2010; Parker & Wright 2016).

With four different initial fractal dimensions and five different initial virial ratios, I run 20 simulations of each of the 20 combinations for a time period of 10 Myr to cover the early phases of the evolution of a star-forming region. The only changes within the simulations sharing the same initial conditions are the random number seed used to initiate the creation of the fractal (i.e. initial positions and velocities of stars) and the sampling of stellar masses from the IMF. For each set of initial conditions, I combine the results of all 20 simulations, thus creating a larger data set for analysis.

The star-forming regions do not have a gas potential and there is no external/tidal field applied. The stars do not undergo stellar evolution and are not in primordial binaries or initially mass-segregated. This allows me to identify the effects of different initial spatial and velocity substructure on the unbound population from young star-forming regions.

### 4.2.2 Unbound stars and fractions by mass class

I consider a star  $i$  to be unbound once it has positive total energy (i.e. its kinetic energy  $T_i$  is larger than the modulus of its potential energy  $\Omega_i$ ). Its kinetic energy is given by:

$$T_i = \frac{1}{2}m_i|\mathbf{v}_i - \mathbf{v}_{cr}|^2, \quad (4.1)$$

where  $m_i$  is the mass of star  $i$  and  $\mathbf{v}_i$  and  $\mathbf{v}_{cr}$  are the velocity vectors of this star and of the centre of the region, respectively. The potential energy of the star  $i$  is given by the

sum of the potential energy between star  $i$  and every other star  $j$ :

$$\Omega_i = - \sum_{i \neq j} \frac{Gm_i m_j}{r_{ij}}, \quad (4.2)$$

where  $G$  is the gravitational constant,  $m_i$  and  $m_j$  are the stellar masses of  $i$  and  $j$  and  $r_{ij}$  is the distance between them. For the regions simulated for this project, the escape velocity is  $\sim 3 \text{ km s}^{-1}$  at which point my unbound stars will have a total positive energy.

After identifying all unbound stars in each snapshot, I divide them up into two mass classes (MC): low/intermediate-mass ( $< 8 M_\odot$ ) and high-mass ( $\geq 8 M_\odot$ ) stars. I then calculate unbound fractions by normalising the number of unbound stars (UB) by the total number of stars (TOT) in that specific mass class:

$$\text{Unbound fraction} = \frac{N_{\text{MC,UB}}}{N_{\text{MC,TOT}}} \quad (4.3)$$

I estimate the standard error of the mean (SE) as a representation of the uncertainty connected to the unbound fractions, where  $s$  is the sample standard deviation and  $n$  is the number of simulations:

$$SE = \frac{s}{\sqrt{n}} \quad (4.4)$$

The uncertainty is caused by the stochastic nature of the underlying dynamical evolution (e.g. [Allison et al. 2010](#); [Parker et al. 2014](#)). In my parameter space study, this different evolution is evident in the different unbound fractions from statistically identical, individual simulations as shown in Fig. 4.1. This figure illustrates how different the unbound fractions considering unbound stars of all masses can be for ten simulations with the same initial conditions (initially subvirial  $\alpha_{\text{vir}} = 0.3$  and a high level of substructure  $D = 1.6$ ). The different lines represent the fractions of all unbound stars as a function of time and in this example, they can increase over the simulation time to values between  $\sim 18$ - $48\%$  after 10 Myr. With smaller sample sizes, e.g. when only considering the rarer higher-mass stars, the unbound fractions for these stars differ to an even larger percentage. In different simulations, the unbound fraction were as low as  $0\%$  or as high as  $80$ - $90\%$ . Fig. 4.4 shows the unbound fractions for different mass classes with different sample sizes and illustrates this with a much higher standard error associated with the unbound fractions from the smaller sample size (high-mass stars).

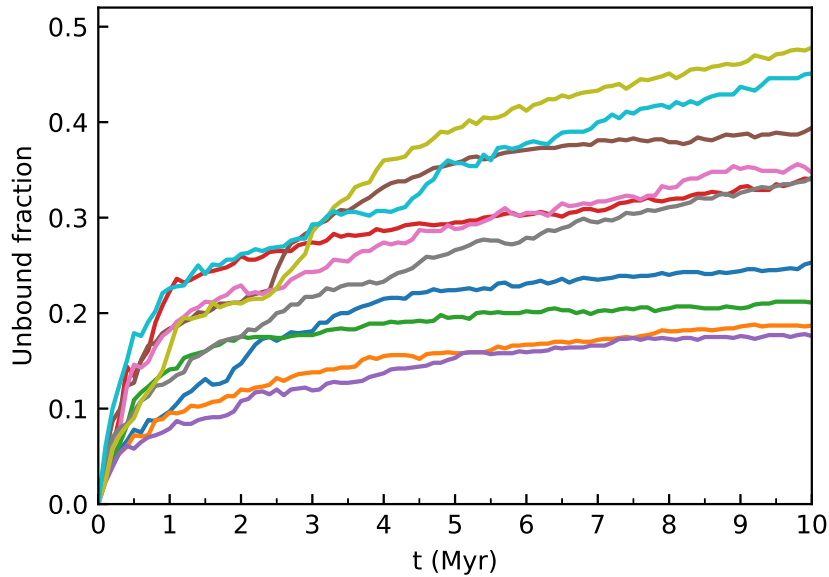


Figure 4.1: Unbound fractions from ten simulations (initially subvirial  $\alpha_{\text{vir}} = 0.3$ , with high level of initial substructure  $D = 1.6$ ) showing the spread of the unbound fractions between statistically identical simulations. These fractions include all unbound stars in the simulation regardless of their mass.

## 4.3 Results

For the following analysis of the velocities, I focus on 2D-velocities to allow me to make predictions for proper motion observations, such as from *Gaia* DR2 (Gaia Collaboration et al. 2018a). In observations, there is a fixed two-dimensional plane, whereas the choice of the 2D-plane from simulations is arbitrary. The 2D-velocity results shown in this section represent the tangential velocity in the xy-plane (i.e. calculated as the motion across the sky would be in observations). However, any other choice of 2D-plane gives the same results after considering statistical noise.

### 4.3.1 Cumulative 2D-velocity distributions of all stars

I first focus on the cumulative distributions of the 2D-velocities and analyse how these evolve over the time period covered by my simulations. For each set of initial conditions, the cumulative distributions contain all stars from 20 simulations. In Fig. 4.2, I show the evolution of the cumulative distributions of the 2D-velocity at four different times from the left to the right column (0, 1, 5 and 10 Myr). From the top row to the bottom, I show that the cumulative velocity distributions for all five initial virial ratios

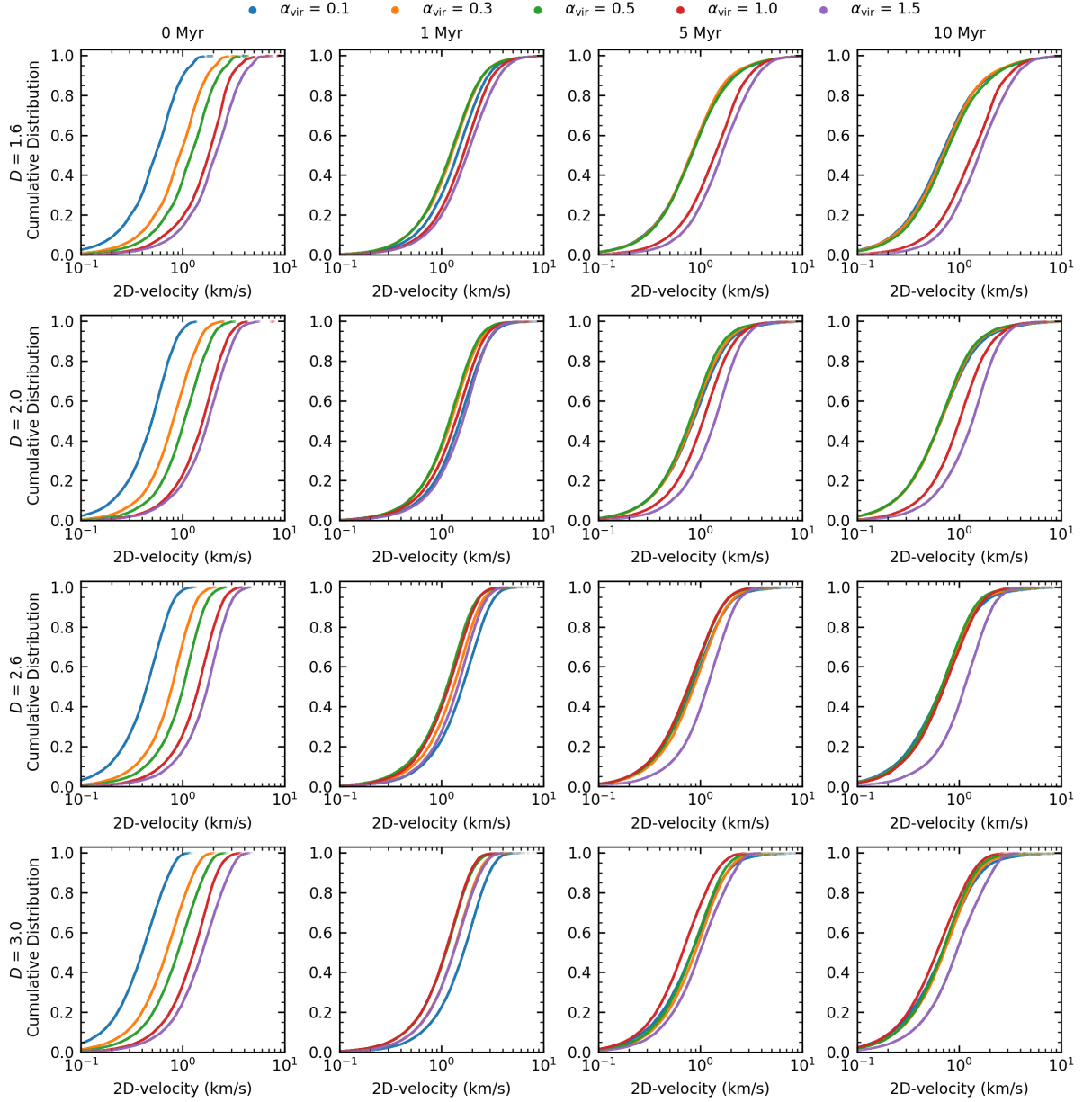


Figure 4.2: Cumulative 2D-velocity distributions at four different simulation times (in columns: 0 Myr, 1 Myr, 5 Myr, 10 Myr) and for the different initial condition sets. Each row represents a different fractal dimension from  $D = 1.6$  (top row) to  $D = 3.0$  (bottom row). The five different initial virial ratios ( $\alpha_{\text{vir}} = 0.1$  (blue),  $\alpha_{\text{vir}} = 0.3$  (orange),  $\alpha_{\text{vir}} = 0.5$  (green),  $\alpha_{\text{vir}} = 1.0$  (red),  $\alpha_{\text{vir}} = 1.5$  (purple)) are shown in each panel for each fractal dimension and time.

are almost identical at 0 Myr regardless of which of the four different fractal dimensions are used. This is to be expected as the virial ratio acts as a scaling factor for the initial velocities and does not depend on the initial spatial distribution. The same virial ratio will therefore create similar cumulative velocity distributions.

During the first 1 Myr, star-forming regions that are initially highly to moderately substructured ( $D \leq 2.0$ ) collapse and undergo violent relaxation (e.g. Lynden-Bell 1967; Van Albada 1982; Funato et al. 1992; McMillan et al. 2007; Moeckel & Bonnell 2009b; Allison et al. 2010; Spera & Capuzzo-Dolcetta 2017) with subvirial regions ( $\alpha_{\text{vir}} < 0.5$ ) collapsing rapidly to form bound, spherical clusters (e.g. Parker et al. 2014). Some of the initially virialized regions ( $\alpha_{\text{vir}} = 0.5$ ) undergo a local collapse in regions of high substructure. Even though they are initially virialized on a global scale, they can be subvirial locally resulting in a localised collapse. Star-forming regions with little or no initial substructure ( $D \geq 2.6$ ) collapse only when they are also initially subvirial.

At 1 Myr (second column), the velocity distributions of different initial virial ratios show similar velocities for identical levels of initial substructure. Initially highly subvirial regions ( $\alpha_{\text{vir}} = 0.1$ ) that are slowest at the start of the simulations attain similar velocities to initially virialized and supervirial regions when  $D \leq 2.0$  or higher velocities when  $D \geq 2.6$ . Violent relaxation leads to an increase in velocity, which is highest in highly subvirial, substructured initial conditions.

After 5 Myr and 10 Myr (third and fourth column), in initially more substructured regions ( $D \leq 2.0$ ) the evolution of the cumulative distributions follows a similar pattern. The bound, initially subvirial or virialized regions ( $\alpha_{\text{vir}} \leq 0.5$ ) have very similar velocity distributions as the initially subvirial regions approach virial equilibrium after violent relaxation. Initially supervirial regions ( $\alpha_{\text{vir}} > 0.5$ ) remain unbound and at higher average velocities. The difference between the subvirial/virial and supervirial distributions becomes clearer the older the simulated regions get, as the initially subvirial/virialized regions slow down compared to the initially supervirial ones.

Star-forming regions with less substructure initially ( $D \geq 2.6$ ) do not show the clear separation of velocity distributions between subvirial/virial and supervirial initial ratios. Only initially highly supervirial regions ( $\alpha_{\text{vir}} = 1.5$ ) have a velocity distribution at later times that can be distinguished from those with lower virial ratios. The initially smooth, sphere-like regions ( $D = 3.0$ ) still show a grouping together of the velocity distributions after 5 Myr. The two initially supervirial distributions ( $\alpha_{\text{vir}} = 1.0$  and  $1.5$ ) are located either side of the initially subvirial and virialized ones. Despite both being supervirial, they exhibit considerably different velocity distributions. Moderately supervirial regions

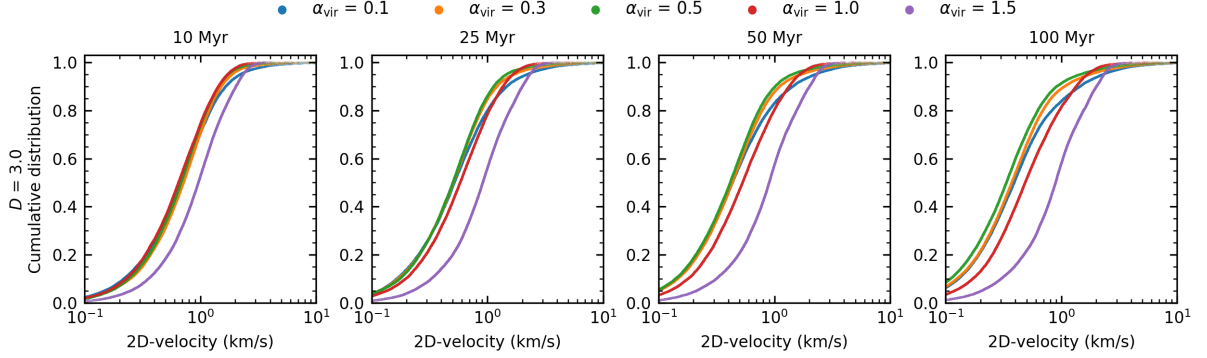


Figure 4.3: Long-term evolution of the cumulative 2D-velocity distributions at four different simulation times (10, 25, 50 and 100 Myr) for the five different initial virial ratios ( $\alpha_{\text{vir}} = 0.1$  (blue),  $\alpha_{\text{vir}} = 0.3$  (orange),  $\alpha_{\text{vir}} = 0.5$  (green),  $\alpha_{\text{vir}} = 1.0$  (red),  $\alpha_{\text{vir}} = 1.5$  (purple)) and constant fractal dimension ( $D = 3.0$ ).

( $\alpha_{\text{vir}} = 1.0$ ) have the slowest, whereas highly supervirial regions ( $\alpha_{\text{vir}} = 1.5$ ) have the fastest cumulative 2D-velocities. This behaviour continues for the remaining 5 Myr and at the end of my simulations the moderately supervirial cases are still indistinguishable from those of initially subvirial/virialized ( $\alpha_{\text{vir}} \leq 0.5$ ) cases.

### Long-term evolution of initially smooth star-forming regions

For these initially smooth star-forming regions ( $D = 3.0$ ), I follow the evolution of their cumulative distributions for a longer time period. I evaluate if they evolve differently or just more slowly than initially more substructured star-forming regions. The evolution of these smooth regions is shown at 10 Myr, 25 Myr, 50 Myr and 100 Myr in Fig. 4.3.

The cumulative distributions for initially subvirial and virialized regions ( $\alpha_{\text{vir}} \leq 0.5$ ) continue to be similar as they are in a state of virial equilibrium. The velocity distribution for the moderately supervirial regions ( $\alpha_{\text{vir}} = 1.0$ , red) starts to become distinguishable from the initially subvirial/virialized regions after 50 Myr, as these regions slow down compared to the moderately supervirial one.

But even after 100 Myr, the velocities of moderately supervirial regions are still much closer to those of initially subvirial/virial star-forming regions than the highly supervirial scenario. Initially smooth, supervirial star-forming regions appear to evolve in a similar fashion than the more substructured regions but on a much longer timescale. The long-term evolution of the cumulative distributions shows that the average velocities decrease at later times for initially subvirial/virialized regions, as the global gravitational field of the bound clusters cause stars to decelerate.



### 4.3.2 Unbound fractions of stars from initially subvirial and virialized regions

In this section, I turn to unbound fractions for initially subvirial and virialized star-forming regions ( $\alpha_{\text{vir}} \leq 0.5$ ). I exclude the two supervirial scenarios as in these globally unbound, expanding regions most stars are born unbound. In my simulations, I do not have any stellar evolution, so stars can only become unbound due to dynamical interactions with other stars (DES, [Poveda et al. 1967](#)) and not from SN kicks (BSS, [Blaauw 1961](#)). In the absence of an external tidal field, lower-mass stars mainly become unbound due to effects of two-body relaxation (e.g. [Binney & Tremaine 2008](#)), whereas high-mass stars require dynamical interactions with other high-mass stars in binaries or higher order multiple systems (e.g. trapezium-like) to become unbound (e.g. [Allison & Goodwin 2011](#); [Parker et al. 2016](#)).

#### Effects of different levels of substructure in regions with the same initial virial ratio

In Fig. 4.4, unbound fractions for star-forming regions with an initially highly subvirial ratio ( $\alpha_{\text{vir}} = 0.1$ ) are shown in the first column. With high levels of initial substructure ( $D = 1.6$ , first row) stars in both mass classes show similar unbound fractions from 5 Myr to the end of the simulations. These regions, regardless of initial degree of substructure, will undergo rapid collapse and violent relaxation. While low/intermediate-mass stars become unbound early in the simulations, high-mass stars show a more gradual increase and match the lower-mass unbound fraction at  $\sim 5$  Myr. The lower-mass unbound fraction decreases with less initial substructure and settles on the same level of  $\sim 20\%$  for more moderate amounts of initial substructure ( $D = 2.0$ - $3.0$ ) after 10 Myr. At the end of my simulations, the high-mass unbound fractions in the four initial substructure scenarios reach final values between  $22 \pm 3\%$  and  $28 \pm 4\%$ .

I see a delay in high-mass stars becoming unbound that increases the lower the level of initial substructure (i.e. higher fractal dimension  $D$ ) in initially highly subvirial regions ( $\alpha_{\text{vir}} = 0.1$ , first column). In these simulations, the degree of collapse reduces with lower amounts of initial substructure, resulting in a longer formation time for multiple star systems that can eject massive stars. The low/intermediate-mass stars also show a delay in stars becoming unbound for  $D = 2.6$ - $3.0$ . The delay is most obvious in regions with no initial substructure ( $D = 3.0$ , bottom row). On average, only 7 stars (all are low/intermediate-mass) per simulation become unbound in the first  $\sim 0.5$  Myr. The lack of initial substructure combined with the low initial virial ratio appears to result in a

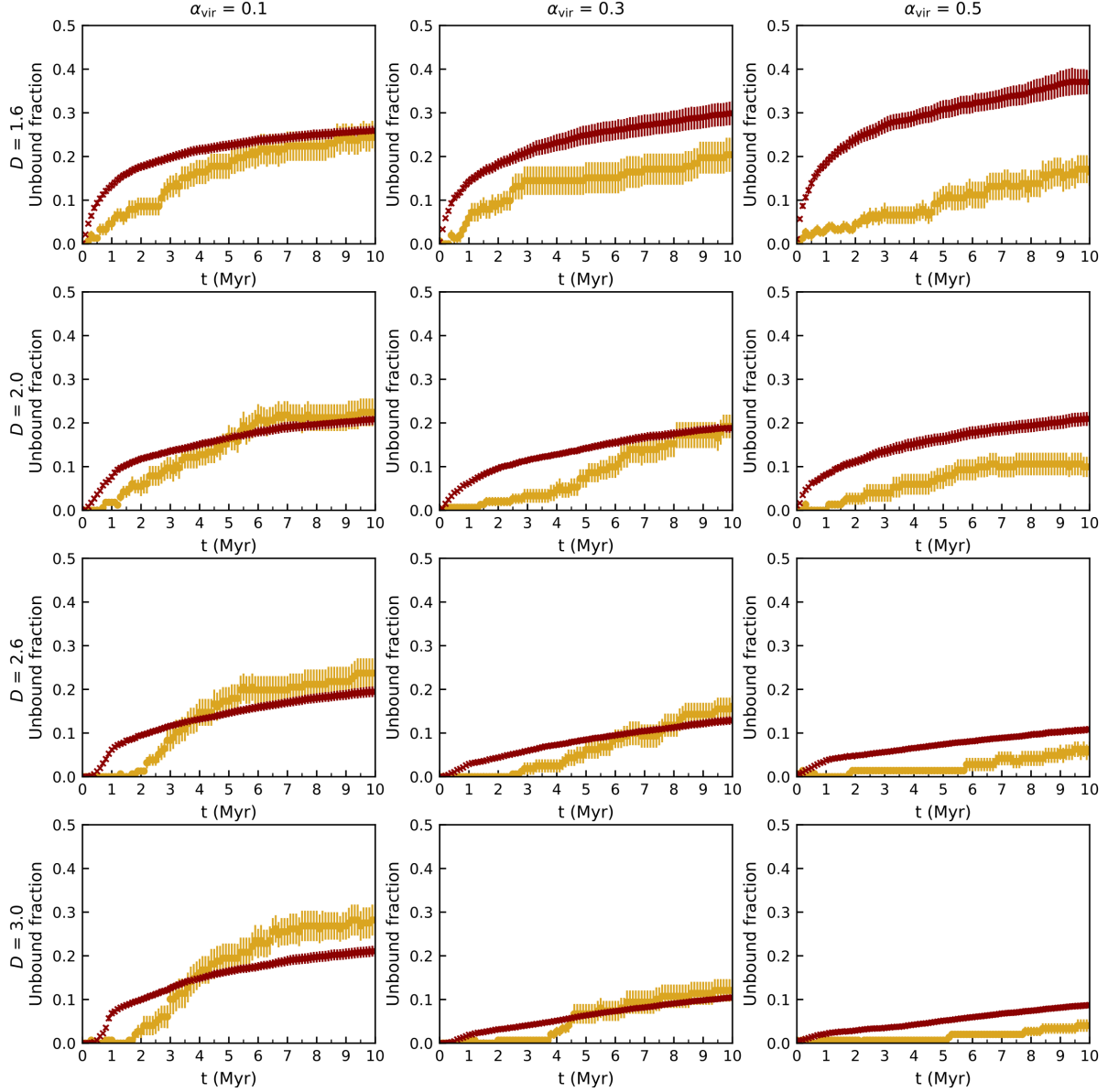


Figure 4.4: Unbound fractions by mass class for initially subvirial and virialized star-forming regions ( $\alpha_{\text{vir}} \leq 0.5$ ). Each row represents a different fractal dimension starting from  $D = 1.6$  (top row) to  $D = 3.0$  (bottom row). The columns show the three subvirial and virial initial ratios. The red points represent the unbound fraction of low/intermediate-mass stars ( $< 8 M_{\odot}$ ) over the simulation time, whereas the yellow points represent the unbound fraction of high-mass stars ( $> 8 M_{\odot}$ ). The uncertainties of the fractions are calculated using the standard error of the mean (Eq. 4.4).

“balanced” collapse that keeps virtually all stars in a bound configuration for a short period of time ( $\sim 0.5$  Myr).

In initially moderately subvirial simulations ( $\alpha_{\text{vir}} = 0.3$ , second column) the star-forming regions undergo an initial collapse, but the degree of collapse is lower when compared to the highly subvirial simulations. I decrease the level of initial substructure and see a significant decrease in the low/intermediate-mass unbound fraction for every change in fractal dimension. After 10 Myr, the high-mass unbound fractions only slightly decrease (from  $20 \pm 4\%$  to  $19 \pm 3\%$ ) for regions with initially high or moderate levels of substructure ( $D \leq 2.0$ ). Further decreasing the initial substructure reduces the high-mass unbound fraction to  $16 \pm 2\%$  ( $D = 2.6$ ) and  $12 \pm 3\%$  ( $D = 3.0$ ). The high-mass unbound fractions are only different for simulations with higher ( $D \leq 2.0$ ) and no initial substructure ( $D = 3.0$ ).

In regions with initial fractal dimensions  $D = 2.0$ - $3.0$ , high-mass stars do not become unbound early in the simulations. The collapse happens fastest in initially highly substructured regions ( $D = 1.6$ ) and high-mass stars can become unbound much earlier than in less substructured star-forming regions. The lower the level of initial substructure, the longer it takes to form dynamical multiples that can eject high-mass stars (e.g. [Allison & Goodwin 2011](#)). My simulations suggest that it can take over 3 Myr longer for high-mass stars to become unbound when there is a lack of initial substructure in moderately subvirial initial conditions (lower, middle panels).

In all simulations that are initially virialized ( $\alpha_{\text{vir}} = 0.5$ , third column), regardless of substructure, the unbound fraction of low/intermediate-mass stars is at least double the fraction of unbound high-mass stars after 10 Myr, which reaches  $16 \pm 3\%$  in initially highly substructured regions ( $D = 1.6$ ). In these star-forming regions,  $37 \pm 3\%$  of all low/intermediate-mass stars become unbound at the end of my simulations.

Initially virialized, highly substructured star-forming regions can collapse locally and binary clusters can form ([Arnold et al. 2017](#)). Binary clusters are a pairing of star clusters that are physically close to each other in space (e.g. [Rozhavskii et al. 1976](#); [Pietrzynski & Udalski 2000](#); [de la Fuente Marcos & de la Fuente Marcos 2009](#); [Priyatikanto et al. 2016](#); [Zhong et al. 2019](#)). They can be a result of the dynamical evolution of young star-forming regions as shown by [Arnold et al. \(2017\)](#). I find these binary clusters at the end of more than half of the 20 simulations and they appear to have an effect on the unbound fractions. The presence of binary clusters lowers the sub-cluster potential energy, effectively creating two smaller clusters with smaller potential wells. In consequence, stars require lower kinetic energy to become unbound. This increases the low/intermediate-mass un-

bound fraction but does not affect the high-mass unbound fraction in the same way. Due to the form of the IMF, there is a much smaller number of high-mass stars present in my simulations. During the localised collapse into binary clusters, these high-mass stars can move to different local regions, reducing the likelihood of creating dynamical multiple systems that can eject high-mass stars from my regions.

I do find a higher unbound fraction for high-mass stars than low/intermediate-mass stars in several simulations with initially virialized, highly substructured conditions ( $\alpha_{\text{vir}} = 0.5$ ,  $D = 1.6$ ) that do not result in the creation of binary clusters. In individual simulations where binary clusters are present, I see a higher-than-average fraction ( $\sim 30\%$  compared to the average value of  $\sim 16\%$ ) of unbound high-mass stars when the low/intermediate-mass unbound fraction is high as well ( $\sim 40\text{-}70\%$ ) or when the absolute number of high-mass stars is high to begin with (i.e. 9 or more high-mass stars per simulation). This increases the chances of forming high-mass dynamical multiple systems, which would lead to more ejections.

Lower levels of initial substructure or smooth regions ( $D = 2.0\text{-}3.0$ ) that are initially virialized ( $\alpha_{\text{vir}} = 0.5$ ) do not form binary clusters (Arnold et al. 2017). In my simulations, this considerably reduces the unbound fractions. Star-forming regions that are initially in virial equilibrium and smooth ( $D = 3.0$ ) undergo very little dynamical evolution and most of the stars ( $\sim 87\%$ ) remain bound throughout the simulations.

### Effects of different initial virial ratios in regions with the same levels of substructure

For star-forming regions with a high degree of initial substructure ( $D = 1.6$ , first row in Fig. 4.4) increasing the initial virial ratio has the opposite effect on the unbound fractions in the two mass classes. The increase in initial kinetic energy (higher virial ratio) in the regions decreases the fraction of unbound high-mass stars whereas it increases the fraction of low/intermediate-mass unbound stars. While an initially highly subvirial region ( $\alpha_{\text{vir}} = 0.1$ ) has the same unbound fraction after 10 Myr in both mass classes, the more virialized a highly substructured region is initially the higher its unbound fraction of low/intermediate-mass stars and the lower its high-mass unbound fraction. The low/intermediate-mass unbound fraction is highest in initially virialized regions due to the presence of binary clusters.

In regions with a lower level of initial substructure ( $D = 2.0$ , second row) differences in initial virial ratio have no effect on the low/intermediate mass unbound fractions, which are virtually the same for all three initial virial ratio scenarios (values between  $19 \pm 1\%$  and  $21 \pm 2\%$ ) at 10 Myr. The high-mass unbound fraction is highest in the initially

most subvirial regions ( $\alpha_{\text{vir}} = 0.1$ ). The degree of collapse is highest here and unstable multiple star systems can form quickly. After about 6 Myr, the high-mass unbound fraction reaches  $21 \pm 3\%$  and starts to level out ( $22 \pm 3\%$  at 10 Myr), suggesting that unstable multiple star systems are no longer present or do not lead to any further high-mass star ejections. The initially more moderate, subvirial ( $\alpha_{\text{vir}} = 0.3$ ) simulations have a similar high-mass unbound fraction than the virialized case in the first  $\sim 6$  Myr of the simulations ( $10 \pm 2\%$  vs.  $9 \pm 2\%$ ). The difference in initial virial ratio ( $\alpha_{\text{vir}} = 0.3$  vs.  $0.5$ ) appears to have no effect on the early evolution of these simulated regions. Later in the simulation, the initially moderately subvirial ( $\alpha_{\text{vir}} = 0.3$ ) regions continue to eject high-mass stars and reach an unbound fraction of  $19 \pm 3\%$  after 10 Myr, which is a similar value than in the highly subvirial case ( $\alpha_{\text{vir}} = 0.1$ ). The high-mass unbound fraction in initially virialized regions levels out after  $\sim 7$  Myr and remains at  $10 \pm 2\%$  to the end of the simulations at 10 Myr.

At low levels of or no initial substructure ( $D = 2.6$  and  $3.0$ , third and fourth row) the low/intermediate-mass unbound fractions are highest when the regions are initially highly subvirial ( $\alpha_{\text{vir}} = 0.1$ ) as these regions collapse initially. Even though the moderately subvirial ( $\alpha_{\text{vir}} = 0.3$ ) regions initially collapse, this does not result in a higher low/intermediate unbound fraction than in the initially virialized regions that do not undergo collapse. When there is little or no initial substructure, star-forming regions will only collapse when the initial virial ratio is subvirial. The collapse increases the likelihood that unstable multiple systems form, which facilitate the ejection of high-mass stars. With higher initial virial ratios, these multiple systems take longer to form or do not form at all. As a result, high-mass stars take longer to become unbound and the final unbound fractions at 10 Myr are lower the more virialized and smooth the regions are initially.

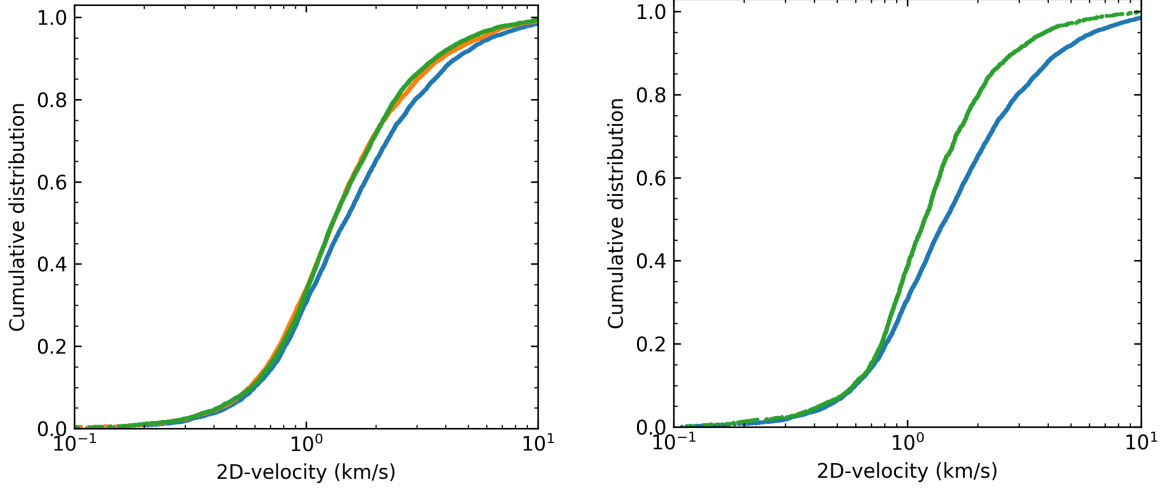


Figure 4.5: Left: Cumulative distributions for unbound stars showing the 2D-velocities at 10 Myr with initial fractal dimension  $D = 1.6$  for initially subvirial and virialized clusters. The distributions for  $\alpha_{\text{vir}} = 0.1$  (blue),  $\alpha_{\text{vir}} = 0.3$  (orange) and  $\alpha_{\text{vir}} = 0.5$  (green) are shown zoomed in to the central part of the curve, highlighting that for these initial conditions, the velocities of the unbound stars do not differ much between different virial ratios for the same degree of initial spatial and kinematic substructure. Right: Cumulative distributions for unbound stars showing the 2D-velocities at 10 Myr for an initially highly substructured and subvirial region with  $D = 1.6$  and  $\alpha_{\text{vir}} = 0.1$  (blue) and an initially almost smooth and virialized region with  $D = 2.6$  and  $\alpha_{\text{vir}} = 0.5$  (green). The comparison illustrates that the more substructured and subvirial a star-forming region is initially, the faster the unbound stars escape.

### 4.3.3 2D-velocity of unbound stars from initially subvirial and virialized star-forming regions

#### Cumulative 2D-velocity distributions

In the left panel of Fig. 4.5, I show the 2D-velocity cumulative distributions for unbound stars from initially subvirial/virialized regions ( $\alpha_{\text{vir}} \leq 0.5$ ) with a high level of initial substructure ( $D = 1.6$ ) at 10 Myr. As I have shown for all (bound and unbound) stars in Fig. 4.2, the cumulative distributions in initially subvirial/virialized simulations are very similar for all four initial fractal dimensions. The cumulative distributions of unbound stars in the left panel of Fig. 4.5 show a similar picture of very similar distributions for a fractal dimension of  $D = 1.6$  (initially highly substructured regions). Even for these initially substructured regions where I see a more dynamic early evolution (i.e. violent relaxation and initial collapse), it is difficult to distinguish between different initial virial ratios at the end of the simulations.

Allison (2012) analysed the spatial and velocity distributions of very different initial conditions after 4 Myr with a smaller, but similar, set of initial conditions. They found that the cumulative velocity distributions differ between the initial conditions and that the initially moderately subvirial, substructured simulations result in higher velocity unbound stars compared with initially virialized simulations with a low level of substructure. The right panel of Fig. 4.5 illustrates the cumulative velocity distributions of very different initial conditions after 10 Myr: initially highly substructured and highly subvirial simulations ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.1$ , blue) compared to simulations with a low level of substructure that are initially virialized ( $D = 2.6$ ,  $\alpha_{\text{vir}} = 0.5$ , green). I also find that unbound stars from substructured, subvirial regions are moving at higher 2D-velocities (after 10 Myr). However, the differences between the distributions are not quite as large as in Allison (2012). This highlights that cumulative velocity distributions can only distinguish between vastly different initial spatial and velocity conditions.

### Violin plots of 2D-velocity distributions

Violin plots are a data visualisation technique that combines a box plot with a density trace or a kernel density estimate (Hintze & Nelson 1998). Like box plots, violin plots also show the median and interquartile range for a variable, as well as any asymmetries and outlier data. They can be useful when comparing distributions of a variable (2D-velocity) over different categories (initial conditions for star-forming regions). Unlike box plots, violin plots include all data from the underlying distribution and give information about its shape. They show all peaks and the position of those peaks, their amplitude, and give insight into any presence of clustering in the data. The outer shape represents all data, with the widest parts corresponding to the value (i.e. 2D-velocity) with the highest probability of occurring in the population (Hintze & Nelson 1998), which can be also interpreted as the most common 2D-velocity in my case.

Fig. 4.6 shows the 2D-velocity distributions on a log-scale for all initially subvirial and virialized regions (left to right) and all four fractal dimensions (decreasing degree of initial substructure - top to bottom) after 10 Myr. The plots include all unbound stars from 20 simulations combined and represent an average. The wider each of the violin plots is at any point, the more stars are likely to have this 2D-velocity. For each fractal dimension (in each row), the width of the violin plot is scaled by the total number of unbound stars for this initial virial ratio. For two violin plots with the same total number of unbound stars, the widest part will have the same width. However, the absolute number of stars with this velocity can be different, e.g. for fractal dimension

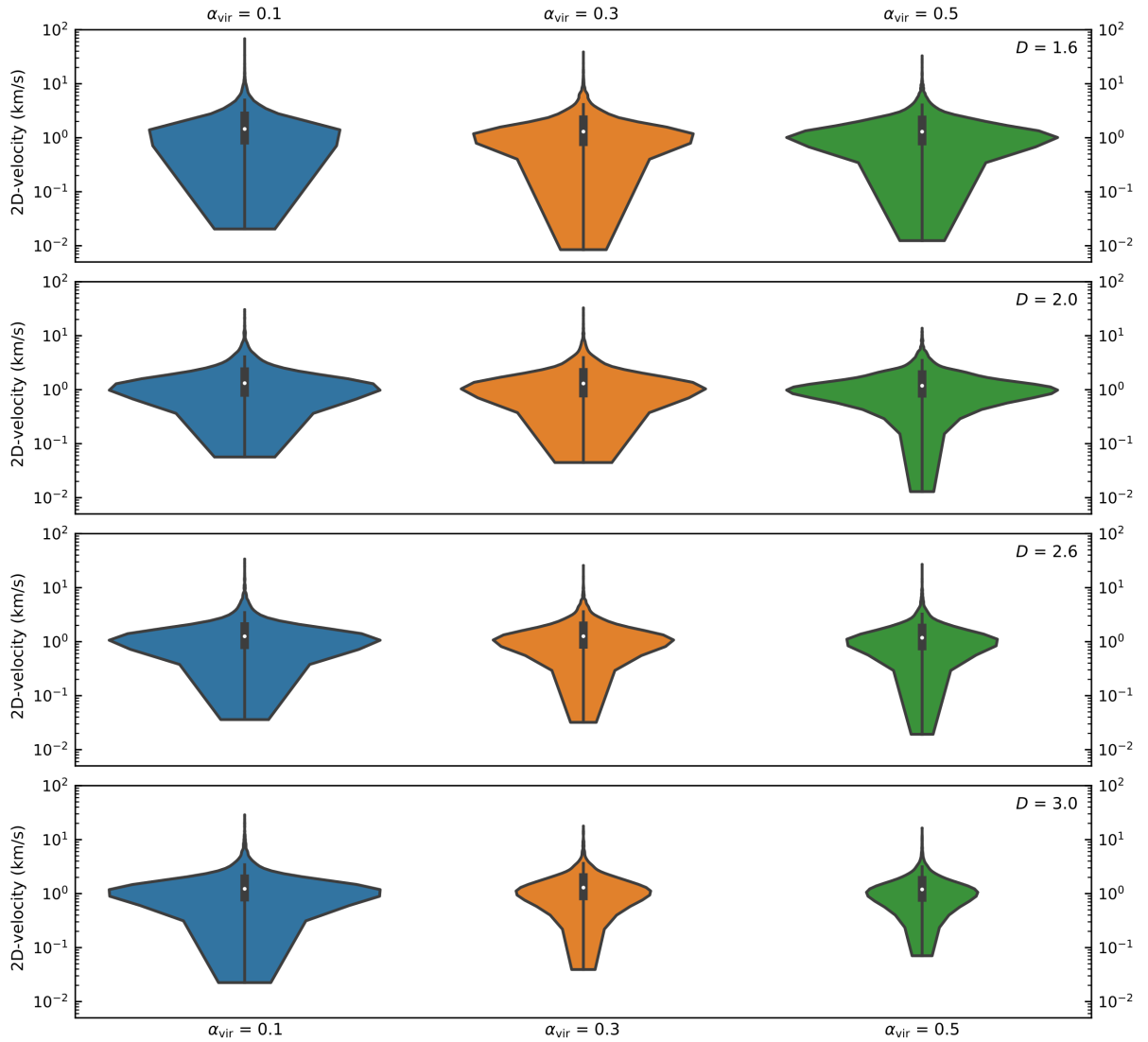


Figure 4.6: Violin plots showing the 2D-velocity distributions of unbound stars at 10 Myr from all initially subvirial ( $\alpha_{\text{vir}} = 0.1$  (blue),  $\alpha_{\text{vir}} = 0.3$  (orange)) and virialized regions ( $\alpha_{\text{vir}} = 0.5$  (green)). The violins are scaled by count, the wider the violins are at any point the more stars in my regions have this 2D-velocity. The larger the violins overall, the more stars have become unbound during the simulation time. The thick vertical bar in the centre shows the interquartile range with the white dot representing the median. The long vertical line represents the 95% confidence interval.



$D = 2.0$  (second row) the blue ( $\alpha_{\text{vir}} = 0.1$ ) and green ( $\alpha_{\text{vir}} = 0.5$ ) violin plots both contain a total of  $\sim 4100$  unbound stars from 20 simulations each, resulting in the widest part of their violin plots having the same width in Fig. 4.6. Due to difference in the distributions, there are  $\sim 80$  more stars at the most common velocity for the initially virialized (green) violin plot.

The thick vertical bar in the centre represents the interquartile range with the white dot representing the median. The thin long vertical line represents the 95% confidence interval. I use a bandwidth following the [Silverman \(1986\)](#) reference rule to smooth my data for the violin plots<sup>1</sup>. The violin plots are cut at the low-velocity end and only show the actual data points there, instead of the tails of the underlying Gaussian kernel density estimate. This allows me to identify the lowest actual 2D-velocity directly from the plot and avoids the appearance of negative 2D-velocities.

Initially highly substructured regions ( $D = 1.6$ , Fig. 4.6, first row) have a large number of unbound stars for all three initial virial ratios. The fastest stars are ejected from initially highly subvirial regions ( $\alpha_{\text{vir}} = 0.1$ , blue) with the peak velocity reaching  $\sim 70 \text{ km s}^{-1}$ . These regions have fewer unbound stars ( $\sim 260$  per simulation) in total and fewer stars at similar velocities with a wider spread of velocities around  $\sim 1 \text{ km s}^{-1}$  compared to the two higher virial ratio scenarios. Despite these differences, the median velocity is similar ( $\sim 1.5 \text{ km s}^{-1}$ ) to the other two scenarios ( $\sim 1.3 \text{ km s}^{-1}$  - both for  $\alpha_{\text{vir}} = 0.3$  and  $0.5$ ). A large number of unbound stars from highly substructured, moderately subvirial regions ( $\alpha_{\text{vir}} = 0.3$ , orange) move at a similar 2D-velocity of  $\sim 1 \text{ km s}^{-1}$  after 10 Myr, creating noticeable arms in the violin plots. The total number of unbound stars increases to  $\sim 300$  per simulation. The arms become most pronounced in the initially virialized case ( $\alpha_{\text{vir}} = 0.5$ , green) with  $\sim 370$  unbound stars per simulation. Despite the increase in the total number of unbound stars, the most common velocity remains around  $\sim 1 \text{ km s}^{-1}$ . The higher the initial virial ratio in initially highly substructured regions, the more likely it is that unbound stars are moving with more similar velocities, whereas unbound stars are more evenly spread over different velocities in initially more subvirial regions.

With a lower level of initial substructure ( $D = 2.0$ , second row) the shape of the distributions changes for all three initial virial ratios. The shape of the velocity distributions of the two initially subvirial scenarios ( $\alpha_{\text{vir}} < 0.5$ ) is now almost identical. The violin plot for highly subvirial regions is wider than the moderately subvirial scenario, meaning more stars become unbound ( $\sim 20$  more stars per simulation). I see the pro-

<sup>1</sup><https://seaborn.pydata.org/generated/seaborn.violinplot.html>

nounced arms in the violin plots now also for highly subvirial regions with less spread in the 2D-velocities. The fastest stars from the two subvirial regions now only reach  $\sim 30 \text{ km s}^{-1}$  and their median velocities are almost identical ( $\sim 1.3 \text{ km s}^{-1}$ ). In initially virialized regions ( $\alpha_{\text{vir}} = 0.5$ , green), the arms in the 2D-velocity become even more pronounced at  $\sim 1 \text{ km s}^{-1}$ . The maximum velocity is lower ( $\sim 13 \text{ km s}^{-1}$ ) than in the subvirial cases ( $\sim 30 \text{ km s}^{-1}$ ); however, the median velocity remains similar ( $\sim 1.2 \text{ km s}^{-1}$ ).

In regions with little or no initial substructure ( $D \geq 2.6$ , third and fourth row), initially highly subvirial regions ( $\alpha_{\text{vir}} = 0.1$ ) show a similar violin shape to the more substructured regions ( $D = 2.0$ ) and also have a similar number of unbound stars ( $\sim 200$ - $210$  unbound stars per simulation). The sizes of the violins shrink considerably (i.e. fewer unbound stars) for initially moderately subvirial ( $\alpha_{\text{vir}} = 0.3$ ) and virialized ( $\alpha_{\text{vir}} = 0.5$ ) regions and I see  $\sim 90$ - $130$  unbound stars per simulation. This indicates a much less dynamical early evolution with the number of unbound stars only  $\sim 30\%$  of what they are in simulations with the highest level of initial substructure. Despite this, the violins retain their overall familiar shape of having arms around the most common velocity of  $\sim 1 \text{ km s}^{-1}$  and a median velocity ( $\sim 1.2$ - $1.3 \text{ km s}^{-1}$ ), which is similar to all other initial conditions.

As stated earlier, my unbound definition is based on stars reaching the escape velocity, which is  $\sim 3 \text{ km s}^{-1}$  in the regions simulated here. In Fig. 4.6, I see that the minimum 2D-velocity of unbound stars can be as low as  $\sim 0.03 \text{ km s}^{-1}$  after 10 Myr. Once unbound stars leave the denser parts of a star-forming region, they interact with fewer or no other stars and slow down gradually. However, the apparent slow-down in my simulations by up to two orders of magnitude is likely due to projection effects. The left panel in Fig. 4.7 shows violin plots for two, very different initial conditions (blue - initially highly subvirial, substructured and green - initially virialized, no substructure) at three different times during the simulations. Already after 1 Myr, a low-velocity tail forms in 2D-space that extends to velocities an order of magnitude lower than the escape velocity. In full 3D-velocity space in the right panel in Fig. 4.7, I see that after 1 Myr the lowest velocities are only  $1$ - $2 \text{ km s}^{-1}$  lower than the escape velocity. This suggests that unbound stars slow down; however, not to the extent suggested by the 2D-velocities.

This 2D projection effect could affect cluster membership identification when observing proper motion (or 1D-radial velocity) in isolation. Depending on relative position to the cluster, these “slow” unbound stars could be identified as not having originated from the cluster at all due to being too far away or still bound due to their central location in the star-forming region. However, my simulations suggest that only  $\sim 1\%$  of these

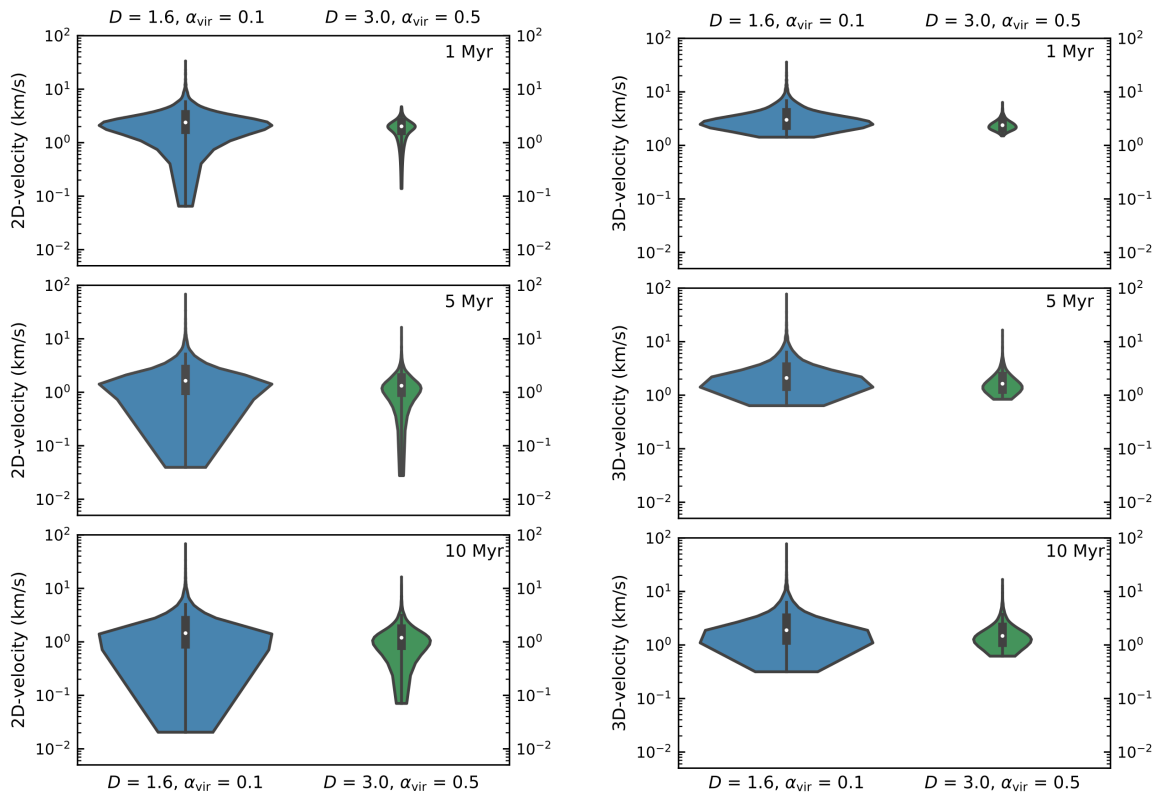


Figure 4.7: Left: Violin plots showing 2D-velocity (xy-plane) distributions of unbound stars at three simulation times for two selected initial conditions (initially subvirial, substructured (blue) and initially virialized with no substructure (green)). Right: Violin plots showing 3D-velocity distributions of unbound stars at three simulation times for two selected initial conditions (initially subvirial, substructured (blue) and initially virialized with no substructure (green)).

unbound stars with low 2D-velocities are located in the central parts of star-forming regions after 10 Myr. This limits the extent of mistakenly assigning membership to “slow” unbound stars, when only proper motion information is available.

In Fig. 4.8, I use split violin plots to show the 2D-velocities separately for the two mass-classes. The plots are now scaled to the same width as I have at most  $\sim 40$  unbound high-mass stars compared with over 7000 lower-mass unbound stars from a set of 20 simulations. The widest part of each half still represents the 2D-velocity with the highest probability of occurring. Dashed lines represent the median and the interquartile range, the 95% confidence interval is no longer identified on the plots. The violin plots are again cut at the low-velocity end and only show the actual data points, instead of the tails of the underlying Gaussian kernel density estimate. This allows me to identify the lowest actual 2D-velocity directly from the plot and avoids the appearance of negative 2D-velocities.

In Fig. 4.8, I find that the shape of the low/intermediate-mass violins is nearly identical to the shape of the total population of unbound stars in Fig. 4.6 as most unbound stars are lower mass. Due to the low number of unbound high-mass stars the velocity distributions of unbound high-mass stars can have a jagged outline depending on the bandwidth used. I use the same bandwidth setting (following [Silverman 1986](#)) as in Fig. 4.6 resulting in the right half (unbound high-mass stars) of my split violin plots in Fig. 4.8 appearing as a smooth distribution despite the small sample size. A small sample size can make conclusions from violin plots unreliable and I limit my interpretation of them to general differences in median, minimum and maximum velocity between the two mass-classes. To gain more insight into the velocity distributions of unbound high-mass stars using violin plots would require an increase in the sample size, i.e. a much higher number of simulations.

For all initial condition scenarios at 10 Myr, high-mass unbound stars have a higher median (and interquartile range) than the low/intermediate-mass stars and also a much higher minimum 2D-velocity. The mechanism for high-mass stars to become unbound is different to that of low/intermediate-mass stars. High-mass stars will only become unbound from my star-forming regions after a dynamical interaction with other massive stars in multiples. These dynamical interactions make unbound high-mass stars move faster on average; however, the fastest stars are in fact from the low-mass end. The differences in 2D-velocities between the mass classes is present in all initial condition combinations, so is not affected by the initial spatial or velocity structure in the star-forming regions.

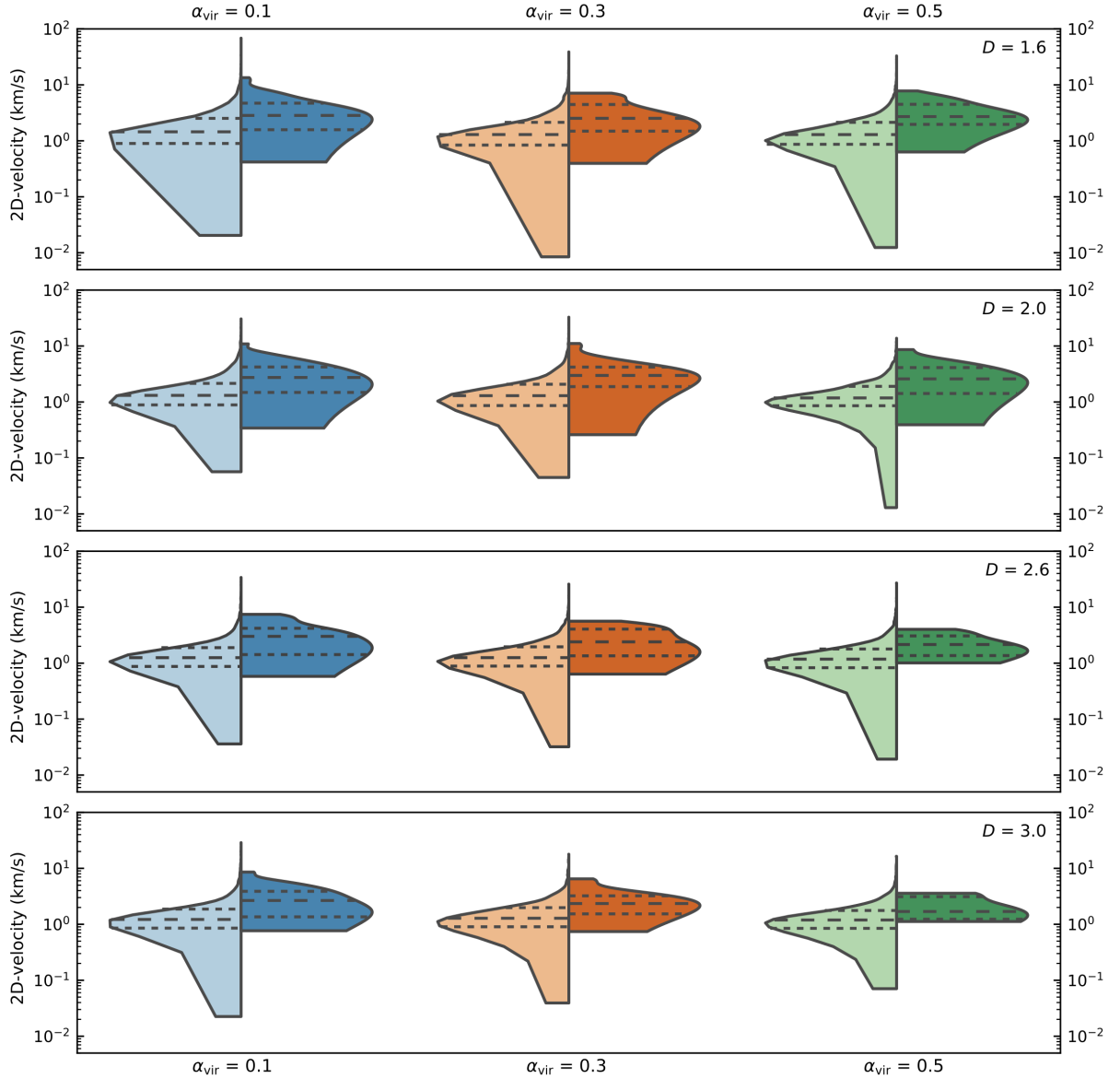


Figure 4.8: Violin plots showing the 2D-velocity distributions of unbound stars at 10 Myr split by mass class (low/intermediate-mass - left half, high-mass - right h) from all initially subvirial and virialized clusters ( $\alpha_{\text{vir}} = 0.1$  (blue),  $\alpha_{\text{vir}} = 0.3$  (orange)) and virialized ( $\alpha_{\text{vir}} = 0.5$  (green)). All plots are scaled to have the same width as there is only a very small number of unbound high-mass stars. The widest part of each violin half represents the 2D-velocity with the highest probability. Dashed lines represent the median and the interquartile range, the 95% confidence interval is no longer shown.

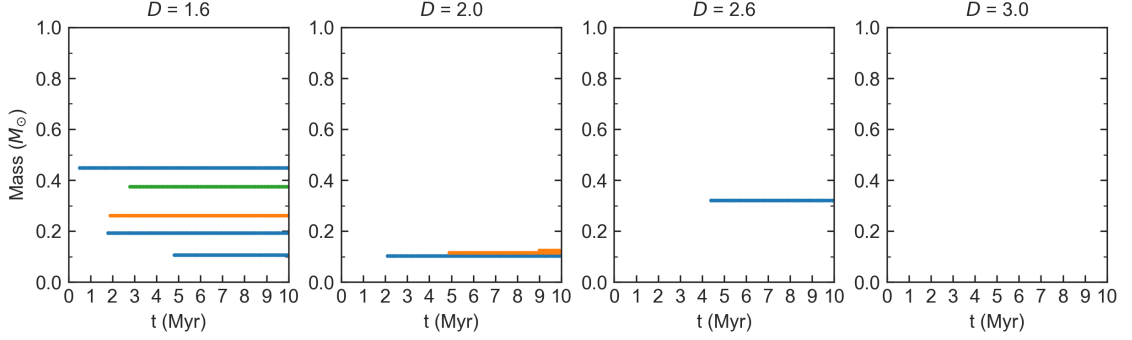


Figure 4.9: RW stars ( $2D\text{-velocity} > 30 \text{ km s}^{-1}$ ) by mass over the simulation time for the four fractal dimensions and the three different initial virial ratios: subvirial ( $\alpha_{\text{vir}} = 0.1$  (blue) and  $\alpha_{\text{vir}} = 0.3$  (orange)) and virialized ( $\alpha_{\text{vir}} = 0.5$  (green)). The y-axis is limited to  $1 M_{\odot}$ , as all of my RW stars have very low mass.

#### 4.3.4 Runaway and walkaway stars

Finally, I analyse how effective star-forming regions with different initial conditions are at ejecting RW and WW stars. I only use 2D-velocity and the lower boundary value of  $30 \text{ km s}^{-1}$  (e.g. Blaauw 1956; Stone 1991; de Wit et al. 2005; Eldridge et al. 2011) for my RW definition and velocities between  $5\text{-}30 \text{ km s}^{-1}$  for WWs (Eldridge et al. 2011; de Mink et al. 2014).

Fig. 4.9 shows all stars from 20 simulations per initial condition moving with a 2D-velocity (xy-plane) above  $30 \text{ km s}^{-1}$ . All of them are from the low end of the mass spectrum, not a single RW star is more massive than  $0.5 M_{\odot}$ . I have the highest number of RW stars from initially highly substructured, subvirial regions ( $\alpha_{\text{vir}} = 0.1$ ,  $D = 1.6$ ) regardless of the choice of 2D-plane. Only the fastest one is present in all three 2D-planes and is moving with a 2D-velocity between  $50\text{-}70 \text{ km s}^{-1}$  depending on the choice of plane. The other two RWs have lower velocities between  $30\text{-}40 \text{ km s}^{-1}$ . With at most three ejected RWs from a set of 20 simulations, I see that regardless of the initial velocity or spatial structure, RW stars are rare from my chosen initial conditions.

Going to WW velocities ( $5\text{-}30 \text{ km s}^{-1}$ ) produces a few high-mass WWs and a large number of low-mass WWs across all initial conditions. Fig. 4.10 shows all WWs from the 20 simulations across each initial condition set. The more violent the early evolution of a star-forming region is, the higher the number of WW stars. In the most violently evolving initial condition set-up - initially highly substructured ( $D = 1.6$ ) and highly subvirial ( $\alpha_{\text{vir}} = 0.1$ ), I have on average  $\sim 0.5$  high-mass WWs per simulation and  $\sim 20$  low/intermediate-mass WWs per simulation.

The lower the initial level of substructure (larger fractal dimension  $D$ ) the lower

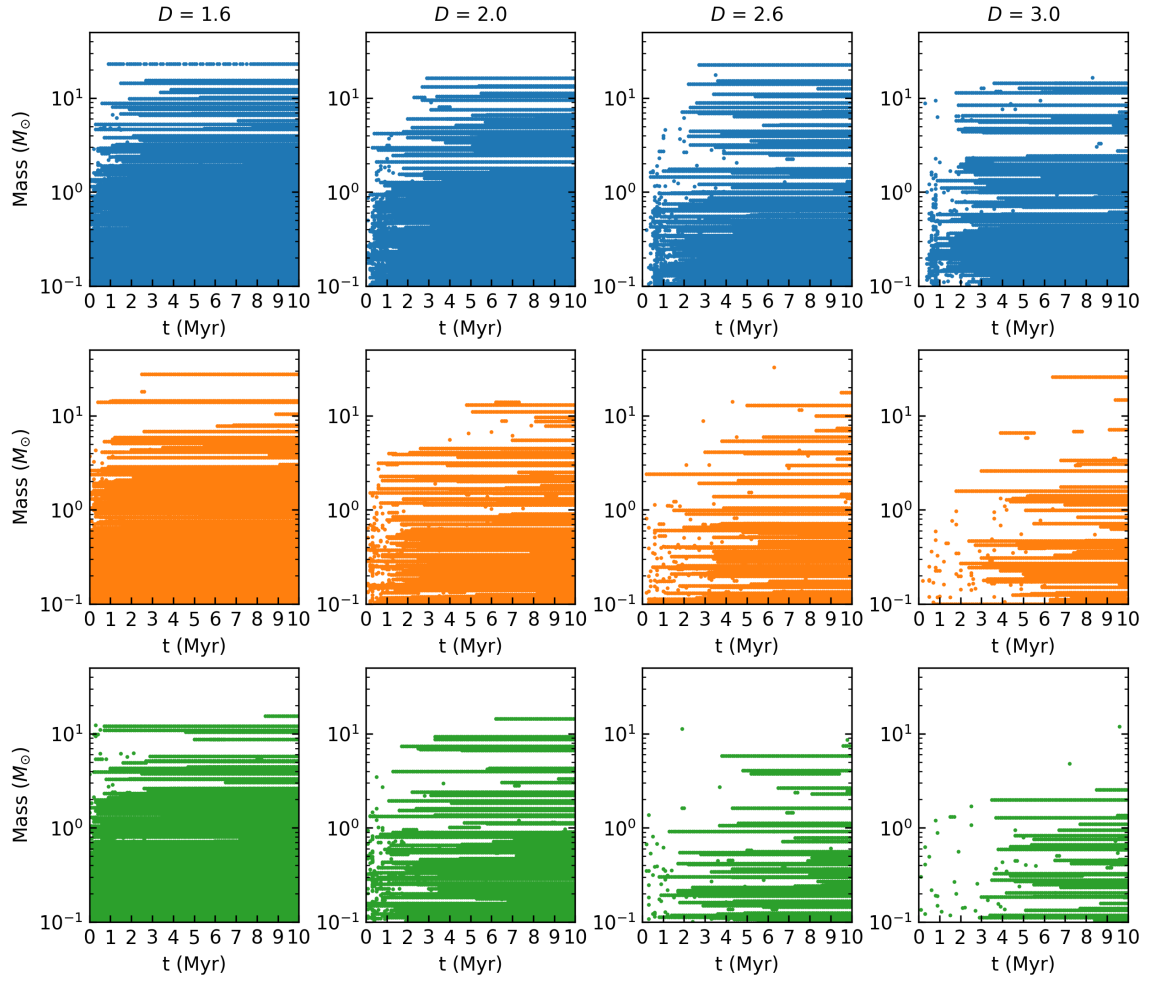


Figure 4.10: WW stars (2D-velocity:  $5\text{-}30\text{ km s}^{-1}$ ) by mass (using a log-scale) over the simulation time for the four fractal dimensions and the three different initial virial ratios: subvirial ( $\alpha_{\text{vir}} = 0.1$  (blue, top row) and  $\alpha_{\text{vir}} = 0.3$  (orange, middle row)) and virialized ( $\alpha_{\text{vir}} = 0.5$  (green, bottom row)). A few stars (single points) are only identified as WWs for a few snapshots. This is due to them being ejected close to the lower WW velocity boundary and slowing down to fall below the boundary shortly after ejection.

the overall number of WW stars, with initially more subvirial regions (Fig. 4.10, top row) producing more WW stars, which are also ejected earlier in the simulations. I see a number of temporary WWs that appear as WWs only for a few snapshots. These are stars ejected just at the minimum WW velocity. After ejection they slow down and disappear from my plots once they drop below  $5 \text{ km s}^{-1}$  (minimum WW velocity). However, this does not mean that they have been recaptured by the star-forming region. Initially virialized star-forming regions with no substructure ( $\alpha_{\text{vir}} = 0.5$ ,  $D = 3.0$  - bottom right panel) produce on average only 2 low/intermediate-mass WWs per simulation. This is an order of magnitude fewer WW stars than in the initial condition scenario (initially highly substructured ( $D = 1.6$ ), highly subvirial ( $\alpha_{\text{vir}} = 0.1$ ) - top left panel) that produces the largest number of WW stars.

## 4.4 Discussion

I summarise the results of my  $N$ -body simulations as follows. Cumulative velocity distributions of star-forming regions with different initial conditions have limited usefulness in clearly distinguishing between different initial spatial and velocity structure. When comparing the long-term evolution of regions with different levels of initial substructure, regions with high levels of initial substructure evolve very quickly kinematically, with supervirial regions (unbound by definition) showing the fastest 2D-velocities. The cumulative velocity distributions of unbound stars from initially subvirial and virialized simulations are difficult to distinguish after 10 Myr and only show differences for extremely different initial conditions (see Fig. 4.5).

The unbound fraction differs considerably for different combinations of initial spatial and velocity structure. This suggests that the unbound population around young, bound star clusters could possibly be used to draw conclusions about their initial conditions. Around initially smooth ( $D = 3.0$ ), virialized ( $\alpha_{\text{vir}} = 0.5$ ) star-forming regions, I find a low number of ejected stars (slow WWs, but no RWs) and virtually no unbound high-mass stars after 10 Myr. Around initially substructured, subvirial regions that have undergone violent relaxation, I find a large number of unbound low/intermediate-mass stars. I also find a few high-mass ejected stars (at WW velocities) and one low-mass RW star in three of the 20 simulations. The unbound fractions are possibly influenced by my choice of initial density as higher densities increase the likelihood of encountering and interacting with other stars.

Initial densities can differ greatly from those currently observed due to the amount



of dynamical evolution that a region undergoes. The level of spatial substructure in a region can constrain the dynamical evolution of regions with different initial densities - the higher the initial density, the quicker substructure is erased (e.g. Parker 2014). My simulated star-forming regions have been set up with a moderate to high, median local initial density ( $10^2$ - $10^4$   $M_{\odot}$   $\text{pc}^{-3}$ ), where initially more substructured regions show the highest densities.

After about 1 Myr, regions with initial spatial substructure have evolved into smooth, centrally concentrated regions, whose densities can be directly compared to observed star-forming regions. The density in my simulations after a few Myr is  $10^1$ - $10^3$   $M_{\odot}$   $\text{pc}^{-3}$  (e.g. Parker 2014) comparable to many nearby star-forming regions where observed present-day densities do not exceed  $\sim 400$   $M_{\odot}$   $\text{pc}^{-3}$  (e.g. Marks & Kroupa 2012).

High-mass stars are less likely to become unbound than low/intermediate-mass stars if a region is not initially very subvirial. When they do escape from their birth environment they do so at higher velocity and become at least WW stars ( $> 5$   $\text{km s}^{-1}$ ). With my chosen initial conditions, high-mass stars do not reach the velocity regime of RW stars. Only the evolution of star-forming regions that are initially subvirial ( $\alpha_{\text{vir}} < 0.5$ ) and/or substructured ( $D \leq 2.0$ ) is dynamic enough to produce any RW stars, all of which are low-mass. This is in apparent contrast to observations, where due to observational bias, predominantly high-mass RWs are found as they are more luminous and easier to observe. Historically, the definition of RW stars is based on OB stars (Blaauw 1961), following the suggestion in Tetzlaff et al. (2011), I also suggest extending this definition to lower-mass stars. Lower-mass stars appear to reach RW velocities more often than higher-mass stars and these could be found around many young star-forming regions when testing my predictions with *Gaia* DR2.

Using data from *Gaia* DR1 (Gaia Collaboration et al. 2016b), Wareing et al. (2018) reported that two of the most massive stars (HD 46223 and HD 46106) in NGC 2244 are moving away from each other and from the centre of this young cluster at a larger velocity than the other cluster stars. They suggested that HD46223 was ejected from the cluster, possibly due to dynamical interactions with other massive stars in the centre. The inferred velocity of  $1.38$   $\text{km s}^{-1}$  from its proper motion (Wareing et al. 2018) is far below the lower velocity boundary for walkway stars and it is unclear if this star is actually unbound. My simulated star-forming regions (1000 single stars) have an escape velocity of  $\sim 3$   $\text{km s}^{-1}$ . NGC 2244 is estimated to have  $\sim 2000$  members (Wang et al. 2008) suggesting that HD 46223 might not have reached escape velocity and might still be bound to the cluster despite its apparent ejection. In my simulations, I also see

massive stars moving outwards after dynamical interactions at velocities higher than their surroundings. If they are moving more slowly than the escape velocity, they will remain bound to the cluster, slow down and eventually return in direction of the cluster centre.

Violin plots show that the velocity distributions do indeed differ between initial conditions, particularly when the regions are initially highly substructured. These distributions also indicate that the vast majority of low/intermediate-mass stars become unbound at just around the escape velocity. I show that 2D-velocity information appears to be an underestimate of the full 3D-velocity for a proportion of unbound stars. This can have implications for membership determination of young star-forming regions, where full velocity parameter space information is not available. The *Gaia* DR2 data set contains a much larger number of stars only with proper motion data, missing information about the radial velocity for many fainter stars. If the 2D-velocity is indeed an underestimate of the full space velocity for some stars, cluster membership might be mistakenly assigned to stars with slow proper motions or ejected stars could not be traced back to their birth cluster.

Escaping, ejected or unbound stars from simulations have been studied previously (e.g. [Weidner et al. 2011](#); [Fujii & Portegies Zwart 2011](#); [Allison 2012](#); [Moyano Loyola & Hurley 2013](#); [Oh & Kroupa 2016](#); [Wang et al. 2019](#)). [Allison \(2012\)](#) found a similar connection between unbound stars (i.e. number, velocity, spatial distribution) and the initial substructure and virial ratio with a more limited set of initial conditions. Other studies ([Weidner et al. 2011](#); [Oh & Kroupa 2016](#)) used Plummer spheres ([Plummer 1911](#)) to set up the initial spatial distribution of the clusters and included primordial binaries. The conclusion from these studies was that the number and mass fraction of unbound stars depend strongly on the initial cluster radius or initial density and to a lesser extent on the parameters of the primordial binaries ([Weidner et al. 2011](#); [Oh & Kroupa 2016](#)) or the initial virial ratio ([Weidner et al. 2011](#)). With their inclusion of primordial binaries, the results of these studies are not directly comparable to my results.

My results show that differences in the initial spatial substructure can have a considerable effect on the fraction, the velocity and the masses of unbound stars. Due to the lack of stellar evolution in my short simulation time of 10 Myr, I miss the effects of SN kicks causing stars to become unbound due to the BSS ([Blaauw 1961](#)). In my simulations, binaries will only form dynamically (i.e. are not present from the beginning of my simulations) and I may therefore be underestimating the impact of the DES ([Poveda et al. 1967](#)) as I only find a few lower-mass RWs stars. [Moyano Loyola & Hur-](#)

ley (2013) showed that a higher fraction of primordial binaries increases the number of higher-velocity ( $20\text{-}100\text{ km s}^{-1}$ ) stars.

## 4.5 Conclusions

In this chapter, I present  $N$ -body simulations of star-forming regions set up with a range of different initial spatial and velocity structures. I investigate if the dynamical evolution results in differences in the unbound population after 10 Myr. The conclusions from my simulations are summarised as follows:

- (i) Cumulative 2D-velocity distributions of all stars in simulated star-forming regions cannot provide strong insights into the long-term evolution of star-forming regions with differing initial spatial and velocity structure. When focussing on unbound stars, clear differences in the cumulative distributions are only found when comparing vastly different initial conditions.
- (ii) Unbound fractions of stars of different masses show clear differences between the initial conditions and could prove useful to distinguish between initial spatial and velocity structures. Only when a region is initially very subvirial can we expect to find a higher fraction of unbound high-mass stars than low/intermediate-mass stars in the vicinity of the region.
- (iii) If high-mass stars manage to escape their birth region, they are likely to reach at least WW velocities. However, based on my simulations, not every young star-forming region will create a high-mass RW or WW star.
- (iv) Most low/intermediate-mass stars leave the regions at velocities just above the escape velocity. However, the fastest stars from my simulations are also low/intermediate-mass stars. I see a number of low/intermediate-mass WW stars from every initial condition set. This number increases for regions that evolve more dynamically (more initial substructure and lower virial ratio). As a result, we should find at least a small number of these stars around virtually every young and high-density star-forming region. The fact that most observed fast stars are still high-mass is very likely due to observational bias/limitations. This changes with *Gaia* DR2 where five-parameter space astrometry for stars down to sub-solar mass is already available for several nearby star-forming regions (Gaia Collaboration et al. 2018a).

## Chapter 5

### Ejected stars from the ONC in *Gaia* DR2 observations

## 5.1 Introduction

This chapter is based on the analysis published in [Schoettler et al. \(2020\)](#). In the work presented here, I adjust the approach from the published version to convert the ICRS coordinates to Cartesian coordinates as done in Chapter 7, which is based on [Schoettler et al. \(2022\)](#). In the following chapter, I use  $N$ -body simulations to predict the number of RW and WW stars for an ONC-like star-forming region. I then use *Gaia* DR2 observations to search for RW and WW stars around the ONC. I first describe the search target. The *Gaia* DR2 data selection and analysis process follows the method described in Chapter 3 and the  $N$ -body simulation set-up follows the approach described in Chapter 2. I then present the predictions from my simulations, followed by the observational results and a brief discussion. Concluding remarks about the implications of the analysis complete this chapter.

## 5.2 ONC specific information used in the analysis

My first target for the observational search for RW and WW stars is the ONC. This star-forming region is well-suited for this purpose due to its proximity to Earth ( $\sim 400$  pc, e.g. [Großschedl et al. 2018](#); [Kuhn et al. 2019](#)). At such a close distance, the faintest stars in my data set will be stars down to sub-solar masses with reasonably accurate *Gaia* DR2 proper motions ([Kuhn et al. 2019](#)). The ONC is a very young region, and its ejected stars will be easier to trace back as they are more likely to still be in close proximity to the region. It has a mean age estimate of 2-3 Myr and a spread of  $\sim 2$ -2.5 Myr (e.g. [Da Rio et al. 2010](#); [Reggiani et al. 2011](#)). For my analysis, I adopt an upper age limit of 4 Myr, thus encompassing the full range of estimated ages of the ONC.

The ONC has an established list of cluster members across different wavelength bands, which allows me to define a clear cluster boundary to trace back ejected stars. [Hillenbrand \(1997\)](#) and [Hillenbrand & Hartmann \(1998\)](#) suggested that the number of stars visible in the optical spectrum is  $\sim 1600$  and a further  $\sim 1900$  stars are visible in the infrared (IR). An updated census by [Da Rio et al. \(2012\)](#) increased the number of known members in the optical spectrum to  $\sim 1750$  stars. This large number of cluster stars and higher local stellar density can increase the likelihood of dynamical ejections (e.g. [Oh & Kroupa 2016](#); [Farias et al. 2019](#); [Schoettler et al. 2019](#)). [Hillenbrand \(1997\)](#) considered the projected size of the ONC in two dimensions to be  $2.5 \times 4.5$  pc and [Kroupa et al. \(2018\)](#) suggested the cluster to have a nominal radius of  $\sim 2.5$  pc. The ONC members in the [Da Rio et al. \(2012\)](#) census also do not extend any further than this radius on the

sky. For my analysis, I therefore use this nominal radius as my cluster boundary on the sky. I consider the ONC to be the region associated with the nebula centred around the Trapezium cluster.

To become a RW/WW, a star should be unbound from its birth region meaning it has to at least reach the respective escape velocity. Kim et al. (2019) suggested an upper limit on the angular escape speed from the ONC of  $\sim 3.1 \text{ mas yr}^{-1}$ . At an adopted distance of 400 pc (e.g. Großschedl et al. 2018; Kuhn et al. 2019) to the region, this angular escape speed translates to a space velocity of  $\sim 5.8 \text{ km s}^{-1}$ . This implies that the suggested lower velocity limit for WWs of  $5 \text{ km s}^{-1}$  (Eldridge et al. 2011) might not be appropriate for star-forming regions with a total mass similar to or higher than the ONC. This lower velocity limit value can then result in considering stars to have “walked away” while still being gravitationally bound to their birth region. For my analysis, I use a velocity boundary for WW candidates of  $10 \text{ km s}^{-1}$  and consider stars above  $30 \text{ km s}^{-1}$  to be RW candidates.

The ONC is thought to have produced high-mass RW stars in the past, most notably AE Aur and  $\mu$  Col, which were among the first identified RWs. Blaauw & Morgan (1954) showed that these two stars are moving in almost opposite direction from the ONC at space velocities of  $\sim 100 \text{ km s}^{-1}$  and suggested that they were ejected in the same event  $\sim 2.6 \text{ Myr}$  ago. Hoogerwerf et al. (2001) used Monte-Carlo simulations to show that observations of these two stars are consistent with having originated in the Trapezium Cluster (at the centre of the ONC) and their RW status being a consequence of a binary-binary dynamical ejection  $\sim 2.5 \text{ Myr}$  ago.

The Becklin-Neugebauer (BN) object (Becklin & Neugebauer 1967) is another fast moving, high-mass star that has been postulated to have been recently ejected from the Orion region; however, its exact origin is still debated (e.g. Tan 2004; Bally & Zinnecker 2005; Rodríguez et al. 2005; Farias & Tan 2018). Unlike AE Aur and  $\mu$  Col, which are both visible in the optical spectrum the BN-object is an IR source and not visible in the wavelength range covered by *Gaia*.

The ONC has also been suggested as the origin of three potential low-mass RWs (Poveda et al. 2005). These candidates were identified based on their high proper motion and converted to tangential velocities ( $38\text{-}69 \text{ km s}^{-1}$ ) based on an assumed distance of 470 pc. However, O’Dell et al. (2005) used Hubble Space Telescope observations to show that these three stars do not actually move fast enough to be classified as RW stars. With these new velocities ( $5.5\text{-}7.9 \text{ km s}^{-1}$ ) I would not even consider them as WW stars. Given the suggested upper limit escape speed (Kim et al. 2019) of  $\sim 5.8 \text{ km s}^{-1}$ , these

Table 5.1: ONC centre parameters used in the analysis; [1] [Kuhn et al. \(2019\)](#)

Right ascension [RA] (ICRS) $\alpha_0$	5h 35m 16s [1]
Declination [Dec] (ICRS) $\delta_0$	-05° 23' 40'' [1]
Proper motion RA $\mu_{\alpha^*,0}$ (mas yr <sup>-1</sup> )	1.51 ± 0.11 [1]
Proper motion Dec $\mu_{\delta_0}$ (mas yr <sup>-1</sup> )	0.50 ± 0.12 [1]
RV (km s <sup>-1</sup> )	21.8 ± 6.6 [1]
Adopted distance (pc)	400
Adopted parallax $\varpi_0$ (mas)	2.50

stars might not be unbound from the ONC if this is in fact their birth region.

There are two more recent searches for ejected stars from the ONC. [McBride & Kounkel \(2019\)](#) compiled a list of known young stellar objects from literature within a search radius of 2° around the ONC and a parallax limit of  $2 < \varpi < 5$  mas. They cross-matched these stars with *Gaia* DR2 data and applied photometric and minimum proper motion cuts. These steps result in 26 potential RW candidates having been ejected from the ONC. After tracing these candidates back, they identify nine stars with an apparent origin close to the Trapezium Cluster at the centre of the ONC. My analysis covers this region as well and I will seek to confirm these candidates in my analysis.

[Farias et al. \(2020\)](#) searched a region out to 45° around the ONC and to a distance of 1 kpc for sources that are coincident with the location of the ONC in the past 10 Myr. This analysis produced a large number of potential sources that the authors reduced down using different quality filters, such as YSO colour, IR-excess and approaches within the half-mass radius of the ONC. They identified 25 new candidates not previously known in the literature (however, several of these were already identified in my work published shortly before in [Schoettler et al. \(2020\)](#)), ten of them passing their strictest requirements.

As described in Chapter 3, I adopt the centre for my star-forming regions as described in [Kuhn et al. \(2019\)](#). The values for the ONC are shown in Table 5.1. For my analysis, I use sources with distances between 300 and 500 pc, adopting a central distance of 400 pc, instead of 403 pc from [Kuhn et al. \(2019\)](#).

The ONC is fairly well constrained on the sky (position and size) allowing me to define an approximate cluster boundary by using the nominal radius of 2.5 pc ([Kroupa et al. 2018](#)). This corresponds to an angular size of  $\sim 0.35^\circ$  around the ONC centre position, which is located at the origin in the ONC rest-frame. However, the ONC is far less constrained in its distance. Older estimates positioned it further away, e.g. [Menten et al. \(2007\)](#) determined it to be at a distance of  $414 \pm 7$  pc. Recent estimates have since reduced the distance to the ONC, [Kounkel et al. \(2018\)](#) derived a distance of  $389 \pm 3$  pc,

whereas [Kuhn et al. \(2019\)](#) located it at  $403^{+7}_{-6}$  pc, both using *Gaia* DR2.

The size of the ONC in the line-of-sight direction is also less constrained than it is on the plane of the sky. [Großschedl et al. \(2018\)](#) used *Gaia* DR2 to investigate the 3D-shape of the Orion A region, which includes the ONC at its “Head”. The authors suggested that the ONC region extends from its centre for about 15-20 pc in either direction. For my analysis, I consider my cluster boundary in the line-of-sight direction to extend 15 pc either direction from the adopted centre at 400 pc, which is located at the origin of the rest frame.

Most of the stars ( $\sim 93\%$ ) in my data set do not have a measured RV in *Gaia* DR2. As a consequence, I start with a search for 2D-candidates and only when confirmed as a 2D-candidate do I proceed in three dimensions for stars with RV. For 2D-candidates without RV-measurements in *Gaia* DR2, I have searched through several RV surveys, i.e. RAVE DR5 ([Kunder et al. 2017](#)), GALAH DR1 ([Martell et al. 2017](#)) and also the Simbad/VizieR databases ([Wenger et al. 2000](#); [Ochsenbein et al. 2000](#)) to complete my data set with secondary RV measurements. I find additional RVs in [Gontcharov \(2006\)](#), [Cottaar et al. \(2015\)](#), and [Kounkel et al. \(2018\)](#) for several 2D RW and WW candidates. I also find secondary, more precise RVs for several sources, where the *Gaia* DR2 RVs have large errors, and use these secondary RVs instead.

For 2D-candidates without RV, I use the radial distance of these candidates to the cluster boundary and each candidate’s minimum flight time to calculate a required RV to reach this distance since ejection. If the resulting velocity is  $> |500| \text{ km s}^{-1}$  I exclude these 2D-candidates from the list. Before tracing back 2D-candidates with RVs in three dimensions, I also exclude those where the RV points towards the ONC as these stars cannot have originated from the ONC.

I use an upper age limit of 4 Myr and consider only stars that are younger than this age to have possibly originated in the ONC. To get age estimates of my candidate stars, I use PARSEC isochrones (version 1.2S, [Bressan et al. 2012](#)) to separate the stars into two age brackets (younger stars plotted above the isochrone, older stars below). I download data<sup>1</sup> to produce an isochrone using a linear age of 4 Myr with a mean metallicity of  $Z = 0.011268$  (e.g. [Santos et al. 2008](#); [Biazzo et al. 2011](#)).

---

<sup>1</sup><http://stev.oapd.inaf.it/cgi-bin/cmd>



### 5.3 $N$ -body simulations of the ONC

To predict the number and velocities of RW and WW stars in the observational data, I have run a set of 20  $N$ -body simulations with initial conditions similar to those the ONC is thought to have evolved from. In Chapter 4, I used  $N$ -body simulations to show that the number and velocity distributions of ejected stars can be used to constrain the initial spatial and kinematic substructure and I use a similar approach here to provide predictions for my search. The general set-up of the simulations is described in Chapter 2 and in the following section I comment on parameters specific to the ONC-like simulations.

#### 5.3.1 Simulation set-up

The ONC is thought to have evolved from an initial state of being spatially and kinematically substructured (e.g. Allison et al. 2010; Allison & Goodwin 2011). Spatial substructure can be created in  $N$ -body simulations by using fractal distributions, as shown in Goodwin & Whitworth (2004). The degree of substructure is defined using only a single parameter, the fractal dimension  $D$ . In my  $N$ -body simulations, I use a fractal dimension  $D = 2.0$ , which produces a moderate amount of spatial substructure. The use of fractals also allows me to set up the initial kinematic substructure. The velocities in my simulations are scaled so the regions are initially subvirial with a virial ratio  $\alpha_{\text{vir}} = 0.3$ . A detailed description of the construction of the fractals for the simulations is shown in Goodwin & Whitworth (2004); Parker et al. (2014); Parker & Wright (2016) and described in Chapter 2.

I use a larger number of systems than in Chapter 4, i.e. 2000 systems compared to 1000 systems per simulation to reflect the higher number of stars in the ONC. The masses for the systems are sampled randomly from a Maschberger (2013) IMF with stellar masses between  $0.1 M_{\odot}$  and  $50 M_{\odot}$ . This upper mass limit of my sample is consistent with the mass estimate of the most massive star in the ONC, which is  $\theta^1$  Ori C, a visual binary system with a total mass of  $\sim 50 M_{\odot}$  and a  $\sim 30\text{-}35 M_{\odot}$  primary (e.g. Hillenbrand & Hartmann 1998; Kraus et al. 2007; Muench et al. 2008; Lehmann et al. 2010).

I also include primordial binaries in my simulations. The set-up of the binaries follows the procedure described in Chapter 2.3, the resulting average total number of stars in my simulations is  $\sim 2800$  stars, I have an average cluster mass of  $\sim 2100 M_{\odot}$  and an escape velocity of  $\sim 6 \text{ km s}^{-1}$ , which is consistent with the estimated total cluster mass in Hillenbrand & Hartmann (1998) and the escape velocity estimate in Kim et al. (2019).

I evolve my star-forming regions over a defined time period of 4 Myr and take snap-

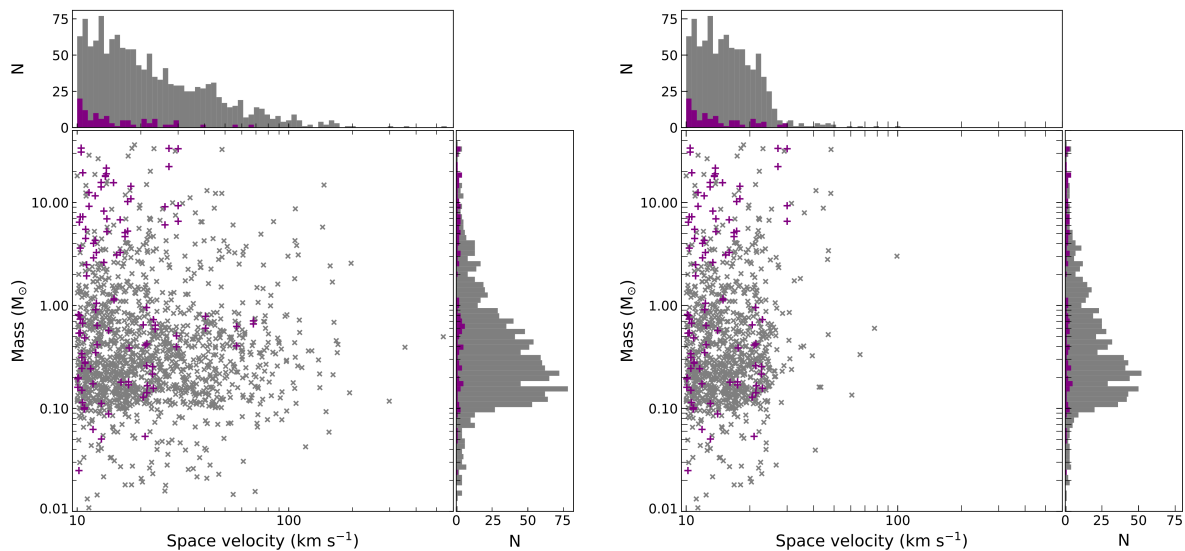


Figure 5.1: Masses and velocities of all RW/WW stars from the 20 simulations after 4 Myr. Stars with masses  $m < 0.1 M_{\odot}$  are current or previous BD binary companion stars. Ejected binaries are marked with a purple “+”, whereas single stars are marked with a grey “x”. Left panel: RW/WW stars at all distances. Right panel: RW/WW stars still located within the 100 pc search boundary.

shots every 0.01 Myr. The initial radius of my star-forming regions is 1.5 pc and I have not applied any external tidal field. The stellar systems in my simulations undergo stellar and binary evolution. I do not see any SNe during the 4 Myr simulations. The furthest evolution step for the highest-mass star is that of a helium star. I do however see several binary mergers as a result of binary evolution.

### 5.3.2 Predictions from the simulations

The left panel of Fig. 5.1 shows the masses and space velocities of all ejected stars after 4 Myr from 20  $N$ -body simulations that reach at least WW velocities ( $> 10 \text{ km s}^{-1}$ ) at time of ejection. The distribution highlights that the highest velocities are achieved by lower-mass stars and that I should find RW and WW stars across the mass spectrum around the ONC.

Most of the RW and WW stars are ejected with a space velocity of  $< 200 \text{ km s}^{-1}$ . However, I have 3 sub-solar mass RW stars travelling with velocities between  $\sim 300$ - $540 \text{ km s}^{-1}$ . These RW-velocities are far above my average; however, they are not improbable for the DES (e.g. Leonard & Duncan 1990; Gvaramadze et al. 2009; Perets & Šubr 2012).

The fastest RW ( $\sim 540 \text{ km s}^{-1}$ ) from my simulations is the result of multiple dynamical interactions starting with a binary-binary interaction between two of the primordial binaries. The fastest RW is the primary (P1,  $0.5 M_{\odot}$ ) in an almost equal-mass binary ( $q = 0.85$ ). After 0.01 Myr, this binary interacts with another binary with a low mass ratio  $q = 0.03$  and a primary (P2) of  $24 M_{\odot}$ . This interaction leads to the ejection of the secondary (S1,  $0.4 M_{\odot}$ ) from the equal-mass binary, turning S1 into a WW star. The now single P1 replaces the secondary (S2,  $0.8 M_{\odot}$ ) in the unequal mass binary. The system continues as a triple system with S2 turning into a tertiary companion on a wider orbit around the close binary P2-P1.

This triple system moves towards the region's centre as the region collapses due to its initial subvirial ratio. The tertiary S2 gets ejected at 0.5 Myr as a WW after further dynamical interactions. The remaining binary forms short-lived dynamical multiples with different stars until it gets fully disrupted at 3 Myr by an encounter with a  $10 M_{\odot}$  star. My fastest star P1 gets ejected with a velocity close to its previous orbital velocity, whereas the high-mass primary P2 forms a new dynamical binary with the disrupting star and becomes an unbound binary just above the escape velocity.

I also find ejected RW/WW binaries across the sampled mass range and these stars are highlighted with a purple “+” in Fig. 5.1. The fraction of RW/WW binaries is low compared to the binary fractions in the cluster. This is consistent with [Leonard & Duncan \(1990\)](#) who showed that  $\simeq 10\%$  of their ejected stars with velocities  $> 30 \text{ km s}^{-1}$  are binaries, compared to an initial binary fraction of 50%. Also, [Perets & Šubr \(2012\)](#) suggested that in general the binary frequency of RW stars is lower than that of the stars still within the cluster.

The maximum velocity of the ejected binaries in my simulations ( $\sim 70 \text{ km s}^{-1}$ ) is lower than for single stars, which is also consistent with the results of [Leonard & Duncan \(1990\)](#) and [Perets & Šubr \(2012\)](#). [Portegies Zwart et al. \(2010\)](#) suggested that an ejected binary, resulting from a binary-single star encounter, received less kinetic energy and travelled at lower velocities than ejected single stars.

The stars in my simulations are sampled from the IMF down to  $0.1 M_{\odot}$ . However, I allow BD binary companions below this mass, so I also find a small number of BD RW and WW stars. All of my RW-BDs are single stars; however, I have an occasional WW-BD ejected in a binary.

The right panel in Fig. 5.1 shows that the velocity distribution of RW/WW stars still within 100 pc of the cluster centre at the end of my simulations is different. Most of the WW stars are still within this radius; however, most RW stars have passed through this

Table 5.2: Ejected RW and WW stars from  $N$ -body simulations at all distances and within the search radius of 100 pc at different times during the simulations. I show the averages from all 20 simulations and maxima in a single simulation. The values in [] indicate the number of RW/WW stars within 100 pc. I count ejected binary systems as one star when calculating averages and maxima. The uncertainties in my averages are the standard deviations.

Mass $m$ ( $M_{\odot}$ )	RW average / maximum	WW average / maximum
$0.01 \leq m < 0.10$		
- after 1 Myr	$1.6 \pm 1.1$ [ $1.5 \pm 1.1$ ] / 4 [4]	$3.5 \pm 2.0$ [ $3.5 \pm 2.0$ ] / 7 [7]
- after 2 Myr	$1.7 \pm 1.1$ [ $1.0 \pm 1.1$ ] / 4 [3]	$3.7 \pm 2.1$ [ $3.7 \pm 2.1$ ] / 7 [7]
- after 3 Myr	$1.8 \pm 1.1$ [ $0.3 \pm 0.4$ ] / 4 [1]	$3.8 \pm 1.9$ [ $3.8 \pm 1.9$ ] / 6 [6]
- after 4 Myr	$1.8 \pm 1.1$ [ $0.1 \pm 0.2$ ] / 4 [1]	$3.8 \pm 1.9$ [ $3.4 \pm 1.6$ ] / 7 [6]
$0.10 \leq m < 8.00$		
- after 1 Myr	$15.5 \pm 4.3$ [ $14.2 \pm 4.2$ ] / 24 [22]	$38.3 \pm 5.1$ [ $38.3 \pm 5.1$ ] / 46 [46]
- after 2 Myr	$16.7 \pm 4.1$ [ $10.8 \pm 3.0$ ] / 25 [15]	$41.7 \pm 5.0$ [ $41.7 \pm 5.0$ ] / 55 [55]
- after 3 Myr	$17.2 \pm 4.5$ [ $4.1 \pm 2.5$ ] / 27 [10]	$44.9 \pm 5.6$ [ $44.9 \pm 5.6$ ] / 57 [57]
- after 4 Myr	$17.6 \pm 4.4$ [ $1.0 \pm 1.0$ ] / 26 [3]	$45.6 \pm 5.0$ [ $41.6 \pm 5.5$ ] / 57 [55]
$m \geq 8.00$		
- after 1 Myr	$0.2 \pm 0.4$ [ $0.1 \pm 0.3$ ] / 1 [1*]	$0.6 \pm 0.9$ [ $0.6 \pm 0.9$ ] / 3 [3]
- after 2 Myr	$0.2 \pm 0.4$ [ $0.1 \pm 0.3$ ] / 1 [1*]	$0.8 \pm 1.1$ [ $0.8 \pm 1.1$ ] / 4 [4]
- after 3 Myr	$0.3 \pm 0.5$ [ $0.2 \pm 0.4$ ] / 1 [1*]	$1.5 \pm 1.7$ [ $1.5 \pm 1.7$ ] / 6 [6]
- after 4 Myr	$0.5 \pm 0.6$ [ $0.3 \pm 0.5$ ] / 2 [1*]	$1.7 \pm 1.5$ [ $1.6 \pm 1.4$ ] / 5 [5]

\*Depending on the simulation, I find 1 RW either within or outside of the 100 pc boundary

region already. The maximum RW velocity drops to  $\sim 100 \text{ km s}^{-1}$  when only considering stars within the 100 pc region.

Table 5.2 provides the average numbers of RW and WW stars across the mass ranges at four different times during the simulations and also gives the maximum I find in a single simulation. Most of my ejected stars are low/intermediate-mass stars ( $0.1 M_{\odot} < m < 8 M_{\odot}$ ) with an average of  $17.6 \pm 4.4$  RW and  $45.6 \pm 5.0$  WW stars ejected at all distances from 20 simulation after 4 Myr. The maximum number of RWs in this mass range from a single simulation is 26, whereas I find a maximum of 57 WW stars.

When I limit myself to ejected low/intermediate-mass stars still within the 100 pc search region, I only find an average of  $1.0 \pm 1.0$  RW star and a maximum of 3 RWs after 4 Myr. At earlier times (younger ages) in my simulations, I see the number of RWs still within the search region increasing to an average of  $14.2 \pm 4.2$  and a maximum of 22 at 1 Myr.

At lower WW velocities, I find that all ejected stars remain within the search region up to 3 Myr. Even after 4 Myr, I still find most WW stars, i.e.  $41.6 \pm 5.5$  stars (average) and 55 stars (maximum), within 100 pc. These findings suggest that I should find very few ejected RW stars, but most of the WW stars within my chosen 100 pc search radius around the ONC.

The number of RW/WW stars from the other two mass ranges (BDs and massive stars) are much lower. I only find a maximum of 2 high-mass RW stars ( $0.5 \pm 0.6$  stars average) and 5 high-mass WW stars ( $1.7 \pm 1.5$  stars average) at all distances. All of the WW stars at these masses are still located within the 100 pc search region, whereas only one of the two RW stars is. I also provide the number of ejected BDs in Table 5.2; however, I do not include them in the following analysis as I am unlikely to be able to observe ejected BDs at the distance of the ONC.

Most of the RW and WW stars are ejected as single stars. In different simulations, I find a maximum of one RW-binary and two WW-binaries composed of higher-mass stars. I also have a maximum of three RW-binaries and three WW-binaries composed of low/intermediate-mass stars ( $0.1 M_{\odot} < m < 8 M_{\odot}$ ).

## 5.4 Results from *Gaia* DR2

### 5.4.1 2D-candidates

Fig. 5.2 and 5.3 show the resulting CAMDs following the procedure described in Chapter 3.3.3 for my 2D-candidate RW and WW stars. A large number of the stars that

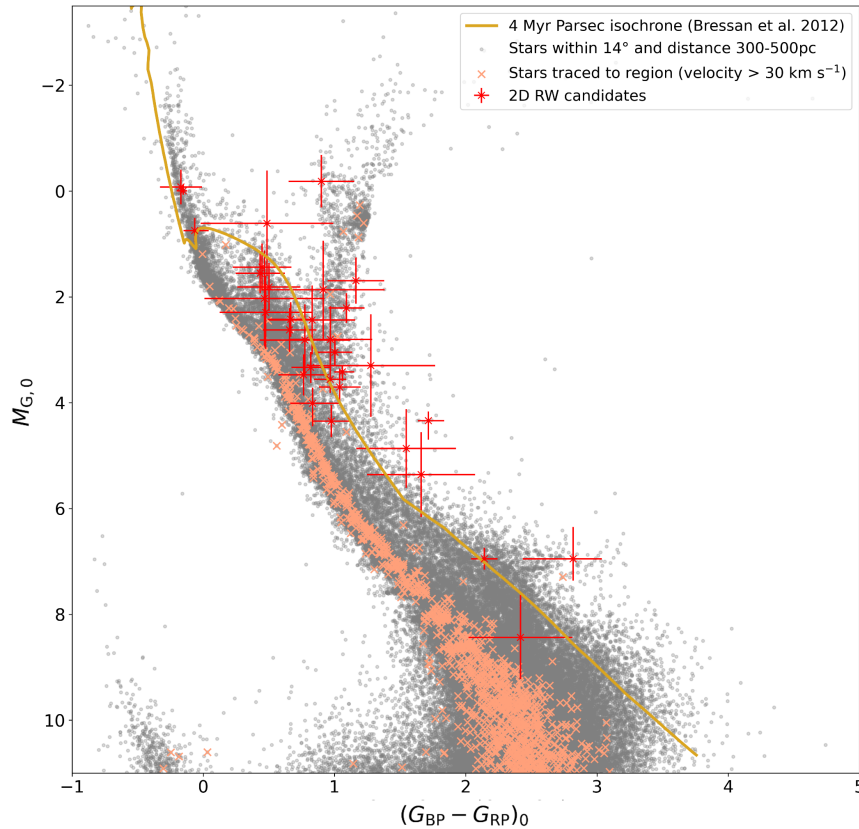


Figure 5.2: CAMD showing all 2D RW candidates ( $> 30 \text{ km s}^{-1}$ ) that can be traced back to the ONC (red “x”). I magnitude-limit the diagram to  $-3.5 \text{ mag} < M_{G,0} < 11 \text{ mag}$ , which corresponds to a  $G$ -magnitude  $\approx 19 \text{ mag}$  at the fainter end. Around this apparent magnitude value, the typical uncertainties in the 5-parameter astrometry increase quickly. The CAMD includes a large number of stars that I have traced back to the ONC but that sit along the MS underneath the 4 Myr isochrone (orange “x”) even when considering their errors, which are not plotted here. These stars are therefore too old to have been born in the ONC. I see a number of 2D-candidates at different absolute magnitudes that correspond to different masses. Many of my identified candidates have large errors in magnitude and colour, which are predominantly driven by the errors in the extinction and reddening. Some of these candidates sit below the 4 Myr isochrone but might be younger than their position suggests due to the large errors.

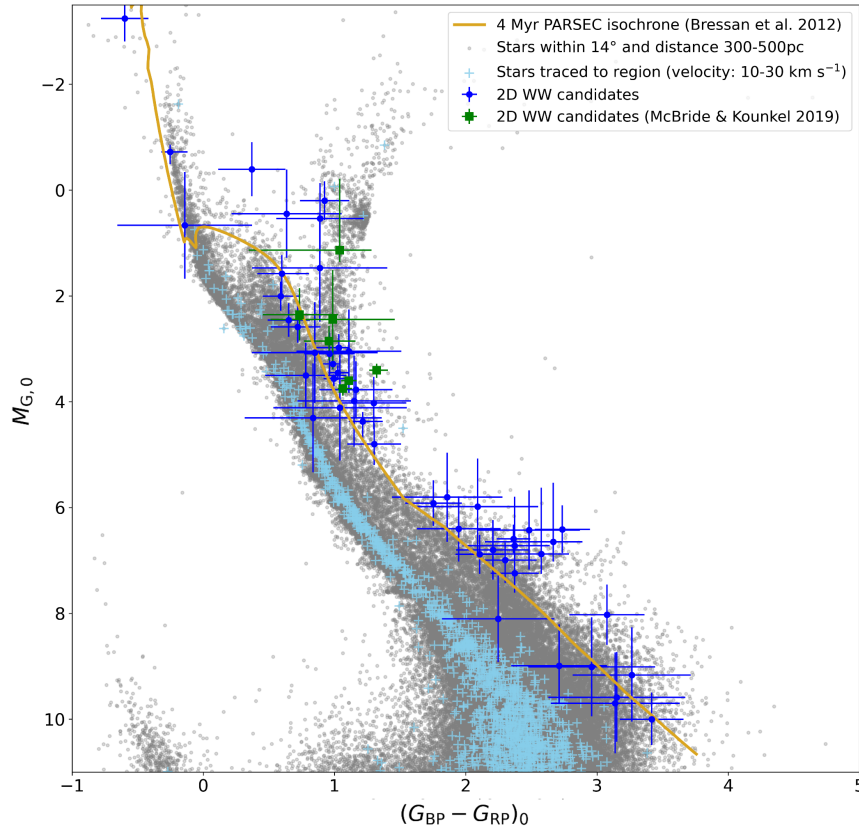


Figure 5.3: CAMD showing all 2D WW candidates (velocity: 10-30 km s<sup>-1</sup>) that can be traced back to the ONC (blue dot). I magnitude-limit the diagram to  $-3.5 \text{ mag} < M_{G,0} < 11 \text{ mag}$ , which corresponds to a  $G$ -magnitude  $\approx 19 \text{ mag}$  at the fainter end. Around this apparent magnitude value, the typical uncertainties in the 5-parameter astrometry increase quickly. The CAMD includes a large number of stars that I have traced back to the ONC but that sit along the MS underneath the 4 Myr isochrone (light-blue “+”). These stars are therefore too old to have been born in the ONC, even when considering their errors, which are not plotted here. I see a number of 2D-candidates above the isochrone at different absolute magnitudes that correspond to different masses. The WW-candidates extend to much lower magnitudes than the RW-candidates shown in Fig. 5.2. On the CAMD, I identify the candidates I have in common with McBride & Kounkel (2019) (green square), which sit towards brighter magnitudes but still within the pre-main sequence part of the isochrone. Many of my identified candidates have large errors in magnitude and colour, which are predominantly driven by the errors in the extinction and reddening. Quite a few of these candidates sit below the 4 Myr isochrone but might be younger than their position suggests due to the large errors.

have been traced back to the ONC search region in the xy-plane (on the sky) are located along the MS underneath the 4 Myr isochrone and are too old to have originated from the ONC. Located above the isochrone are all traced-back candidates that are young enough to have been born in the ONC. In addition, I also find potential candidates that are located below the isochrone but where the photometric error bars cross the isochrone, indicating a possibly younger age.

I find 29 RW and 57 WW 2D candidates with an isochronal age  $< 4$  Myr after applying my criteria for candidate identification. I exclude one RW and one WW candidate without RV as based on their radial distance to the ONC the RV required to get to their current position since ejection from the ONC is unreasonably large for RW or WW stars ( $> |500| \text{ km s}^{-1}$ ). A few 2D-candidates (seven RWs and eight WWs) with known RVs are excluded as their RVs point towards the ONC instead of away from it, so these stars cannot have originated from the ONC. I also identify four RWs and one WWs located clearly in a red giant position. These stars are also excluded from the result table and the CAMD plots. Further information on the excluded stars can be found in Appendix A Tables A.1, A.2 and A.4

Table 5.3 provides an overview of all my identified RW candidates in 2D with information about their velocities in the ONC rest frame, flight times and approximate isochronal ages. I also identify whether any of the stars with RV have been traced back successfully in 3D. Table 5.4 provides the same information for the WW candidates. Both tables also include information gathered from literature sources about my candidates, such as age, mass and spectral type.

The brightest 2D RW candidate HD 288089 (*Gaia* DR2 3222673430030590592) has an absolute magnitude of  $\sim -0.2$  mag. Very little is known about this star, apart from its spectral type listed as K0 (Nesterov et al. 1995). It is the only 2D RW candidate whose flight time is larger than its isochronal age and as a consequence it is unlikely to have originated from the ONC. The second brightest star in this list HD 41288 (*Gaia* DR2 3122639449820663040) has an absolute magnitude of  $\sim 0$  mag and its spectral type from literature is a B5 (Houk & Swift 1999). This candidate is the most massive 2D RW candidate in my data set and its position is consistent with an isochrone age of  $\sim 2$  Myr. Its 2D-velocity is  $\sim 44 \text{ km s}^{-1}$ .

My brightest 2D WW candidate  $\nu$  Ori (*Gaia* DR2 3016424530632449280) is a known O-star (Sota et al. 2011) with an absolute magnitude of  $\sim -3.3$  mag on my CAMD. This star is one of only a few stars in my 2D-candidate list that has reached the MS. It is located slightly underneath the MS due to an over-correction for extinction as a conse-



Table 5.3: RW star 2D candidates sorted by decreasing 2D-velocity. Column 2+3: velocity in ONC rest frame [rf]; Column 3: RV sources - <sup>a</sup>Gaia DR2, <sup>b</sup>Cottaar et al. (2015); Column 4: indication of 3D-candidate status; Column 5: minimum flight time since ejection (crossing of search boundary); Column 6: age from PARSEC isochrones (Bressan et al. 2012); Column 7-9: from literature sources - <sup>1</sup>Van Altena et al. (1988), <sup>2</sup>Hillenbrand (1997), <sup>3</sup>Houk & Swift (1999), <sup>4</sup>Rebull et al. (2000), <sup>5</sup>Da Rio et al. (2016), <sup>6</sup>Nesterov et al. (1995), <sup>7</sup>Cannon & Pickering (1993).

Gaia DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3216203177762381952	108.0 ±0.8	59.1 ±13.7 <sup>a</sup>	yes	0.2	0.4 <sup>+9.0</sup> <sub>-0.3</sub>	-	-	-
3015321754828860928	82.9 ±1.0	-	-	0.3	0.4 <sup>+1.1</sup> <sub>-0.3</sub>	-	-	-
3013484917577226240	68.6±0.8	-	-	0.6	3.0 <sup>+5.0</sup> <sub>-2.0</sub>	-	-	-
2998984592590172288	60.3 ±0.5	-80.5 ±6.6 <sup>a</sup>	no	-	0.2 <sup>+0.3</sup> <sub>-0.1</sub>	-	-	-
3012438796685305728	59.9 ±0.5	-4.6 ±6.8 <sup>a</sup>	yes	0.5	10.0 <sup>+40.0</sup> <sub>-7.5</sub>	-	-	-
2986587942582891264	56.0 ±0.4	-31.3 ±6.7 <sup>a</sup>	yes	1.3	9.0 <sup>+10.0</sup> <sub>-6.0</sub>	-	-	-
3016780428803888768	55.4 ±0.4	-	-	0.2	1.0 <sup>+5.0</sup> <sub>-0.7</sub>	-	-	-
3209498394512739968	47.2 ±0.5	-	-	0.1	2.0 ±1.0	-	-	M3 <sup>4</sup>
3016792935748254336	46.0 ±0.6	-	-	0.2	0.3 ±0.2	-	-	-
3003060825792025088	44.5 ±0.5	28.4±6.8 <sup>a</sup>	no	-	8.0 ±5.0	-	-	-
3122639449820663040	43.6 ±1.3	-	-	1.6	2.0 <sup>+2.0</sup> <sub>-0.5</sub>	-	-	B5 <sup>3</sup>
3320554665258533376	37.0 ±1.5	-	-	2.0	5.0 ±2.0	-	-	-
2998697894931641600	36.5 ±1.1	-	-	1.2	2.0 ±0.2	-	-	A2-A9 <sup>3</sup>
3023329085698084992	34.2 ±0.5	-	-	0.2	1.0 <sup>+6.0</sup> <sub>-0.9</sub>	-	-	-
2994091353528613120	33.1 ±1.5	-	-	2.4	4.0 pm1.8	-	-	-
3209936343738052992	32.9 ±0.5	-	-	0.3	1.5 <sup>+8.5</sup> <sub>-1.0</sub>	-	-	-
3017250019053914368	32.0 ±0.5	32.4 ±6.7 <sup>a</sup>	yes	in cluster	8.0 <sup>+42.0</sup> <sub>-7.0</sub>	1.9-5.7 <sup>2,5</sup>	1.9-2.4 <sup>2,5</sup>	G6 <sup>2</sup>
3222673430030590592	30.3 ±0.4	4.2 ±6.6 <sup>a</sup>	yes*	1.4*	0.3 ±0.2	-	-	K0 <sup>6</sup>
3017265515291765760	30.1 ±0.4	12.3 ±6.6 <sup>a</sup>	yes	in cluster	2.2 <sup>+4.8</sup> <sub>-2.0</sub>	0.3 <sup>2</sup>	2.5 <sup>2</sup>	K1 <sup>1</sup>
3016354436766366848	28.1 ±0.4	**b	**	0.4	4.0 ±1.0	-	-	A0 <sup>3</sup>
3016016714897329152	27.9 ±0.3	-29.0 ±6.6 <sup>a</sup>	no	-	7.0 ±4.0	-	-	-
3122561556293863552	27.8 ±1.5	-22.0±11.3 <sup>a</sup>	yes	2.0	1.8 <sup>+2.1</sup> <sub>-1.1</sub>	-	-	-
3122421987035894784	27.5 ±0.4	49.3 ±6.6 <sup>a</sup>	no	-	0.4 <sup>+0.6</sup> <sub>-0.2</sub>	-	-	-
2993831456467526016	23.8 ±0.4	52.0 ±6.6 <sup>a</sup>	no	-	4.0 <sup>+5.0</sup> <sub>-2.0</sub>	-	-	-
2982855577345989504	22.9 ±0.6	-24.7 ±6.9 <sup>a</sup>	yes	3.6	5.5 <sup>+2.5</sup> <sub>-1.6</sub>	-	-	-
3216868764551493504	22.6 ±0.4	44.2 ±6.8 <sup>a</sup>	no	-	5.0 <sup>+40.0</sup> <sub>-4.5</sub>	-	-	-
3208970285334738944	21.8 ±0.7	-30.5 ±19.7 <sup>a</sup>	yes	0.5	2.8 <sup>+8.2</sup> <sub>-2.4</sub>	-	-	-
3218763120006174976	20.4 ±1.0	62.1 ±10.3 <sup>a</sup>	no	-	1.5 <sup>+8.5</sup> <sub>-1.3</sub>	-	-	-
3015532208227085824	19.0 ±0.4	-55.5 ±6.7 <sup>a</sup>	yes	0.9	6.5 ±3.5	-	-	-

\*Age estimate is smaller than the flight time; \*\*multiple and varying RV measurements, possibly indicating a binary system

Table 5.4: WW star 2D candidates. Column 2+3: velocity in ONC rest frame [rf]; Column 3: RV sources - <sup>a</sup>Gaia DR2, <sup>b</sup>Cottaar et al. (2015), <sup>c</sup>Gontcharov (2006), <sup>d</sup>Kounkel et al. (2018); Column 4: indication of 3D-candidate status; Column 5: minimum flight time since ejection (crossing of search boundary); Column 6: age from PARSEC isochrones (Bressan et al. 2012); Column 7-9: from literature sources - <sup>1</sup>Van Altena et al. (1988), <sup>2</sup>Hillenbrand (1997), <sup>3</sup>Houk & Swift (1999), <sup>4</sup>Rebull et al. (2000), <sup>5</sup>Da Rio et al. (2010), <sup>6</sup>Tetzlaff et al. (2011), <sup>7</sup>Da Rio et al. (2012), <sup>8</sup>Hsu et al. (2012), <sup>9</sup>Hsu et al. (2013), <sup>10</sup>Da Rio et al. (2016), <sup>11</sup>Sota et al. (2011), <sup>12</sup>Hohle et al. (2010).

Gaia DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3218162816720967040	29.7 ±1.3	-	-	1.3	15.0 <sup>+40.0</sup> <sub>-11.2</sub>	-	-	-
3209653627514662528	26.0 ±0.5	12.4 ±6.6 <sup>b</sup>	no	-	0.4 ±0.1	0.2 <sup>10</sup>	0.3 <sup>10</sup>	-
3222368036380921600	25.9 ±1.4	-11.5 ±14.7 <sup>a</sup>	yes	1.4	3.0 <sup>+6.0</sup> <sub>-2.0</sub>	-	-	-
3231583219428074752	25.7 ±1.5	-	-	2.8	8.0 <sup>+4.0</sup> <sub>-5.0</sub>	-	-	-
3013902388397518208	25.7 ±0.7	-	-	1.0	1.0±0.5	-	-	-
3209590577396377856	25.2 ±1.1	-	-	0.1	5.0 <sup>+15.0</sup> <sub>-4.0</sub>	-	-	M6 <sup>4</sup>
3209228906787768832	24.4 ±0.4	-9.4 ±6.8 <sup>a</sup>	yes	0.8	10.0 <sup>+41.0</sup> <sub>-8.3</sub>	-	-	-
3181732702253990144	23.1 ±1.5	-	-	2.8	1.0 <sup>+2.0</sup> <sub>-0.5</sub>	-	-	B9 <sup>3</sup>
3334687998177678080	22.6 ±1.8	-	-	3.4	4.0 <sup>+5.5</sup> <sub>-2.1</sub>	-	-	-
3017260292611534848	21.7 ±0.4	3.5 ±6.7 <sup>a</sup>	yes	0.1	3.0 <sup>+7.0</sup> <sub>-1.2</sub>	7.7 <sup>10</sup>	1.9 <sup>10</sup>	-
3004263966389331456	20.0 ±0.4	-18.6 ±6.6 <sup>a</sup>	yes	3.0	6.0±2.5	-	-	-
3023589704313257600	19.5 ±0.6	-	-	1.0*	0.2 <sup>+0.4</sup> <sub>-0.1</sub>	-	-	-
3023540054490074752	18.6 ±1.7	-	-	0.6	10.0 <sup>+20.0</sup> <sub>-8.0</sub>	-	-	-
3014981937018182144	18.3 ±0.9	-	-	1.7	1.2 <sup>+2.9</sup> <sub>-0.7</sub>	-	-	-
3209624872711454976	18.2 ±0.4	14.5 ±6.6 <sup>b</sup>	yes	0.1	1.7 ±0.6	0.4 <sup>10</sup>	0.5 <sup>10</sup>	-
3209836391259368960	17.4 ±1.3	-	-	0.3	10.0 <sup>+20.0</sup> <sub>-8.0</sub>	-	-	-
3017166907140904320	17.3 ±0.4	5.1 ±6.6 <sup>b</sup>	yes	0.2	2.8 <sup>+0.3</sup> <sub>-0.9</sub>	1.0 <sup>10</sup>	0.6 <sup>10</sup>	K7.5 <sup>8</sup>
3015018014743100544	17.1 ±1.5	-22.1 ±22.6 <sup>a</sup>	yes	1.6	0.9 <sup>+3.1</sup> <sub>-0.4</sub>	-	-	-
3017242051888552704	16.7 ±0.4	-5.6 ±7.0 <sup>b</sup>	yes	in cluster	4.0 <sup>+20.0</sup> <sub>-3.5</sub>	1.8 <sup>10</sup>	0.7 <sup>10</sup>	-
3209497088842680704	16.6 ±0.4	-	-	0.2	3.0 ±1.5	-	-	M2 <sup>4</sup>
3209424108758593408	16.4 ±0.4	8.2 ±6.6 <sup>b</sup>	yes	in cluster	1.0 <sup>+9.0</sup> <sub>-0.9</sub>	0.5-2.5 <sup>5,11</sup>	1.1-2.3 <sup>5,11</sup>	G9 <sup>8</sup>
3214878167468186880	16.4 ±1.0	-	-	2.4	2.1±1.1	-	-	-
3015625563635553024	16.4 ±0.4	-4.2 ±6.6 <sup>b</sup>	yes	0.9	0.8 <sup>+3.2</sup> <sub>-0.4</sub>	1.2 <sup>10</sup>	0.3 <sup>10</sup>	M2.9 <sup>8</sup>
3016070590967059968	16.2 ±0.6	-5.1 ±6.6 <sup>d</sup>	yes	1.0	0.5 <sup>+1.5</sup> <sub>-0.4</sub>	-	-	-
3220151695816273152	16.1 ±0.4	8.9 ±6.6 <sup>a</sup>	yes	2.2	2.1 <sup>+3.0</sup> <sub>-1.5</sub>	-	-	-
2984454031031531008	15.0 ±1.2	-	-	3.7	2.8 <sup>+3.2</sup> <sub>-1.3</sub>	-	-	-
3208349129984108800	14.9 ±0.8	-	-	2.5	2.1±1.1	-	-	-
3017402614955763200	14.7 ±0.4	-13.8 ±8.3 <sup>a</sup>	yes	0.1	6.0 ±4.0	-	-	K7 <sup>4</sup>
3015334914608642688	14.2 ±0.6	16.6 ±7.8 <sup>b</sup>	yes	1.6	1.5 <sup>+3.5</sup> <sub>-1.0</sub>	-	-	M1.6 <sup>8</sup>
3209424108758593536	14.1 ±0.4	-4.4 ±6.6 <sup>b</sup>	yes	in cluster	4.0 <sup>+6.0</sup> <sub>-3.3</sub>	0.5-2.5 <sup>5,11</sup>	0.7 <sup>5,11</sup>	K7 <sup>2</sup>
3017384129418196992	14.1 ±0.4	-	-	0.1	2.5 <sup>+1.5</sup> <sub>-0.5</sub>	-	-	M2 <sup>4</sup>
3012142379518284288	13.5 ±0.7	-	-	2.0	2.1 <sup>+3.0</sup> <sub>-1.1</sub>	-	-	-
3209074803362165888	13.2 ±0.5	-	-	0.6	0.8 <sup>+9.0</sup> <sub>-0.7</sub>	-	-	-
3014834946056441984	13.2 ±0.6	-	-	1.6	1.0 <sup>+0.8</sup> <sub>-0.4</sub>	1 <sup>6</sup>	3.8 <sup>6</sup>	A5 <sup>6</sup>
3209531650444835840	13.0 ±0.4	-4.2 ±8.0 <sup>a</sup>	yes	in cluster	0.3 <sup>+3.7</sup> <sub>-0.2</sub>	-	3.8 <sup>2</sup>	K0 <sup>1</sup>
2984723926777044480	13.0 ±0.4	-11.4 ±6.6 <sup>a</sup>	yes*	3.8*	0.4±0.3	-	-	-
3015714967674577024	12.7 ±0.8	-	-	1.0	20.0 <sup>+80.0</sup> <sub>-17.0</sub>	-	-	K8 <sup>8</sup>
3012280432650658304	12.6 ±3.6	-	-	3.0	14.0 <sup>+36.0</sup> <sub>-10.5</sub>	-	-	-
3216889827071056896	12.6 ±0.7	-	-	1.4	1.0 <sup>+3.0</sup> <sub>-0.9</sub>	-	-	-
3013899158582179712	12.4 ±1.5	-	-	1.9	5.0 <sup>+7.0</sup> <sub>-4.0</sub>	-	-	-
3216174629116142336	12.1 ±0.6	-	-	1.7	2.5 ±1.5	-	-	-
3016101579155228928	12.1 ±1.4	-	-	1.1	1.0 <sup>+10.0</sup> <sub>-0.8</sub>	-	-	-
3219378365481960832	11.7 ±0.9	-	-	3.3	2.8 <sup>+7.0</sup> <sub>-1.8</sub>	-	-	-
3017341385903759744	11.7 ±0.4	1.0 ±6.6 <sup>d</sup>	yes	in cluster	3.0±1.0	0.7-0.9 <sup>5,11</sup>	0.5 <sup>5,11</sup>	K6 <sup>4</sup>
3017340664349130368	11.5 ±0.5	14.4 ±18.4 <sup>a</sup>	yes	0.3	2.0 <sup>+10.0</sup> <sub>-1.7</sub>	-	-	K5 <sup>4</sup>
3209527291054667136	11.1 ±0.5	10.6 ±6.6 <sup>b</sup>	no	-	1.2 <sup>+1.8</sup> <sub>-1.1</sub>	0.8-2 <sup>5,11</sup>	2.5-3.2 <sup>5,11</sup>	-

Table 5.4 - continued

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3017260022031719040	10.6 ± 0.4	8.7 ± 6.6 <sup>d</sup>	yes	in cluster	2.5 ± 1.5	0.7-1.5 <sup>5,7</sup>	0.3-0.5 <sup>5,7</sup>	M3 <sup>5</sup>
3017270879709003520	10.5 ± 0.4	-1.2 ± 6.8 <sup>b</sup>	yes	in cluster	0.9 <sup>+5.0</sup> <sub>-0.8</sub>	0.3 <sup>10</sup>	1.4 <sup>10</sup>	K4 <sup>1</sup>
3016961676421884672	10.4 ± 0.6	-	-	0.8	0.6 <sup>+1.3</sup> <sub>-0.5</sub>	-	-	-
3016424530632449280	10.2 ± 1.5	-5.2 ± 7.1 <sup>c</sup>	no	-	~4	-	24.1 <sup>12</sup>	O9.7 <sup>11</sup>
3215804677813294976	10.0 ± 0.8	-	-	1.2	0.5 <sup>+1.5</sup> <sub>-0.1</sub>	-	-	-
3209680466766984448	9.5 ± 0.4	15.9 ± 6.6 <sup>a</sup>	yes	1.4	0.4 <sup>+1.5</sup> <sub>-0.3</sub>	-	-	-
3017252600328207104	9.3 ± 0.4	-5.0 ± 6.6 <sup>d</sup>	yes	0.1	1.2 <sup>+2.8</sup> <sub>-0.5</sub>	0.1 <sup>10</sup>	0.3 <sup>10</sup>	M3 <sup>2</sup>
3017174741161205760	8.5 ± 0.4	-17.3 ± 9.6 <sup>a</sup>	yes	0.3	4.0 <sup>+20.0</sup> <sub>-3.5</sub>	-	-	K3 <sup>9</sup>
3217017610938439552	5.9 ± 0.7	-8.4 ± 11.8 <sup>a</sup>	yes	3.2	1.5 <sup>+6.5</sup> <sub>-1.2</sub>	-	-	-
3017367151399567872	3.6 ± 0.4	12.2 ± 9.4 <sup>a</sup>	yes	in cluster	4.0 ± 2.0	1.7 <sup>10</sup>	2.7 <sup>10</sup>	-
3209529112120792320	3.0 ± 0.4	12.1 ± 6.6 <sup>b</sup>	yes	in cluster	20.0 <sup>+50.0</sup> <sub>-19.0</sub>	6.2 <sup>10</sup>	1.1 <sup>10</sup>	-

\*Age estimate is smaller than the flight time

quence of my chosen approach. Its mass is reported to be  $\sim 24 M_{\odot}$  (Hohle et al. 2010). It is located at the very edge of my search field with a distance of just under 100 pc to the ONC centre.

Several of my faintest 2D candidates appear not to have any further information than that contained within *Gaia* DR2. The faintest candidates will also likely be the stars that have the lowest mass, so information from other sources is critical. The faintest 2D RW candidate with a spectral type identification is V\* HP Ori (*Gaia* DR2 3209498394512739968). Rebull et al. (2000) suggested it to be a M3-type star. Its absolute magnitude is  $\sim 3.5$  mag. The faintest 2D WW candidate 2MASS J05332200-0458321 (*Gaia* DR2 3209590577396377856) has an absolute magnitude of  $\sim 10$  mag and a suggested spectral type of M6 (Rebull et al. 2000).

### Comparison to McBride & Kounkel 2019

McBride & Kounkel (2019) identified 9 “RW stars” with an origin close to the Trapezium cluster at the centre of the ONC. Of these stars, I successfully trace back 7 in my analysis. However, none fit my velocity requirement of a RW (velocity  $> 30$  km s<sup>-1</sup>), and all are identified as WW candidates instead:

- V1961 Ori (*Gaia* DR2 3209424108758593408): this star has previously been identified as a RW candidate by Kounkel et al. (2017). It has been suggested as a spectroscopic binary by Kounkel et al. (2019) but has a RUWE  $\approx 1.1$  in *Gaia* DR2 indicating a good fit to a single-star model or that this system’s binary status does not affect its astrometric quality. It has the second lowest  $M_{G,0}$  value of  $\sim 3.7$  mag of all McBride & Kounkel (2019) stars found in my analysis, but as a binary it will be brighter than each of the individual stars. Its 2D-velocity in the ONC rest

frame is  $\sim 16 \text{ km s}^{-1}$ .

- Brun 259 (*Gaia* DR2 3209424108758593536): [Duchêne et al. \(2018\)](#) show it is unlikely to be binary system. It has a rest-frame 2D-velocity of  $\sim 15 \text{ km s}^{-1}$  and an absolute magnitude of  $\sim 2.4 \text{ mag}$ .
- V1321 Ori (*Gaia* DR2 3209531650444835840): [Janson et al. \(2012a\)](#) suggest this star is possibly a binary, its RUWE  $\approx 1$  indicates a good fit to the single star model. This low value indicates that the binarity of this system is not affecting the astrometry quality. It has a 2D-velocity of  $\sim 14 \text{ km s}^{-1}$  and an absolute magnitude  $\sim 1.1 \text{ mag}$ .
- V1440 Ori (*Gaia* DR2 3209624872711454976): this star has an absolute magnitude  $\sim 1.1 \text{ mag}$  and a 2D-velocity of  $\sim 18 \text{ km s}^{-1}$  in the ONC rest frame.
- 2MASS J05351295-0417499 (*Gaia* DR2 3209653627514662528): this star is one of my fastest 2D-candidates with a 2D-velocity of  $\sim 26 \text{ km s}^{-1}$ . It is one of the youngest in my WW list with an isochronal age of  $\sim 0.4 \text{ Myr}$  and a magnitude of  $\sim 3.4 \text{ mag}$ .
- CRTS J053223.9-050523 (*Gaia* DR2 3209497088842680704): the 2D-velocity of this star is  $\sim 17 \text{ km s}^{-1}$  and it has a magnitude of  $\sim 2.9 \text{ mag}$ .
- Haro 4-379 (*Gaia* DR2 3017166907140904320): this star is the faintest in this [McBride & Kounkel \(2019\)](#) group with a  $M_{G,0}$  value of  $\sim 3.8 \text{ mag}$  and a 2D-velocity of  $\sim 17 \text{ km s}^{-1}$ .

The final two candidates of [McBride & Kounkel \(2019\)](#) have been excluded from my data set from the outset due to their high RUWE-values, which indicate that the astrometry might be unreliable. 2MASS J05382070-0610007 (*Gaia* DR2 3016971567730386432) has a RUWE  $\approx 8$ , and V360 Ori (*Gaia* DR2 3209528081326372864) has a RUWE  $\approx 28$ , this latter star is also a known binary ([Daemgen et al. 2012](#)).

### 5.4.2 3D-candidates

Fig. 5.4 shows the CAMD of RW and WW stars that can be traced back in 3D to my search region. Ten of the 29 RW 2D-candidates are successfully traced back in all three dimensions using *Gaia* DR2 RV.

The fastest RW star (*Gaia* DR2 3216203177762381952) in my sample has an ONC rest frame space velocity of  $\sim 123 \text{ km s}^{-1}$  and an absolute magnitude of  $\sim 4.4 \text{ mag}$ . However,

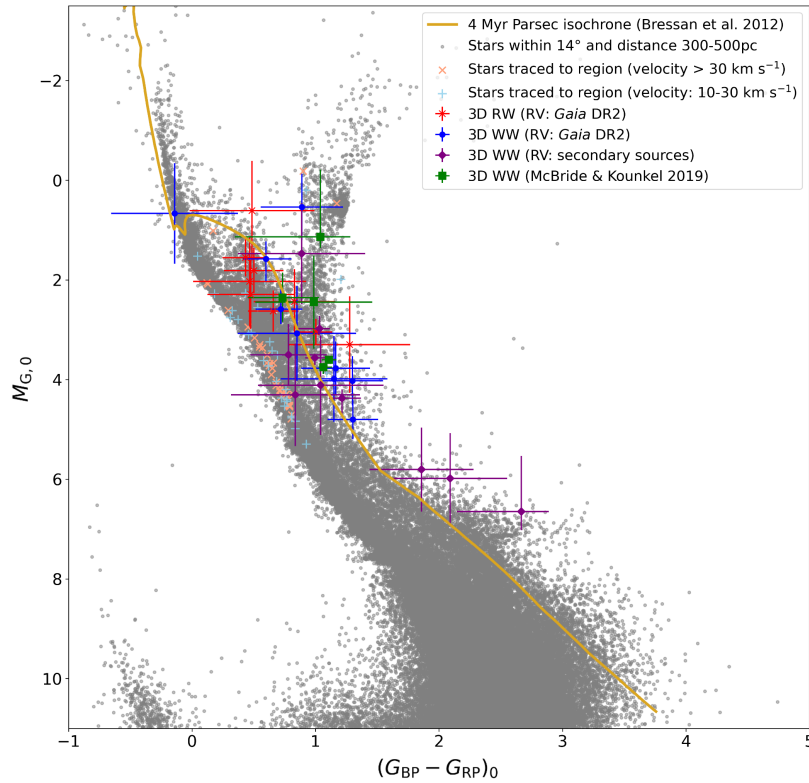


Figure 5.4: CAMD showing all 3D-candidates (at RW and WW velocities). I magnitude limit the diagram to  $-3.5 \text{ mag} < M_{G,0} < 11 \text{ mag}$ , which corresponds to a  $G$ -magnitude  $\approx 19 \text{ mag}$  at the fainter end. Around this apparent magnitude value, the typical uncertainties in the 5-parameter astrometry increase quickly. However, my faintest 3D-candidate is much brighter than this limit and has an absolute magnitude of  $\sim 7 \text{ mag}$ . My confirmed 3D-candidates are all located towards the upper end of the 4 Myr isochrone. All of my 3D RW stars (red “x”) are identified using *Gaia* DR2 RV. Among the 3D WW stars I separately show the candidates with RV from *Gaia* DR2 (blue square) and from secondary sources (Cottaar et al. 2015; Kounkel et al. 2018) (purple dot). Finally I show the five 3D WW stars what I have in common with McBride & Kounkel (2019) (green square). Many of my identified candidates have large errors in magnitude and colour, which are predominantly driven by the errors in the extinction and reddening. Interestingly, I also find a small number of 3D trace-backs that sit along the MS, below the isochrone (with RW-velocity: orange “x”, with WW-velocity: light-blue “+”) and even when considering their errors they are too old to have been born in the ONC. These might be past visitors to the ONC or could possibly belong to an older population of stars, as their estimated isochronal ages are between 5-40 Myr.

there is no additional information available about this star. Two of the 10 3D RWs have information available from the literature about their spectral type, mass and/or age. Brun 609 (*Gaia* DR2 3017250019053914368) has a G6 spectral type (Van Altena et al. 1988), a mass estimate of  $\sim 1.9\text{-}2.4 M_{\odot}$  and an age estimate of 1.9-5.7 Myr (Hillenbrand 1997; Da Rio et al. 2016). It has a rest frame space velocity of  $\sim 46 \text{ km s}^{-1}$  and is still located within the search boundary.

BD-05 1307 (*Gaia* DR2 3017265515291765760) is a K1-type star (Van Altena et al. 1988) with a mass estimate of  $\sim 2.5 M_{\odot}$  and an estimated age of  $\sim 0.3$  Myr (Hillenbrand 1997). It has a rest frame space velocity of  $\sim 33 \text{ km s}^{-1}$ . It is also the brightest 3D RW star in my data set with an absolute magnitude of  $\sim 0.7$  mag and is still located within the central 2.5 pc search boundary at a distance of  $\sim 400$  pc.

I find another potential 3D RW star *Gaia* DR2 3209532135777678208 (TYC 4774-868-1). This star could be an example of a special case described in Chapter 4. This case is a star that appears bound in proper motion or 2D-velocity space and is still located in the cluster; however, a very high RV turns it into RWs. If I only consider its 2D-velocity of  $\sim 3.4 \text{ km s}^{-1}$ , TYC 4774-868-1 appears to be still bound to the ONC. It is also still located in the central region. However, its *Gaia* DR2 RV in the ONC rest frame would be high enough to turn it into a RW star. There is a caveat as Cottaar et al. (2015) stated a much lower RV of  $\sim 28.3$  km/s, which results in a RV of  $\sim 6.5$  km/s in the ONC rest frame. This leads me to not consider this star as a RW until further clarification of its RV. The data for this star is contained in Table A.2 in Appendix A.

The brightest 2D RW candidate HD 288089 (*Gaia* DR2 3222673430030590592) also traces back in three dimensions; however, its isochronal age is smaller than its flight time. While it traces back to the ONC, it cannot have been born there, but must have instead come from somewhere in-between the ONC and its current position. It is not counted in the final list of 3D RWs, and this leaves me with nine 3D RWs.

Of the 57 WW 2D-candidates, 27 are also 3D-candidates. Fourteen of these using *Gaia* DR2 RV, another nine candidates using RVs from Cottaar et al. (2015) and four candidates using RVs from Kounkel et al. (2018). Five of the seven 2D WW candidates I have in common with McBride & Kounkel (2019) are also 3D WW candidates:

- V\* V1961 Ori *Gaia* DR2 3209424108758593408): has a G9 spectral type (Hsu et al. 2012) with an estimated mass of  $\sim 1.1\text{-}2.3 M_{\odot}$  and an age of  $\sim 0.1\text{-}2.5$  Myr (Da Rio et al. 2010, 2016). Its space velocity in my rest-frame is  $\sim 18 \text{ km s}^{-1}$ .
- Brun 259 *Gaia* DR2 3209424108758593536): has a spectral type K7 (Hillenbrand 1997) with a mass of  $\sim 0.7 M_{\odot}$  and an age estimate of  $\sim 0.5\text{-}2.5$  Myr (Da Rio et al.

- 2010, 2016). It has a rest-frame space velocity of  $\sim 15 \text{ km s}^{-1}$ .
- V1321 Ori *Gaia* DR2 3209531650444835840): this star has a spectral type K0 (Van Altena et al. 1988) and a mass estimate of  $3.8 M_{\odot}$  (Hillenbrand 1997). Its space velocity is  $\sim 14 \text{ km s}^{-1}$ .
  - V1440 Ori *Gaia* DR2 3209624872711454976) is one of the faster 3D WW stars with an ONC rest frame space velocity of  $\sim 23 \text{ km s}^{-1}$ , it is a sub-solar mass ( $\sim 0.5 M_{\odot}$ ) star with an age estimate of  $\sim 0.4 \text{ Myr}$  (Da Rio et al. 2016), suggesting it left the cluster shortly after its birth.
  - Haro 4-379 *Gaia* DR2 3017166907140904320): has a space velocity  $\sim 18 \text{ km s}^{-1}$  with a spectral type of K7.5 (Hsu et al. 2012), has a sub-solar mass ( $\sim 0.6 M_{\odot}$ ) and an age estimate from literature of  $\sim 1 \text{ Myr}$  (Da Rio et al. 2016).

The slowest confirmed 3D WW star is Brun 519 (*Gaia* DR2 3017252600328207104), and it has a rest frame space velocity just above the  $10 \text{ km s}^{-1}$  lower velocity limit. It is also one of the youngest 3D WW stars in my sample with an age and mass estimate of  $\sim 0.1 \text{ Myr}$  and  $\sim 0.3 M_{\odot}$ , respectively (Da Rio et al. 2016). Its spectral type is suggested to be a M3 (Hillenbrand 1997).

Like I have for the 3D RW candidates, I also find one 3D WW candidate BD-13 1169 (*Gaia* DR2 2984723926777044480) where the flight time since ejection is considerably larger than the estimated isochronal age, strongly indicating that this star was not born in the ONC. I exclude this star from the 3D WW list and am left with 26 3D WW stars.

I also find three further potential 3D WW stars that I might still consider to be bound to the ONC if I only had information about their proper motion (or 2D-velocities). Here a larger RV (from *Gaia* DR2 or secondary sources) can turn these stars into 3D candidates. These WW candidates (like the RW candidate TYC 4774-868-1) show very different RV measurements in different surveys, which could have an alternative explanation of a bound binary system. These candidates are shown separately in Table A.2 in Appendix A.

I also find several older 3D RW/WW stars below the isochrone in Fig. 5.4 along the main sequence. These stars are possibly past “visitors” to the ONC, having travelled through the ONC from their origin. Information about these visitors is provided in Appendix A in Table A.3.

## 5.5 Discussion

The ejected RW stars in my  $N$ -body simulations quickly leave my 100 pc search region, as most of them have been ejected during the very early dynamical evolution of my simulated star-forming regions. After 2 Myr in the simulations, an average of  $\sim 11$  (and maximum of 15) RWs are still located within the search area but this average reduces to  $\sim 4$  (maximum of 10) and  $\sim 1$  (maximum of 3) RWs after 3 and 4 Myr, respectively. In contrast, all WW stars remain within my 100 pc boundary until at least 3 Myr, only travelling past this boundary towards the end of my simulations.

My analysis of *Gaia* DR2 finds a total of 29 RW 2D-candidates by tracing back the positions on the sky for up to 4 Myr. Of these RW 2D-candidates, seven have RV measurements but do not trace back to the ONC in 3D. Nine RW stars can be traced back in three dimensions, all using RVs from *Gaia* DR2.

I find a further potential RW star TYC 4774-868-1 (*Gaia* DR2 3209532135777678208), which appears bound when considering only its proper motion, but it turns into a RW in three dimensions using its *Gaia* DR2 RV measurement. However, [Cottaar et al. \(2015\)](#) stated a much lower RV of  $\sim 28.3$  km/s, which results in a RV of  $\sim 6.5$  km/s in the ONC rest frame instead of 58.5 km/s using *Gaia* DR2 RV. Using this secondary RV source results in a rest-frame space velocity  $< 10$  km/s, which is below my lower boundary for WW stars. Detailed information about this star is shown in Table A.2 in Appendix A.

While simulations show (see Chapter 4) that I can expect to find RW stars where only one of their velocity components would identify them as a RW or WW star, there can be other explanations such as the possible binarity of the system. In this case, the difference in RV measurements from two surveys hint at this star being part of a binary, located in the ONC centre.

I also find a potential ejected binary HD 36697 (*Gaia* DR2 3016354436766366848) based on the RV measurements by [Cottaar et al. \(2015\)](#). This could be either a 3D RW or WW system. Its identification depends on the system's radial velocity, as its 2D-velocity is high enough to put it at least into the WW-velocity regime.

Two of these 9 RWs have mass estimates from the literature putting them all in the low/intermediate-mass category with masses of 1.9-2.5  $M_{\odot}$  ([Hillenbrand 1997](#); [Da Rio et al. 2016](#)). Based on their position on the CAMD, the other 7 3D RW stars are also within the low/intermediate-mass range, as none are consistent with the part of the isochrone that has already reached the MS, i.e. quickly evolving high-mass stars. When I compare the CAMD-positions of the 3D RWs to the 3D WWs, I see that they do not extend below the lowest masses of any identified WWs, so I can conclude that



none of the RWs identified in my search have a mass  $< 0.3 M_{\odot}$ . This is a consequence of the magnitude-limit I applied to the CAMD, removing any fainter candidates with unreasonably large uncertainties in their astrometric and photometric measurements.

Fig. 5.1 shows that a large number (about half) of all RW and WW stars in my simulations have a mass below  $0.3 M_{\odot}$ . This suggests that I can expect to find a large number of RWs/WWs at these low masses. To compare my findings from *Gaia* DR2 to simulations covering the same mass range, I consider only low/intermediate-mass RWs within 100 pc with masses between  $0.3\text{--}8 M_{\odot}$ . I find a RW average of  $7.3 \pm 2.3$  at 1 Myr, decreasing to  $5.5 \pm 2.5$  at 2 Myr. The number further reduces to  $2.2 \pm 1.4$  ( $0.8 \pm 0.9$ ) RWs at 3 (4) Myr. The maximum number of RWs from a single simulation reduces to ten at 2 Myr and five (three) RWs at 3 (4) Myr.

Comparing the nine 3D RWs I find in *Gaia* DR2 with my simulation results for the number of RWs gives me an age estimate of  $\sim 1.3$  Myr when I compare it to the average, and  $\sim 2.4$  Myr when compared to the maximum number from my simulations. This age range is in good agreement with the mean age range (2-3 Myr) of [Da Rio et al. \(2010\)](#) and [Reggiani et al. \(2011\)](#). However, there are caveats related to the age estimate from my gas-free  $N$ -body simulations.  $N$ -body simulations track the dynamical evolution of a star-forming region under the effect of gravity. In addition, I assume that at time  $t = 0$  Myr all my stars have instantaneously formed and the cluster is gas-free. These assumptions neglect the time it takes for stars to form from the GMC itself and any effect the gas might have on the dynamical evolution (e.g. [Sills et al. 2018](#)). At best, my age estimate could be considered as a lower age estimate for my star-forming regions.

Four of the identified RWs are located within my applied cluster boundary on the sky or are still located within just a few pc of the centre. Several of my RWs only trace back to the region when I consider their large radial velocity and distance errors, making these candidates less certain.

*Gaia* DR2 2982855577345989504 is an interesting RW star traced back to the ONC. Based on its flight time, it has left the ONC  $\sim 3.6$  Myr ago. However, my age estimates suggests that the ONC has an upper age less than this age. This would make this star a past (young) visitor instead of a star ejected from the ONC. Leaving this star out of my list of RWs leaves me with eight RWs. This change in successfully traced back RWs, increases the age estimate from the  $N$ -body simulations slightly to  $\sim 1.6$  Myr from the averages and  $\sim 2.5$  Myr from the maximum.

I have not found any high-mass ( $> 8 M_{\odot}$ ) RW stars in my *Gaia* analysis (neither 2D, nor 3D). I find a B5-star HD 41288 (*Gaia* DR2 3122639449820663040) in 2D ([Houk](#)

& Swift 1999); however, without an explicit mass estimate from literature. Based on its spectral type, its mass is unlikely to be within my high-mass category. This star does not have RV information so cannot be confirmed in 3D.

It is possible to estimate the mass of a main-sequence star using a mass-luminosity relationship (MLR). This requires an estimate of the star's luminosity  $L$ , which can be derived from the absolute  $G$ -magnitude value  $M_G$  of a star using a temperature-dependent bolometric correction  $BC_G$  and the absolute solar bolometric magnitude  $M_{\text{bol}\odot}$ . *Gaia* DR2 provides luminosity estimates that are derived using Eq. 4 in Andrae et al. (2018). This equation can be written as:

$$\log_{10} \frac{L}{L_{\odot}} = 0.4 (M_{\text{bol}\odot} - M_{\text{bol}}), \quad (5.1)$$

where  $M_{\text{bol}} = M_G + BC_G(T_{\text{eff}})$ .

For HD 41288, the luminosity is unfortunately not provided in DR2, so I calculate a rough  $BC_G$  estimate based on the  $BC_G$ ,  $T_{\text{eff}}$  relationship shown in Fig. 8 in Andrae et al. (2018). The provided  $T_{\text{eff}}$ -value of  $\sim 8941$  K is outside the range of values covered in Fig. 8, however, at higher temperatures  $BC_G$  appears to be close to 0 mag, so I use  $BC_G = 0$  mag. As stated in Andrae et al. (2018), the value for  $M_{\text{bol}\odot}$  is set to 4.74 mag as defined by the IAU. I derived a  $M_{G,0}$  of  $\sim 0$  mag from the CAMD. Using these values would result in  $L \approx 80 L_{\odot}$ . Compared to typical luminosity values for the stated spectral type B5, this is much lower than one would expect. Also its  $T_{\text{eff}}$  is lower than one would expect for this spectral type. Regardless of these issues with the existing spectral type identification, the luminosity allows me to derive a mass estimate. The CAMD in Fig. 5.2 shows that this star is already located on the main-sequence, so its mass  $M_{\text{star}}$  can be estimated from its luminosity  $L$ . Using the classical MLR:

$$\frac{L}{L_{\odot}} = \left( \frac{M_{\text{star}}}{M_{\odot}} \right)^{3.5} \quad (5.2)$$

results in a mass estimate for HD 41288 of  $\sim 3.5 M_{\odot}$ , which shows that this star does not belong into my defined high-mass class ( $> 8 M_{\odot}$ ). The estimates for  $T_{\text{eff}}$ , luminosity and mass all point towards this star being an A-type star instead of a B5.

Two high-mass OB-RWs (AE Aur and  $\mu$  Col) have previously been postulated to have originated in the ONC, and Blaauw & Morgan (1954) and Hoogerwerf et al. (2001) suggested these stars were ejected  $\sim 2.5$  Myr ago. I find a maximum of two high-mass

RW stars at the end of one of my simulations; however, these two RWs are not ejected at the same time. The second RW only gets ejected at the very end of this simulation and is in fact a RW-binary. The ejections of high-mass RW stars are rare events in my simulations (as in reality), and in several of my simulations I do not see any high-mass RWs. Due to the IMF, the number of high-mass stars in my simulated regions is low to begin with, which in turn results in only a small number of high-mass RWs, which also differs in-between individual simulations.

At lower WW velocities, I find 57 2D-candidates with 30 of them having RV information (either from *Gaia* DR2 or literature sources). I find one O-star  $\nu$  Ori (*Gaia* DR2 3016424530632449280) WW-candidate in 2D. However, its RV (Gontcharov 2006) prevents it from turning into a 3D-candidate. It is not a known binary (Bodensteiner et al. 2018). The Simbad database (Wenger et al. 2000) lists further RV measurements for this star, none of which are large enough to change this. To become a WW star from the ONC, this O-star would require its RV to point opposite to its current direction.

I can successfully trace back 26 WW stars to the ONC in 3D, with stellar masses (where available from the literature) between  $\sim 0.3$ - $2.7 M_{\odot}$ . My upper age estimate using 3D RWs is 2.5 Myr and I find two 3D WWs (*Gaia* DR2 3004263966389331456 and *Gaia* DR2 3217017610938439552) that have flight times since ejection that are larger than this age estimate. These two candidates are excluded from the results as they might have come from a different young star-forming region in the neighbourhood of the ONC and I am left with 24 3D WWs.

Within the low/intermediate-mass range  $> 0.3 M_{\odot}$ , the simulations produce an average of  $19.2 \pm 3.9$  WWs within 100 pc at 1 Myr, which increases to  $23.1 \pm 4.2$  at 3 Myr. By 4 Myr, I see a reduction in the WW numbers within 100 pc to  $21.4 \pm 4.5$ , due to stars travelling past this boundary. The maximum number of WWs within 100 pc from a single simulation is 27 WWs at 1 Myr increasing to 32 WWs at 3 Myr then reducing to 30 WWs by 4 Myr. The number of 3D WWs I find in *Gaia* DR2 matches the average number of WWs  $> 0.3 M_{\odot}$  in my simulations at  $\sim 1.5$  Myr. It matches the maximum number found in a single simulation only at an age of  $\sim 0.3$  Myr; however, further identified 3D WWs will increase these age estimates.

Eleven of the identified 29 RW and 27 of the 57 WW 2D-candidates are missing RV information. From Table 5.3 and 5.4, I see that even 2D-candidates with RV are not all confirmed as full 3D-candidates. However, 2D WW candidates with RV measurements appear more often to be 3D WW stars than their RW counterparts. Due to the missing RV information, I am unable to draw any further conclusions from the list of

2D-candidates as RVs are required to make an unambiguous RW/WW identification. I have also seen that a high RV might change the RW/WW status of a star that is identified as still bound to the ONC only considering its proper motion, and high RVs can turn a 2D WW into a 3D RW star.

*Gaia* does not detect any stars in the IR and more than half of the members of the ONC are only detectable in this range (e.g. Hillenbrand 1997). Fig. 5.5 shows the location of all identified 3D RW and WWs in relation to the ONC, which is located at the origin in this figure. Several of these are still within the close vicinity (a few pc) of the ONC, I expect to find further candidates at IR wavelengths.

Regardless of these limitations, I show with this analysis that the ONC has produced RW and WW stars across the full stellar mass range. In addition to the two known OB-RW stars, I find many more low/intermediate-mass RW and WW stars. This is consistent with the predictions made in my simulations in Chapter 4.

A clear 3D-identification is affected by missing or lower-quality radial velocities for many of my 2D-candidates and also by uncertainties in their distances. Furthermore, my analysis is influenced by the uncertainties about the radial extent and distance to the centre of the ONC. My search region projected on the sky has a diameter of 5 pc, based on the location of existing members. In contrast in the radial direction my search region has a size of 30 pc (15 pc in either direction of the adopted ONC distance of  $\sim 400$  pc). When constructing the CAMD, I correct for extinction and reddening; however, only a subset of my data has individual  $A_G$  and  $E(G_{BP} - G_{RP})$ . This results in me having to estimate values for the remaining stars by averaging over neighbouring stars, leading to highly uncertain age estimates as shown in Table 5.3 and 5.4. Even where stars have *Gaia* DR2 measured extinction and reddening values, some have very large errors in these quantities, which can lead to up to a magnitude of error on the CAMD. The consequence of these errors are large age ranges for my candidate stars, in particular upper age ranges. Andrae et al. (2018) also noted that the *Gaia* DR2 extinction and reddening itself are inaccurate on a star-by-star level. While the age estimates from literature (Da Rio et al. 2010, 2012, 2016) are not always consistent with my isochronal age estimates for the confirmed 3D-stars, they still confirm that most of the candidates are younger than the upper age of the ONC.

The CAMDs in Fig. 5.2 and 5.3 reveal a large number of 2D trace-backs that are located along the main-sequence and are therefore much older than the ONC's upper age limit. This highlights further that a trace-back on the sky is no indication of the origin of a star without an age estimate. More surprising is the trace-back of several

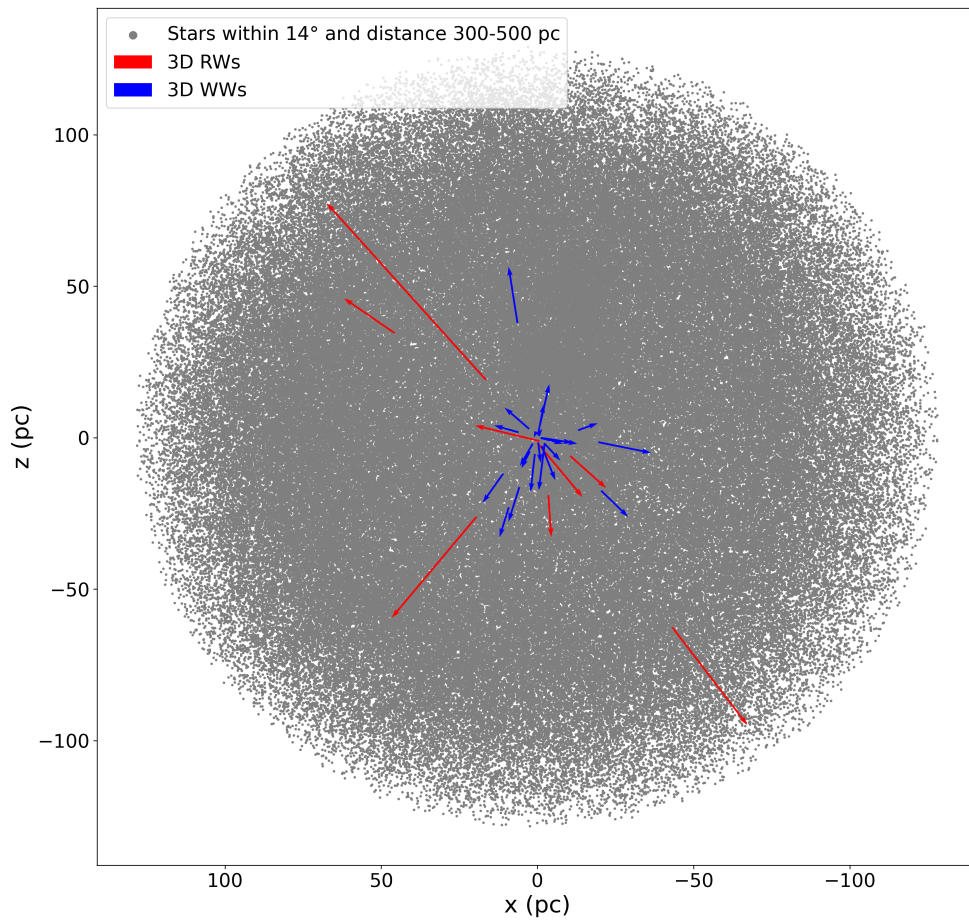


Figure 5.5: Location of the identified 3D RW and WW stars. The ONC is located in the centre of the plot and extends to a radius of 2.5 pc. I invert the x-axis to replicate the orientation on the sky (i.e. decreasing right ascension from left to right). The eight RW stars are plotted in “red”, with the length of the arrows indicating their 2D-velocity in the ONC rest frame, the 24 WW stars are plotted in “blue”.

older stars (on the main-sequence with ages of  $\sim 5\text{-}40$  Myr) in 3D, which qualifies these stars as having “visited” the ONC in the past.

It is possible that some of these stars might have even originated in the ONC. [Palla et al. \(2007\)](#) found lithium depletion in a small group of low-mass stars within the ONC. The authors suggest that this is an indication of these stars being older ( $>30$  Myr) than the rest of the stars in the ONC. However, [Sergison et al. \(2013\)](#) suggested that these differences in lithium are not necessarily evidence of an older population.

## 5.6 Conclusions

In this chapter, I combine *Gaia* DR2 observations with predictions from  $N$ -body simulations to search for RW and WW stars from the ONC within a distance of 100 pc in an attempt to constrain the region's initial conditions. The conclusions from my simulations and the search in *Gaia* DR2 are summarised as follows:

- (i) I find a number of 3D RW ( $>30 \text{ km s}^{-1}$ ) and 3D WW ( $10\text{-}30 \text{ km s}^{-1}$ ) stars originating from the ONC using *Gaia* DR2 astrometry and photometry in the low/intermediate-mass range ( $<8 M_{\odot}$ ). However, I find no high-mass stars ( $>8 M_{\odot}$ ) in either of the velocity ranges in all three dimensions. About 40% of my 2D-candidates are missing RVs and cannot be confirmed in 3D until this information is available.
- (ii) I trace back nine RWs to the ONC in *Gaia* DR2 that are still within my 100 pc search boundary. However, one of these stars appears to have left the ONC  $\sim 3.6$  Myr ago and my age estimates suggest that the ONC has an upper age below this age, reducing the number to eight RWs. My  $N$ -body simulations suggest that the older a star-forming region is, the fewer RWs are still found within this boundary. The number of RWs I find in my simulations agrees with those in *Gaia* DR2 when my simulated regions have an age of  $\sim 1.6\text{-}2.5$  Myr (based on eight RWs). This age estimate for the ONC is in agreement with others from literature (e.g. [Da Rio et al. 2010](#); [Reggiani et al. 2011](#)); however, the age estimate from my  $N$ -body simulations are likely an underestimate.
- (iii) My simulations predict that all WWs are still to be found within the search region until at least 3 Myr and that the number of WWs increases up to this age. In *Gaia* DR2, I find 26 WWs all with masses between  $0.3$  and  $2.7 M_{\odot}$ . Twenty-four of those have been ejected within the past 2.5 Myr (upper age implied from my RW findings). This agrees with the average number of WWs at an age of  $\sim 1.5$  Myr from my simulations but is below the maximum from a single simulation at virtually any age. However, future *Gaia* data releases and complementary IR surveys may enable me to identify further WWs, which will increase these age estimates.
- (iv) My analysis shows that ejected stars might be useful in constraining the initial conditions of star-forming regions. However, to use this method to its full extent

requires further improvements in *Gaia* or other observations, e.g. more measurements of radial velocities in addition to proper motion; extinction and reddening values for a larger number of stars.

- (v) The update of the method for converting from ICRS into Cartesian coordinates and velocities that I have applied in this chapter in comparison to the published version of [Schoettler et al. \(2020\)](#) did not lead to a major change in my conclusions. Instead of nine RWs, I now only identify eight RWs, which increases the age estimate by 0.1 Myr, but does not have any further effects on the conclusions. This highlights that the RW/WW candidates I have identified are likely high-confidence candidates.



## Chapter 6

**Protoplanetary systems that may experience more than one dense star-forming environment**

## 6.1 Introduction

In this chapter, which has been published in [Schoettler & Parker \(2021\)](#), I search for evidence of circumstellar discs around recently ejected stars from the ONC found in *Gaia* DR2 ([Gaia Collaboration et al. 2018a](#)), as well as around future and past visitors to determine whether any protoplanetary systems could experience more than one dense stellar environment. I briefly describe the ONC and its known population of stars with circumstellar discs. This is followed by a description of my data analysis method. I then present my results and conclude with a discussion/conclusion.

## 6.2 Circumstellar discs in the ONC

The ONC is a well-studied star-forming region at a distance of  $\sim 400$  pc (e.g. [Großschedl et al. 2018](#); [Kuhn et al. 2019](#)). It is still very young with an estimated mean age of 2-3 Myr (e.g. [Da Rio et al. 2010](#); [Reggiani et al. 2011](#)). Observations at different wavelengths have shown that this cluster has a population of  $\sim 3500$  stars (e.g. [Hillenbrand 1997](#); [Hillenbrand & Hartmann 1998](#); [Da Rio et al. 2012](#)). Its current average volume stellar density is approximately  $4 \times 10^2 \text{ M}_{\odot} \text{ pc}^{-3}$ , while its initial average volume stellar density is thought to have been much higher at  $10^3\text{-}10^4 \text{ M}_{\odot} \text{ pc}^{-3}$  (e.g. [Marks & Kroupa 2012](#); [Parker 2014](#)). This implies that the number of dynamical encounters may have been higher in the early stages of planet formation. While higher extinction might have protected discs against radiation at these times, simulations have shown that massive stars can quickly clear out the large cavities in their immediate surroundings reducing this shielding effect (e.g. [Dale et al. 2014](#)). As a result, the radiation fields experienced by discs are likely to have been much stronger than at present.

Several authors have searched for circumstellar discs around young stars in the ONC. [Hillenbrand et al. \(1998\)](#) used near-infrared (nIR) photometry combined with optical photometry and spectroscopy. They found evidence for circumstellar discs in 55-90 % of their sample of young stars within the mass range of  $0.1\text{-}50 \text{ M}_{\odot}$ . [Rebull et al. \(2000\)](#) used UV-excess emission in dereddened photometry to investigate the evidence for circumstellar accretion discs in the flanking fields of the ONC. They found that at least 40 % of the stars in their sample have a disc. [Sicilia-Aguilar et al. \(2005\)](#) used  $\text{H}\alpha$  profiles to find 15 new accreting member stars in the ONC. [Rebull et al. \(2006\)](#) then used mid-infrared observations to study a correlation of stars with circumstellar accretion discs with their rotation period and found a clear correlation. [Megeath et al. \(2012\)](#) classified YSOs with/without a disc via mid-IR observations. Most recently, [Großschedl](#)

et al. (2019) identified almost 300 young stars with discs and refined existing catalogues with new measurements. Fűrész et al. (2008) measured  $H\alpha$  to identify accreting stars and briefly mentioned high-velocity stars escaping the ONC but found no clear disc candidates among these ejected stars.

The Becklin-Neugebauer (BN) object (Becklin & Neugebauer 1967) is a fast moving, high-mass star. It is thought to have been recently ejected from the Orion region after a dynamical interaction with other ejected stars, known as Src I and Src x (e.g. Tan 2004; Bally & Zinnecker 2005; Rodríguez et al. 2005; Farias & Tan 2018). Src I is of specific interest as it appears to have retained part of its disc throughout the dynamical interaction and ejection process and might even be an ejected binary system moving with a proper motion of  $\sim 10 \text{ km s}^{-1}$  (e.g. Goddi et al. 2011; Moeckel & Goddi 2012; Bally et al. 2020).

Olczak et al. (2008) investigated the effect of dynamical encounters and the impact on the circumstellar discs around higher velocity stars escaping the ONC. Their result suggested that the location of the dynamical encounters (cluster centre or outer region) and the resulting velocity of the escapers can affect the amount of disc material that remains after ejection. More recently, McBride & Kounkel (2019) searched for high proper motion stars in the close vicinity of the Orion Nebula Cluster and found that seven out of their 26 candidates (including visitors and stars tracing back to other dense groups in the region) are clearly disc-bearing. The authors concluded that higher velocity stars are slightly more likely to be disc-less after ejection; however, they did not consider the difference to be significant.

While it does not appear to be a common occurrence, these examples show that circumstellar discs can survive the ejection process and that I should in principle be able to find RW stars with intact discs.

### 6.3 Method

I use the 3D RW and slower WW results of Schoettler et al. (2020) as a basis to investigate if any of these stars get ejected from the ONC with an intact circumstellar disc. In addition, I also search for disc-hosting stars in the list of past visitors to the ONC. I find secondary RVs for several 2D past visitors in Kounkel et al. (2018) and Cottaar et al. (2015), where no *Gaia* DR2 RVs are available. I then use this data and repeat the trace-back process described in Schoettler et al. (2020) to find additional past visitors to the ONC, i.e. older stars that trace back to the ONC but have not been born there.

Following [McBride & Kounkel \(2019\)](#), I also search for future visitors to the ONC and use the search approach described in [Schoettler et al. \(2020\)](#) but trace the stars' motion forwards instead of backwards in time. Given an estimated age of the ONC of  $\sim 2.5$  Myr (e.g. [Da Rio et al. 2010](#); [Reggiani et al. 2011](#); [Schoettler et al. 2020](#)), I search for any visitors that will travel through the ONC in the next 7.5 Myr and are currently already within 100 pc of it. This future time limit is driven by the knowledge that most young star-forming regions do not live past an age of 10 Myr ([Lada & Lada 2003](#)). I also search for secondary RVs to complement those provided in *Gaia* DR2.

Using these three lists of candidates, I search through [Hillenbrand et al. \(1998\)](#), [Rebull et al. \(2000, 2006\)](#), [Sicilia-Aguilar et al. \(2005\)](#), [Fűrész et al. \(2008\)](#), [Megeath et al. \(2012\)](#) and [Großschedl et al. \(2019\)](#) to check for the presence of a circumstellar disc around any of my past/future visitors or ejected ONC stars.

## 6.4 Results

Most of my fast ejected stars do not appear in any catalogue when searching for disc signatures. Nevertheless, I have identified several candidates with potential disc signatures in all three groups that might warrant further observations. I have double-checked the candidates identified in the published version of this chapter ([Schoettler & Parker 2021](#)) against the updated search results from the ONC presented in Chapter 5, and the candidates have remained unchanged by the update in the coordinate conversion procedure. In Table 6.1, I present my ejected candidates with a disc.

*Gaia* DR2 3017265515291765760 (BD-05 1307) is the brightest 3D RW star identified in [Schoettler et al. \(2020\)](#). It is a very young star ( $\sim 0.3$  Myr, [Hillenbrand 1997](#)) and is still located within the central ONC region but will leave this region due to its high space velocity. I find this star in [Hillenbrand et al. \(1998\)](#), who used excess emission in the nIR  $\Delta(I_C-K)$  to identify possible circumstellar discs around young stars. This disc indicator measures hot dust at a wavelength of  $\sim 2 \mu\text{m}$ . The authors quoted two limiting values for the IR-excess above which a disc could be present. The first, more conservative value is  $\Delta(I_C-K) = 0.30$  mag. This RW-star has an IR-excess of  $\Delta(I_C-K) = 0.21$  mag, which falls below this higher IR-excess value. However, it satisfies the lower limit of  $\Delta(I_C-K) = 0.10$  mag, so a disc could be present. The IR-excess for this star is weaker than the mean and median values for stars in its local environment, suggesting it might not have originated there.

*Gaia* DR2 3209424108758593536 (Brun 259) is a 3D WW star identified in [Schoettler](#)

Table 6.1: Stars around the ONC possibly with a circumstellar disc. Column 2+3: velocity in ONC rest frame [rf] from [Schoettler et al. \(2020\)](#); Column 4: status identifier - ONC ejected, future or past visitor; Column 5-7: from literature sources - <sup>1</sup>[Van Altena et al. \(1988\)](#), <sup>2</sup>[Hillenbrand \(1997\)](#), <sup>3</sup>[Da Rio et al. \(2010\)](#), <sup>4</sup>[Da Rio et al. \(2016\)](#); Column 8: Disc information from literature sources: [5] [Hillenbrand et al. \(1998\)](#), [6] [Rebull et al. \(2000\)](#), [7] [Rebull et al. \(2006\)](#), [8] [Fűrész et al. \(2008\)](#), [9] [Sicilia-Aguilar et al. \(2005\)](#).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Status identifier	Age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type	Source
3017265515291765760	30.1	12.3 <sup>a</sup>	ONC ejected	0.3 <sup>2</sup>	2.5 <sup>2</sup>	K1 <sup>1</sup>	[5]
3209424108758593536	14.1	-4.4 <sup>b</sup>	ONC ejected	0.5-2.5 <sup>3,4</sup>	0.7 <sup>3,4</sup>	K7 <sup>2</sup>	[5, 9]
3017367151399567872	3.6	12.2 <sup>a</sup>	ONC ejected	1.7 <sup>4</sup>	2.7 <sup>4</sup>	-	[5, 7]
3209637203559481728	16.3	10.1 <sup>c</sup>	Future visitor	2.5 <sup>4</sup>	0.2 <sup>4</sup>	-	[6, 8]
3017376325447976576	26.5	-63.3 <sup>c</sup>	Past visitor	18 <sup>4</sup>	1.4 <sup>4</sup>	-	[6]

<sup>a</sup>*Gaia* DR2, <sup>b</sup>[Cottaar et al. \(2015\)](#), <sup>c</sup>[Kounkel et al. \(2018\)](#)

[et al. \(2020\)](#) and also [McBride & Kounkel \(2019\)](#). It has an age of 0.5-2.5 Myr ([Da Rio et al. 2010, 2016](#)) and is also still located within the ONC. It appears in [Hillenbrand et al. \(1998\)](#) with a value  $\Delta(I_C-K) = 0.19$  mag, fulfilling their lower limit for the presence of a disc. The IR-excess for this star is weaker than the mean and median values for stars in its local environment, suggesting it might not have originated there. However, [Sicilia-Aguilar et al. \(2005\)](#) stated that it is a WTTS (narrow H $\alpha$  emission), which are not expected to have much, or any circumstellar material left.

*Gaia* DR2 3017367151399567872 (Par 1799) is another 3D WW candidate from [Schoettler et al. \(2020\)](#); however, its high-velocity status is driven by the RV. According to [Rebull et al. \(2006\)](#), it has indicators of a disc. These authors use a mid-IR colour index  $[3.6 \mu\text{m}] - [8 \mu\text{m}] > 1$  mag to infer a disc, and this star has a value just larger than 1 mag ([Rebull et al. 2006](#)). This star also appears in [Hillenbrand et al. \(1998\)](#) but has a negative  $\Delta(I_C-K)$ , which is due to measurement errors and photometric variability.

I have also searched for disc signatures in the identified 2D-RW/WW candidates in [Schoettler et al. \(2020\)](#) and find four stars with measurements from observations indicating the presence of discs. These stars are *Gaia* DR2 3209497088842680704 and *Gaia* DR2 3209498394512739968 with a small, negative UV-excess as shown in [Rebull et al. \(2000\)](#), *Gaia* DR2 3014834946056441984 with a flat disc identified in [Juhász et al. \(2010\)](#) and *Gaia* DR2 3015714967674577024 with a transitional disc identified in [Kim et al. \(2013\)](#). These four candidates could add to the number of ejected stars with discs; however, they are missing an RV measurement and I cannot confirm their origin in the ONC yet.

Searching through my list of future visitors to the ONC, I do not find any that have appeared in any surveys/papers searching for circumstellar discs and most of these

candidates do not appear in any literature sources at all.

McBride & Kounkel (2019) identified several 2D-visitors to the ONC in their paper. *Gaia* DR2 3209637203559481728 (2MASS J05350504-0432334) is one such future visitor. It does not appear in my search for future visitors as I apply an astrometric quality indicator, the renormalized unit weight error (RUWE)  $< 1.3$ . This star has a much higher RUWE value indicating issues with its astrometry. It is in fact a spectroscopic binary (e.g. Tobin et al. 2009; Kounkel et al. 2016). The proper motion errors are just  $\sim 10\%$ , which is still acceptable. I can trace this star forward to the ONC in 3D using the RV from Kounkel et al. (2018) but only when considering its large distance errors ( $377^{+57}_{-44}$  pc). It appears in Rebull et al. (2000) with an UV-excess clearly indicative of a disc, i.e. smaller than  $-0.5$  mag, even though it is classed as a non-accreting WTTS in Fűrész et al. (2008).

There is another 2D-visitor *Gaia* DR2 3017199755050720384 (V1589 Ori) identified in McBride & Kounkel (2019) with a very clear disc identification (Megeath et al. 2012; Großschedl et al. 2019). Its position on the sky (within the ONC boundaries) and proper motion (moving away from the ONC) on the sky suggests a strong connection to the ONC. However, based on its estimated distance of  $\sim 250$ - $350$  pc (Bailer-Jones et al. 2018) and its RV (Kounkel et al. 2018) pointing in direction of the ONC, the star does not trace to the ONC in 3D. Based on its kinematics and position, I suggest it is not associated with the ONC (or even Orion A) at all, despite often being linked with it in literature (e.g. Fűrész et al. 2008; Da Rio et al. 2016; Kounkel et al. 2018).

I have not found any stars with a disc in the past visitors list in Schoettler et al. (2020). However, amongst the past visitors that I trace back using secondary RVs for this project, I find one possible disc candidate.

*Gaia* DR2 3017376325447976576 (Brun 944) appeared in Rebull et al. (2000) and showed a clear UV-excess ( $\sim -0.25$  mag), which is not small enough to be directly identified as a disc candidate in that paper using their  $-0.5$  mag upper limit. I use RV from Kounkel et al. (2018) to trace back this star. It is still located inside the ONC at a distance of  $\sim 390$  pc (Bailer-Jones et al. 2018) but has passed through the central, denser part already. With an estimated age of  $\sim 18$  Myr (Da Rio et al. 2016) it cannot have been born in the ONC.

## 6.5 Discussion and Conclusion

I find three recently ejected young stars from the ONC (one RW and two WW stars) with some evidence of a disc, based on the IR-excess observed in their emission (Hillenbrand et al. 1998; Rebull et al. 2006). However, not all of these stars satisfy the more conservative excess limits stated in the above papers for a clear disc identification.

Hillenbrand et al. (1998) stated two IR-excess limits for the identification of discs and the authors highlighted in their paper, that the stricter limit of  $\Delta(I_C-K) = 0.30$  mag might be too conservative and cause discarding of actual disc candidates. Sicilia-Aguilar et al. (2005) stated that imposing an even higher IR-excess  $\Delta(I - K) > 0.5$  mag (originally suggested by Rhode et al. 2001) to detect disc-bearing CTTSs can be considered a safe approach. However, they also mentioned that this can lead to missing out on a significant fraction of young stars with discs that have an IR-excess value between 0-0.5 mag and found more than one-third of the stars in this IR-excess range are CTTSs in their study (Sicilia-Aguilar et al. 2005). The  $\Delta(I_C-K)$  values for the two stars in Table 6.1 are on average lower than those in their local environment. This could be additional evidence that they do not originate there. However, the dynamical evolution means that the positions of individual stars are highly likely to be transient with respect to each other.

The only WW star (*Gaia* DR2 3017367151399567872) that is identified as a disc candidate within the boundaries given in Rebull et al. (2006) is not a clear-cut ejected star. It is still located within the central ONC region at a distance of  $\sim 386$  pc (Bailer-Jones et al. 2018) and has a 2D-velocity in the ONC reference frame that is below the escape velocity calculated by Kim et al. (2019). Its RV pushes this star into the ejected star category, which could alternatively be explained as having a binary origin. Its RV has been measured by *Gaia* DR2 (which I use) and Cottaar et al. (2015) and the values are consistent with each other, therefore not supporting a binary identification. This star has also featured in many studies of the ONC stellar population over the years, none of which identify a binary companion. Köhler et al. (2006) included this star as one of their targets in their search for binaries but do not find a companion.

Just as higher RVs can be due to binary motion, proper motion can also be affected by the binary motion. The orbital motion of binaries can lead to a photocentre wobble in observations. This centroid displacement can be identified by higher RUWE values for shorter period binaries, i.e. less than the observational baseline of the survey (22 months for *Gaia* DR2). For longer period binaries (several to  $\sim 10$  yr), this can instead lead to excess proper motion and a lower RUWE (e.g. Belokurov et al. 2020; Penoyre et al. 2020).

Four of the five identified star-disc candidates have a low RUWE value ( $< 1.3$ ), and it is theoretically possible that the high velocities measured are at least partially due to an unknown binary companion. However, even if any of these stars were in an equal-mass binary, the average separation between the stars would not exceed 10 au (using a 10 yr period). According to estimates by [Belokurov et al. \(2020\)](#), binaries with a semi-major axis up to 10 au should be detectable by RUWE up to a distance of 2 kpc, so these binaries would show up with a higher RUWE in addition to excess proper motion.

One of the five stars (*Gaia* DR2 3209637203559481728) is a known spectroscopic binary with a high RUWE. This suggests a shorter period binary, where the measured higher proper motion is unlikely due to the binary status. The binary separation is likely to be on a similar scale to that of the disc diagnostics used ( $\sim 1$  au). A small separation between stars in a binary can affect any circumstellar discs present around the stars (e.g. [Jensen & Akeson 2014](#); [Benisty et al. 2018](#))

While I have not found many ejected stars that show excess emission indicative of a disc, my findings suggest that stars ejected from their birth regions due to dynamical interactions might retain some circumstellar material. I find further ejected candidates that feature in papers searching for discs, but that do not show excess emission. These stars are shown in Table A.5 in Appendix A for information. While these young stars (mostly WTTSs) show no indication of accretion, they could still feature harder to find debris discs or even young planetary systems.

Most of the ejected stars with or without a disc are unlikely to encounter a second, dense star-forming region during their lifetime. However, the location of the ONC within the Orion A molecular cloud provides several opportunities for a second encounter with such a region. In this chapter, I do not trace forward the trajectories of my ejected star-disc systems that originate in the ONC. However, I have searched for future visitors approaching the ONC from other regions, but I have found no candidates using my more conservative search requirements.

*Gaia* DR2 3209637203559481728 is a future visitor identified in [McBride & Kounkel \(2019\)](#), for which I find disc indicators in literature. It fulfils the strict UV-excess limit of [Rebull et al. \(2000\)](#); however, was classed as a WTTS by [Fűrész et al. \(2008\)](#). This young star ( $\sim 2.5$  Myr, [Da Rio et al. 2016](#)) is on approach to the ONC with a velocity of  $19 \text{ km s}^{-1}$  (ONC reference frame) appearing to have been ejected from its birth region with a partial disc. Which part of the ONC it will encounter when passing through it in the future depends strongly on its current location. While its current position on the sky is fairly well constrained, its distance has a large margin of error ( $377^{+57}_{-44}$  pc). It



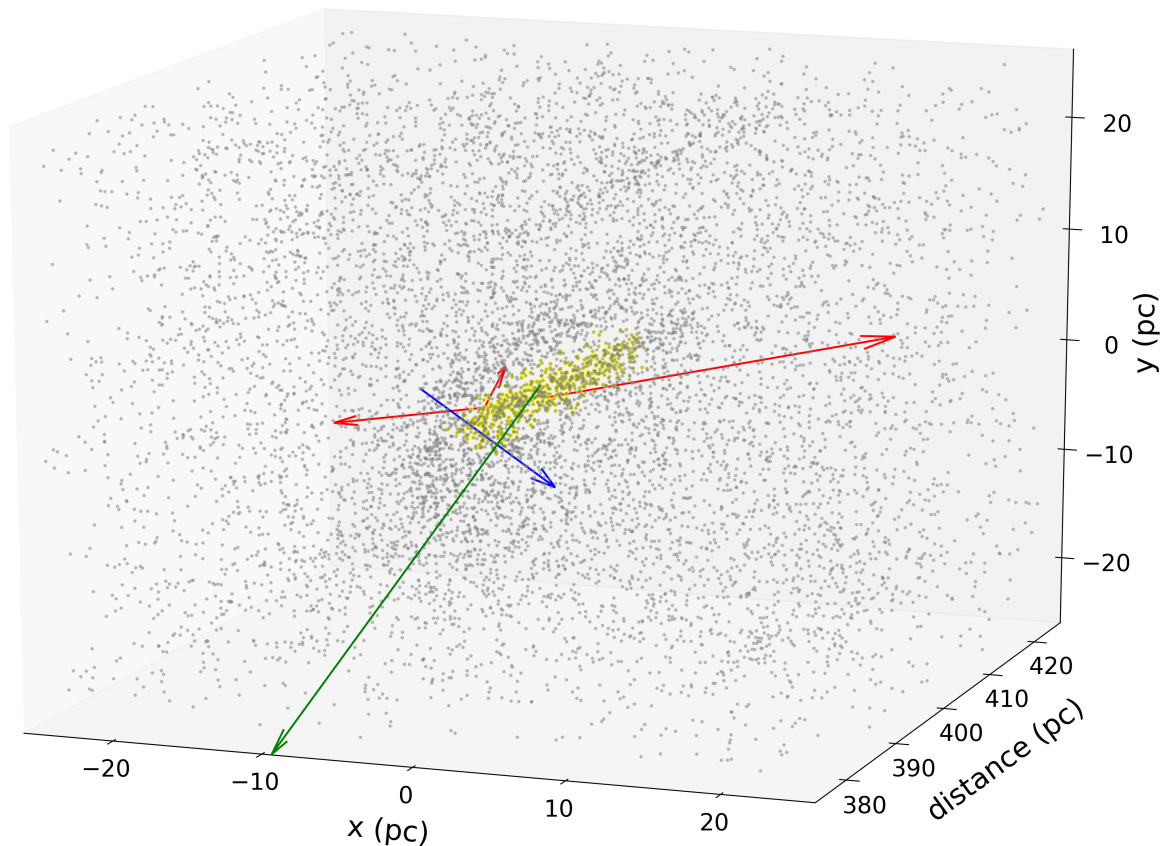


Figure 6.1: Location and motion of the five identified stars with possible discs. The ONC is located in the centre of the plot extending to a radius of 2.5 pc in  $xy$ -direction (plane of the sky) and 15 pc in  $z$ -direction (stars within this boundary in “yellow”) (Schoettler et al. 2020). I plot fast ejected stars born in the ONC in “red”, fast future visitors in “blue” and older visitors passing through the ONC in “green”. The tail of the arrows indicates the current position of the stars, with the length of the arrows representing the scaled velocity.

might miss the ONC completely (if the distance  $< \sim 380$  pc or  $> \sim 415$  pc), encounter only the more sparsely populated outskirts, or it may encounter the central, dense parts of the ONC. Depending on its trajectory, it might retain all of its existing circumstellar material or lose it all.

Finally, I find an older visitor to the ONC, that based on its estimated age ( $\sim 18$  Myr, Da Rio et al. 2016) does not originate in the ONC. While it is still located just within the ONC’s boundary, it has already passed through the densest parts on its trajectory. It is the fastest of all stars identified in my search and still retains a small UV-excess, but larger than the upper limit of Rebull et al. (2000). This type of excess emission is often used as an indicator of an accretion disc but can also point to magnetic activity in

a WTTS without a disc (e.g. Venuti et al. 2015). Rebull et al. (2000) stated that their upper limit for UV-excess of -0.5 mag used to distinguish disc and non-disc candidates is likely a conservative approach that might exclude stars with discs.

If this UV-excess does in fact indicate the presence of a disc, then it has survived several factors that can often destroy a disc completely. Its higher peculiar velocity points to a fairly dynamical encounter leading to the ejection from its birth environment. It has then encountered a second dense region (the ONC), where it once again was subject to dynamical interactions and possibly photoevaporating radiation. Finally, its advanced age makes it less likely to still have an accretion disc. Unfortunately, this star does not appear in any other disc searches, so no final verdict on the presence/absence of a disc can be made.

While I have found a small number of possibly disc-bearing, ejected RW/WW stars, future and past visitors in my star-disc searches in literature, there is a key aspect hindering my search. Circumstellar discs are found predominantly around young stars, which are usually located in star-forming regions. As a consequence, observations to search for discs focus on these regions, omitting areas further away. Fig. 6.1 indicates the positions and motions of the five identified high-velocity stars around the ONC (zoomed in to the central  $50 \times 50$  pc). The figure illustrates the likely observational bias as all identified disc candidates (stars just ejected from the ONC, future visitors and older stars not born in the ONC that have visited in the past) are located in close proximity to the ONC, even though the data used cover a much larger area.

In this chapter, I set out to investigate if circumstellar discs can survive the ejection from young star-forming regions using the ONC as an example. I find that there are stars at RW and WW velocities that have been ejected, which show some evidence of a disc. I also find a disc-bearing visitor from another star-forming region that is on approach to the ONC about to encounter a second higher density region. Finally, I find an older visitor that has just passed through the ONC and could possibly still have retained some of its circumstellar disc.

Whilst limited to a handful of stars, I have demonstrated that planet formation around these stars could have been hindered by external effects in more than one dense star-forming environment.

## Chapter 7

### Ejected stars from NGC 2264 in *Gaia* DR2 observations

## 7.1 Introduction

This chapter is based on my work that was published as [Schoettler et al. \(2022\)](#). In the discussion, I also refer to results from a companion study ([Parker & Schoettler 2022](#)) that I contributed to by compiling the mass census of the member stars in NGC 2264 from literature. For the analysis presented in this chapter, I use *Gaia* DR2 and also EDR3 observations to search for RW and WW stars around NGC 2264. I compare these numbers to  $N$ -body simulations with different initial conditions to constrain the initial conditions of NGC 2264. NGC 2264 is another region within 1 kpc to Earth showing on-going star formation, which should allow me to probe the ejected population down to sub-solar masses. The region is less centrally concentrated with obvious subclustering, and fewer studies have investigated the initial conditions of this star-forming region. The detailed study of the kinematics in this chapter allows me to more broadly test my approach using RW/WW stars to constrain the initial conditions.

## 7.2 NGC 2264 specific information used in the analysis

I am searching for RW and WW stars from the young star-forming region NGC 2264, which is located in the Monoceros OB1 association cloud complex. Recent estimates put this region at distances ranging from approximately 720 pc (719 pc: [Maíz Apellániz 2019](#); 723 pc: [Cantat-Gaudin et al. 2019](#)) up to 750 pc ([Kuhn et al. 2019](#)). All of these estimates used *Gaia* DR2 and are subject to considerable error margins ( $\pm 16$  pc to  $\pm 60$  pc). Another recent estimate by [Zhao et al. \(2018\)](#) used nIR extinction from 2MASS photometry to derive a much larger distance estimate of 1.2 kpc. Older estimates also encompassed a large range from 667 pc ([Dias et al. 2002, 2012](#)) and 760 pc ([Dahm 2008](#)) up to 913 pc ([Baxter et al. 2009](#)) with similarly high margins of error.

NGC 2264 is not a centrally concentrated cluster but spatially elongated along a NW-SE orientation with different subclusters spread along  $\sim 8$  pc (e.g. [Tobin et al. 2015a](#); [Buckner et al. 2020](#)). [Zwintz et al. \(2017\)](#) suggested a diameter of  $\sim 39$  arcmin, which translates to 8-14 pc depending on the distance used to convert from arcmin to parsec. The northern region is dominated by S Monocerotis (S Mon), an O7 spectroscopic binary ( $\sim 40$ -50  $M_{\odot}$ ; [Tokovinin 2018](#); [Maíz Apellániz 2019](#)). The southern region consists of two subclusters. These subclusters have differing designations in literature but are most often referred to as NGC 2264-C or IRS 1 and NGC 2264-D, IRS 2 or the Spokes cluster (e.g. [Teixeira et al. 2006](#)). IRS 1 is thought to centre around a B2-type star with a mass estimate of  $\sim 10 M_{\odot}$  (IRAS 06384+0932; [Allen 1972](#); [Thompson et al. 1998](#); [Peretto et al.](#)

2006). IRS 2 is thought to centre around a young Class 1 type source, which is possibly a B-type binary with a primary star of  $\sim 8 M_{\odot}$  (IRAS 06382+0939; [Castelaz & Grasdalen 1988](#); [Margulis et al. 1989](#); [Teixeira et al. 2006](#)).

[Lada & Lada \(2003\)](#) suggested that NGC 2264 appears to be a cluster containing at least two different areas of increased surface density in a hierarchical structure. [Hetem & Gregorio-Hetem \(2019\)](#) calculated a surface density for the whole region of NGC 2264 of  $\sim 7$  stars  $\text{pc}^{-2}$  and a radius for the region of  $\sim 3.8$  pc, which is roughly in line with the diameter values from other authors. Slightly higher surface density values were recorded by [Rapson et al. \(2014\)](#), who showed that most of the YSOs in NGC 2264 are found in regions with densities above  $\sim 10$  stars  $\text{pc}^{-2}$  with a peak between  $\sim 10$ – $25$  stars  $\text{pc}^{-2}$ . [Peretto et al. \(2006\)](#) suggested a stellar density for the southern region around IRS 1/2 of  $\sim 80$  stars  $\text{pc}^{-2}$ , based on a figure from [Lada et al. \(1993\)](#). [Mariñas et al. \(2013\)](#) calculated a value for the stellar density of  $\sim 30$  stars  $\text{pc}^{-2}$  for the southern region.

There is very little foreground extinction and the dark cloud located directly in the background reduces the contamination from non-member stars (e.g. [Herbig 1954](#); [Venuti et al. 2018](#)). While this is helpful when identifying member stars still located in NGC 2264, it makes it more difficult to find stars ejected in radial direction moving away behind the cluster.

The age of the cluster differs across the different regions of NGC 2264. The average age of the cluster is suggested to be 3–5 Myr with an age spread of 4–5 Myr, with the region around S Mon containing older stars. The two subclusters in the south are thought to be younger as a result of sequential star formation (e.g. [Dahm 2008](#); [Mayne & Naylor 2008](#); [Sung et al. 2009](#); [Naylor 2009](#); [Venuti et al. 2018](#)). A recent study by [Hetem & Gregorio-Hetem \(2019\)](#) used a sample of stars spread across the whole region of NGC 2264 and calculated an average age of  $\sim 2$  Myr. [Getman et al. \(2014\)](#) also calculated ages for stars in NGC 2264 with the mean value suggesting an age of  $\sim 2.6$  Myr. [Peretto et al. \(2007\)](#) calculated very young ages for the southern subclusters IRS 1 and IRS 2 of  $\sim 0.1$  Myr, whereas [Mariñas et al. \(2013\)](#) calculated a median age of  $\sim 1$  Myr for this region.

All of these age estimates suggest that star formation activity started first in the northern region where S Mon is now located over 5 Myr ago and in the southern region more recently  $\sim 1.5$ – $2$  Myr ago. Furthermore, in the Cone nebula located around IRS 1 there are also a number of embedded sources further supporting a young age estimate (e.g. [Venuti et al. 2018](#)).

I separately analyse the northern (S Mon) and the two southern (IRS 1 and IRS 2)

Table 7.1: NGC 2264 centre parameters used in the analysis from [Kuhn et al. \(2019\)](#)

	S Mon	IRS 1	IRS 2
Right ascension [RA] (ICRS) $\alpha_0$	6h 40m 50s	6h 41m 07s	6h 41m 01s
Declination [Dec] (ICRS) $\delta_0$	09° 51' 03"	09° 28' 09"	09° 35' 56"
Proper motion RA $\mu_{\alpha^*,0}$ (mas yr <sup>-1</sup> )	-1.62 ± 0.08	-2.05 ± 0.18	-2.29 ± 0.14
Proper motion Dec $\mu_{\delta_0}$ (mas yr <sup>-1</sup> )	-3.71 ± 0.07	-3.90 ± 0.09	-3.61 ± 0.08
RV (km s <sup>-1</sup> )	15.8 ± 2.9	15.8 ± 2.9	15.8 ± 2.9
Adopted distance (pc)	738 ± 23	736 ± 23	748 ± 24

subclusters of NGC 2264 due to the difference in age and search for RW and WW stars that have been ejected from all three. I use an upper age of 5 Myr for the search around S Mon and 2 Myr for the search around IRS 1 and IRS 2, thereby covering most of the available age estimates. I use the distances to these subclusters from [Kuhn et al. \(2019\)](#) for my analysis, as shown in Table 7.1. These distances are at the upper end of the distance estimates for NGC 2264 from *Gaia* DR2 measurements and have considerable uncertainties.

[Teixeira et al. \(2012\)](#) used IR-luminosity functions to estimate the size of the stellar population in NGC 2264 within their search fields that contain all three of the regions of interest for this analysis (S Mon, IRS 1 and IRS 2). Their results suggested that the whole cluster contains  $1436 \pm 242$  members. They also provided an estimate of the stellar mass of the region using a simple assumption that each star has a mass of  $0.5 M_{\odot}$ , which they adopted from [Muench et al. \(2007\)](#). This results in their stellar mass estimate of  $718 \pm 121 M_{\odot}$  ([Teixeira et al. 2012](#)).

This mass estimate means that NGC 2264 is considerably less massive than the ONC. This lower mass estimate reduces the required escape velocity from the cluster compared to the value for the ONC (see [Kim et al. 2019](#)), and therefore also affects my choice for the lower velocity limit for WW stars (compared to [Schoettler et al. 2020](#)). Given the much lower mass, I choose the original velocity of  $5 \text{ km s}^{-1}$  from [Eldridge et al. \(2011\)](#) as my lower velocity limit for WW stars and consider RWs to have velocities exceeding  $30 \text{ km s}^{-1}$ .

There has not yet been any comprehensive search for RW/WW stars from NGC 2264. [Maíz Apellániz et al. \(2020\)](#) found two potential RW 2D-candidates that appear to have been ejected from the northern region of the cluster: *Gaia* DR2 3326734332924414976 and *Gaia* DR2 3326951215889632128. I will seek to confirm these stars in my analysis.

In [Schoettler & Parker \(2021\)](#), I searched for circumstellar discs around RW/WW stars from the ONC, as well as around future and past visitors to determine the re-

silence of protoplanetary discs to multiple encounters within dense environments. For NGC 2264, I repeat this search in the Simbad/VizieR databases (Wenger et al. 2000; Ochsenbein et al. 2000) for studies that have covered this cluster. I search for circumstellar discs to increase my sample size of ejected star-disc systems that have or will encounter a second dense star-forming region.

As for the work in Chapter 5, I use the centre parameters for position and velocity as defined in Kuhn et al. (2019) with 100 pc as the outer boundary for my search region. Choosing an outer boundary for my search region, both in the observations and the simulations, allows me to compare similar volumes of space. 100 pc translates to  $\sim 8^\circ$  around each of the centres of the three subclusters evaluated here. Instead of parallax, I use the distances from the catalogue by Bailer-Jones et al. (2018) and reduce the sample by selecting sources that are within  $\pm 100$  pc of each of the centre distances, as shown in Table 7.1.

Jackson et al. (2020) suggested central velocity parameters for NGC 2264, which are close to the values of Kuhn et al. (2019) in proper motion ( $\sim -1.95$  mas yr $^{-1}$  and  $\sim -3.76$  mas yr $^{-1}$ , when converted from km s $^{-1}$  using the distance to S Mon in Table 7.1). However, the values differ more in the radial velocity (RV), where the higher Jackson et al. (2020) value of  $\sim 20.3$  km s $^{-1}$  is closer to the higher radial velocities quoted in other studies (e.g. Kharchenko et al. 2005; Fűrész et al. 2006). I have run my analysis with this higher RV, as well as the one in Kuhn et al. (2019) and find only minor differences in the results. I use the RV from Kuhn et al. (2019) throughout this analysis to be consistent with the choice of the other velocity parameters for the centre regions.

The *Gaia* analysis steps are described in Chapter 3, and I filter my data for astrometric and photometric quality. For defining the RUWE cut-off value for the astrometric quality, I plot my data set as described in the technical note and as shown in Schoettler et al. (2020). I find that an upper RUWE limit of 1.3 to be the best fitting value, which is the same as I used in the ONC analysis.

In my search for RW (velocity  $> 30$  km s $^{-1}$ ) and WW (velocity: 5–30 km s $^{-1}$ ) candidates, I first trace back their positions in the xy-plane for up to 2 and 5 Myr depending on the subcluster using the converted proper motions. I define a cluster boundary based on existing member lists and once the backwards path of a star crosses this boundary, a star becomes a 2D-candidate. I use the time at which this path intersects the boundary as the minimum time since ejection, i.e. the flight time.

As I am interested in ejections from the subclusters, I define a radius of 2 pc in the xy-direction as my subcluster radius. When considering the whole star-forming region,

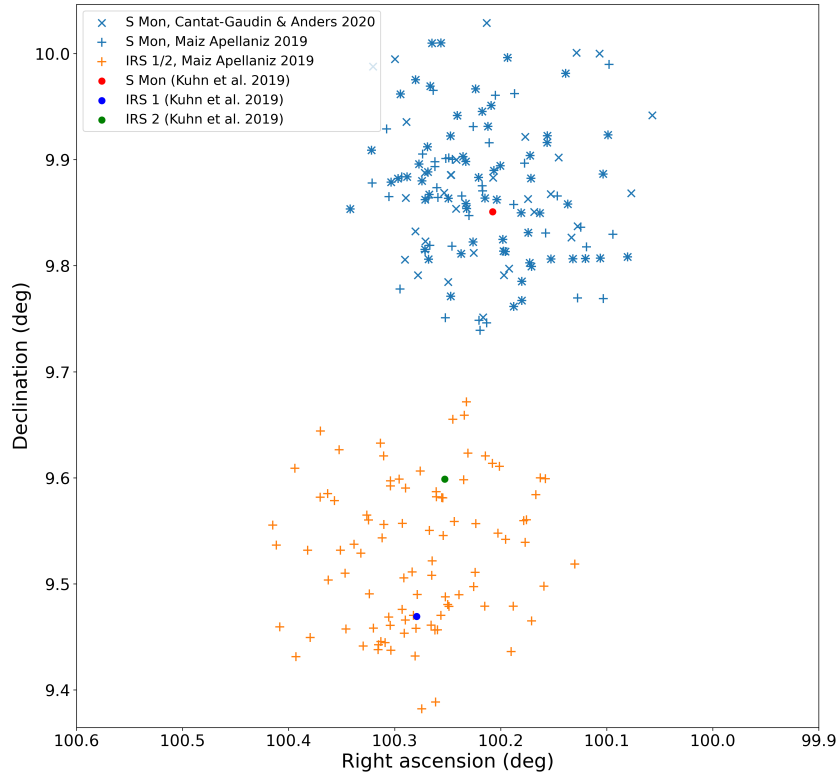


Figure 7.1: *Gaia* DR2 right ascension and declination of NGC 2264 members identified in [Cantat-Gaudin & Anders \(2020\)](#), members with 100% probability, only around S Mon, blue “x”) and [Maíz Apellániz \(2019\)](#), North cluster: blue “+”, South cluster: orange “+”). The centres of S Mon, IRS 1 and IRS 2 are plotted based on their positions in [Kuhn et al. \(2019\)](#).

this will amount to it extending across  $\sim 2$  pc in the x-direction (right ascension), while extending to over  $\sim 8$  pc in the y-direction (declination), due to the NW-SE elongation of the whole cluster. This 2 pc radius corresponds to an angular size of  $\sim 0.16^\circ$  around the subcluster centre positions, which are located at the origin in the rest-frame for each converted data set.

Fig. 7.1 illustrates this choice, where I plot cluster members identified in [Cantat-Gaudin & Anders \(2020\)](#), members with 100% probability, only around S Mon, no members identified around IRS 1/2) and in [Maíz Apellániz \(2019\)](#), North and South cluster) using their *Gaia* DR2 measurements for the right ascension and declination. These regions are each well enclosed by a circle with a radius of  $\sim 0.16^\circ$ .

To constrain the size of NGC 2264 in the radial direction (distance), one of the issues is the uncertainties in the distance of the subcluster centres ([Kuhn et al. 2019](#)), another is the lack of information about their radial extent. The available membership lists



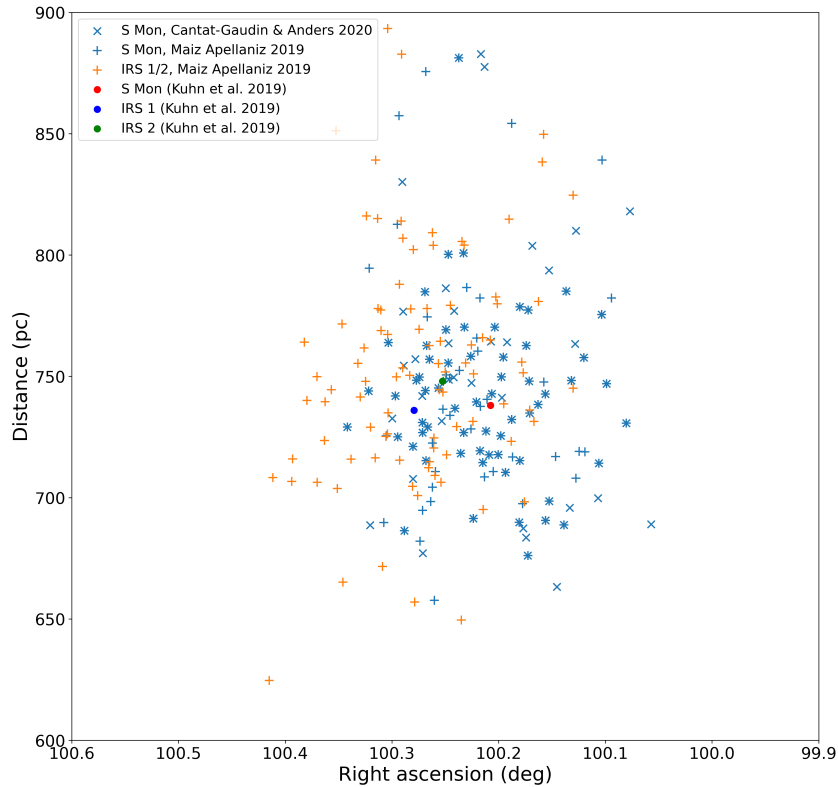


Figure 7.2: *Gaia* DR2 right ascension and distance (inverted parallax) of NGC 2264 members identified in Cantat-Gaudin & Anders (2020, members with 100% probability, only around S Mon, blue “x”) and Maíz Apellániz (2019, North cluster: blue “+”, South cluster: orange “+”). The centres of S Mon, IRS 1 and IRS 2 are plotted based on positions in Kuhn et al. (2019).

could be helpful in constraining the radial extent of NGC 2264. In Fig. 7.2, I again plot the members from Maíz Apellániz (2019) and Cantat-Gaudin & Anders (2020) using their *Gaia* DR2 right ascension and distance (inverted parallax). There appears to be a higher stellar surface density between 700–800 pc. This could suggest a radial extent of NGC 2264 of several tens of pc with a centre distance of 740–750 pc but could also be driven by the parallax errors.

As Fig. 7.2 does not allow me to put better constraints on the size in radial direction, I use the subcluster centre distance estimates in Kuhn et al. (2019, 736–748 pc), which show upper uncertainties of 23–24 pc in the subcluster centres. I add these uncertainties to the assumed 2 pc on-the-sky cluster radius turning my chosen search radius in the radial direction to 25–26 pc depending on subcluster. However, by using this geometry for my analysis, I do not suggest that this value represents the actual radial extent of NGC 2264. The circular cluster boundary used in the xy-plane turns to an elliptical

cluster boundary in the  $xz$ - and  $yz$ -plane with a semi-minor axis of 2 pc (in the  $x$  and  $y$  axes) and a semi-major axis of 25–26 pc (in the  $z$  axis) depending on the subcluster.

Only  $\sim 2\%$  of the stars in my NGC 2264 data set have RVs in *Gaia* DR2, which makes the search for ejected stars using 3D kinematics difficult. For my 2D-candidates without RVs, I have searched through the Simbad/VizieR databases (Wenger et al. 2000; Ochsenbein et al. 2000) and find several additional RVs in Jackson et al. (2016), Kounkel et al. (2019) and individual measurements in other sources. I also cross-match positions with Lamost DR5 data (Luo et al. 2019) via VizieR (using a target match radius of 2 arcsec) to complete my data set with RV measurements from secondary literature sources.

For all my 2D-candidates, I search for mass estimates, ages and spectral types from literature sources. I find mass estimates for most of my candidates in the StarHorse database (Anders et al. 2019), where I choose the 50th percentile and in the TESS Input Catalogue (v8.0, Stassun et al. 2019). I also find additional masses for individual stars in other sources. For any stars with more than one mass estimate, I quote the whole range of the values in the data tables.

I use an upper age limit of 5 Myr for S Mon and 2 Myr for IRS1 and IRS2 and only stars that are younger than these ages are considered as ejected stars from these subclusters. I use PARSEC isochrones (version 1.2S, Bressan et al. 2012) to separate the stars into two age brackets with younger stars either fully located above the isochrones or when located below the isochrone with error bars crossing the age boundary.

Not much is known about the metallicity of NGC 2264. Baratella et al. (2020) gave an abundance ratio  $[\text{Fe}/\text{H}] = 0.11$  but used only one star to determine this value. The average value given in Spina et al. (2017) was  $[\text{Fe}/\text{H}] = -0.06$ . King et al. (2000) suggested  $[\text{Fe}/\text{H}] = -0.15$  but calculated this value from only three stars. Heiter et al. (2014) re-evaluated the observational data of King et al. (2000) and applied restrictions on certain parameters to get an abundance  $[\text{Fe}/\text{H}] = -0.13$ . This value was also found by Netopil et al. (2016). I use the abundance ratio  $[\text{Fe}/\text{H}] = -0.13$  for my isochrones, which results in a metallicity  $Z = 0.011$ . Using a different value such as  $[\text{Fe}/\text{H}] = -0.06$  (Spina et al. 2017) has no effect on the results as the position of the isochrone on the CAMDs barely changes.

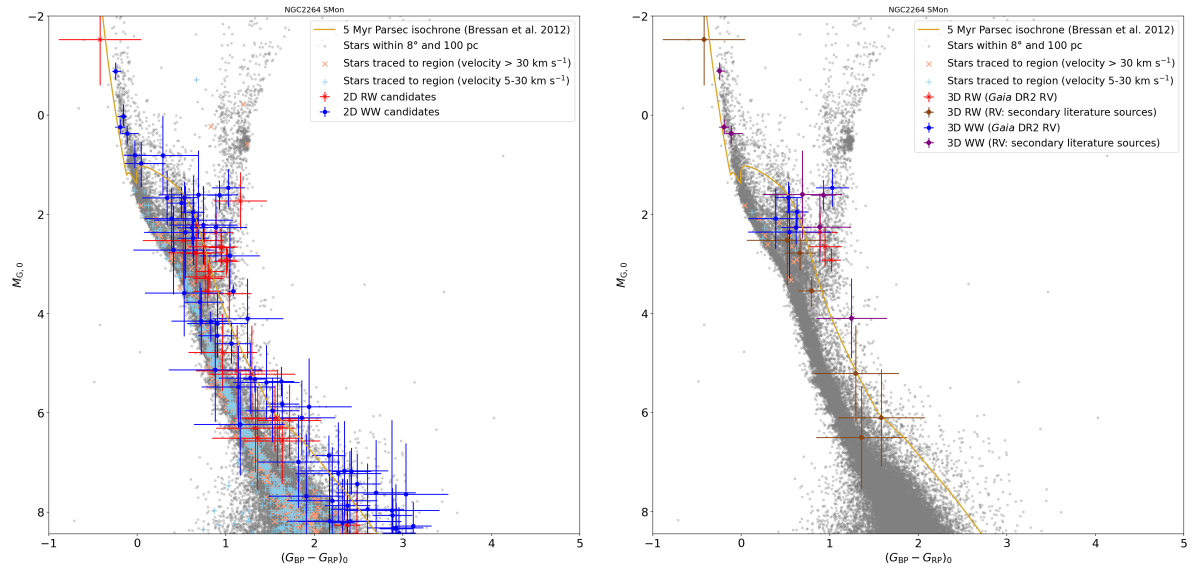


Figure 7.3: CAMDs showing all 2D (top) and 3D (bottom) RW ( $> 30 \text{ km s}^{-1}$ ) and WW ( $5\text{--}30 \text{ km s}^{-1}$ ) candidates that can be traced back to the cluster around S Mon (RW: red “x”, WW: blue square). I magnitude-limit the diagram to  $-2 \text{ mag} < M_{G,0} < 8.5 \text{ mag}$ , which corresponds to a  $G$ -magnitude  $\approx 18 \text{ mag}$  at the fainter end. Around this apparent magnitude value, the typical uncertainties in the 5-parameter astrometry increase quickly. The CAMDs include a large number of stars that I have traced back to the cluster around S Mon but that sit along the MS underneath the 5 Myr isochrone (RW-velocity: orange “x”, WW-velocity: lightblue “+”) even when considering their errors, which are not plotted here. These stars are therefore too old to have been born in S Mon. Many of my identified candidates have large errors in magnitude and colour, which are predominantly driven by the errors in the extinction and reddening. Some of these candidates sit below the 5 Myr isochrone but might be younger than their position suggests due to the large errors. In the right plot, I show the 3D-candidates with RV from *Gaia* DR2 (RW: red “x”, WW: blue square) and from secondary literature sources (RW: brown “x”, WW: purple dot).

Table 7.2: S Mon 3D RW and WW stars sorted by decreasing 3D-velocity. Column 2+3: velocity in S Mon rest frame [rf]; Column 3: RV sources - <sup>a</sup>Gaia DR2, <sup>b</sup>Jackson et al. (2016), <sup>c</sup>Luo et al. (2019), <sup>d</sup>Kounkel et al. (2019), <sup>e</sup>Fehrenbach et al. (1992), <sup>f</sup>Duflot et al. (1995); Column 4: minimum flight time since ejection (crossing of search boundary); Column 5: age from PARSEC isochrones (Bressan et al. 2012); Column 6–7: from literature sources - <sup>1</sup>Luo et al. (2019), <sup>2</sup>Venuti et al. (2015), <sup>3</sup>Venuti et al. (2017), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Kounkel et al. (2019), <sup>6</sup>Lamm et al. (2004), <sup>7</sup>Karlssohn (1972), <sup>8</sup>Cannon & Pickering (1993), <sup>9</sup>Voroshilov et al. (1985), <sup>10</sup>Paunzen et al. (2001), <sup>11</sup>Rebull et al. (2002), <sup>12</sup>Stassun et al. (2019), <sup>13</sup>Anders et al. (2019), <sup>14</sup>Venuti et al. (2018).

Gaia DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Flight time (Myr)	Iso. age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3D RW stars						
3132380637509393920	54.2 ± 0.7	-34.2 ± 6.3 <sup>a</sup>	0.9**	1.1 <sup>+1.0</sup> <sub>-0.6</sub>	1.1 <sup>13</sup>	G9 <sup>1</sup>
3326893624672681216	17.5 ± 1.1	59.5 ± 2.9 <sup>b</sup>	0.2	31.0 <sup>+19.0</sup> <sub>-26.0</sub>	0.7 <sup>12,13</sup>	-
3326630811329448576	8.4 ± 0.6	54.9 ± 15.3 <sup>d</sup>	0.8**	>0.5	2.3–2.9 <sup>12,13</sup>	B2 <sup>8</sup>
3132474680112352128	42.7 ± 0.6	-18.1 ± 6.0 <sup>c</sup>	1.3	10.0 <sup>+5.0</sup> <sub>-6.0</sub>	1.0 <sup>12,13</sup>	G5/6 <sup>1</sup>
3331678394335985152	41.1 ± 0.7	21.3 ± 4.0 <sup>a</sup>	2.0	1.8 <sup>+3.2</sup> <sub>-1.3</sub>	1.1 <sup>13</sup>	-
3134176728413264896	41.9 ± 0.6	8.0 ± 3.1 <sup>a</sup>	0.6***	1.8 <sup>+3.2</sup> <sub>-1.3</sub>	1.1 <sup>13</sup>	-
3326654725707134464	8.6 ± 0.5	34.1 ± 7.4 <sup>c</sup>	1.0***	12.0 <sup>+7.0</sup> <sub>-11.0</sub>	1.1–1.2 <sup>12,13</sup>	F9 <sup>1</sup>
3326632082639739264	27.7 ± 0.7	21.6 ± 2.9 <sup>b</sup>	0.3**	5.0 <sup>+35.0</sup> <sub>-4.3</sub>	0.9–1.1 <sup>12,13</sup>	-
3326908781612828544	30.5 ± 0.8	17.4 ± 2.9 <sup>b</sup>	0.1	5.0 <sup>+35.0</sup> <sub>-4.0</sub>	0.9–1.0 <sup>12,13</sup>	-
3131997187129420672	16.5 ± 1.1	28.8 ± 12.2 <sup>c</sup>	3.8	10.0 <sup>+3.0</sup> <sub>-6.0</sub>	1.2–1.3 <sup>12,13</sup>	F8 <sup>1</sup>
3D WW stars						
3134179335455713408	11.5 ± 0.7	26.5 ± 6.6 <sup>a</sup>	2.0	7.0 <sup>+3.0</sup> <sub>-4.0</sub>	1.2–2.1 <sup>12,13</sup>	-
3327008867233046528	6.7 ± 0.8	-24.7 ± 3.8 <sup>e</sup>	0.9	>0.7	3.3 <sup>13</sup>	B5 <sup>9</sup>
3331597816450524288	22.4 ± 0.6	-9.0 ± 3.2 <sup>a</sup>	3.5	5.0 ± 2.0	1.2–1.5 <sup>12,13</sup>	-
3132933764876638848	19.1 ± 0.5	-13.6 ± 3.1 <sup>a</sup>	2.6*	0.7 <sup>+1.3</sup> <sub>-0.5</sub>	1.1 <sup>13</sup>	G5 <sup>1</sup>
3326713442204844160	8.1 ± 0.5	-16.1 ± 9.4 <sup>b</sup>	0.1	>2.5	2.0 <sup>12,13</sup>	A2/3 <sup>7</sup>
3351602404024775168	16.8 ± 0.6	6.3 ± 3.0 <sup>a</sup>	2.1*	0.3 <sup>+0.7</sup> <sub>-0.1</sub>	1.1 <sup>13</sup>	G5/K1 <sup>1</sup>
3326938567209095936	8.7 ± 0.6	15.5 ± 3.0 <sup>b</sup>	0.2	3.0 <sup>+4.0</sup> <sub>-2.7</sub>	2.1 <sup>12,13</sup>	B8 <sup>9</sup>
3327203588170236672	13.3 ± 0.6	-7.2 ± 5.8 <sup>a</sup>	2.0	10.0 <sup>+5.0</sup> <sub>-16.0</sub>	1.5–1.9 <sup>12,13</sup>	-
3351770835461871360	14.6 ± 0.7	-0.2 ± 5.6 <sup>a</sup>	3.1	5.0 <sup>+3.0</sup> <sub>-2.0</sub>	1.5–1.6 <sup>12,13</sup>	-
3326693857153492736	4.2 ± 0.5	11.7 ± 2.9 <sup>d</sup>	0.6**	1.3 <sup>+13.7</sup> <sub>-1.0</sub>	0.8–1.3 <sup>4,13,14</sup>	-
3326576656084295296	10.7 ± 0.6	3.4 ± 7.2 <sup>c</sup>	0.9	1.8 <sup>+8.0</sup> <sub>-1.5</sub>	1.0 <sup>13</sup>	K1 <sup>1</sup>
3326739933562218496	1.5 ± 0.5	8.8 ± 9.3 <sup>a</sup>	in cluster	10.0 <sup>+8.0</sup> <sub>-9.0</sub>	0.9–2.1 <sup>4,12,13</sup>	K0 <sup>4,11</sup>
3326740693772293248	6.0 ± 0.8	0.2 ± 2.9 <sup>f</sup>	in cluster	>2.0	2.3 <sup>13</sup>	A1 <sup>10</sup>

\*Age estimate is smaller than the flight time; \*\*more likely from IRS 1, \*\*\*more likely from IRS 2

## 7.3 Results from *Gaia* DR2

I analyse the *Gaia* DR2 trace-backs separately for S Mon and IRS 1/2, due to the difference in age estimates for these regions, requiring different PARSEC isochrones and time limits for the flight times to be applied. While the regions encompassed by the search radius around IRS 1 and IRS 2 overlap in my analysis, I draw the search radius separately around each of the subcluster centres and trace back stars to each of them. However, when I compare the results from IRS 1/2 to the simulations later, I consider them as one subcluster having evolved together.

### 7.3.1 RW/WW stars from S Mon

Fig. 7.3 shows the CAMDs for the RW/WW candidates that can be traced back to S Mon in the past 5 Myr. On the left, I show all 2D-candidates with absolute magnitudes of  $-2 \text{ mag} < M_{G,0} < 8.5 \text{ mag}$ . I find further candidates below the 8.5 mag limit; however, the uncertainties for these candidates are too large for them to be included in my results. On the right in this figure, I show all 3D-candidates within the same magnitude range. In addition to 3D-candidates traced back using *Gaia* DR2 RV, I also trace back several 3D-candidates using RVs from secondary literature sources.

I can trace back 30 2D RW and 65 2D WW candidates to S Mon. Not all of these 2D-candidates turn into 3D-candidates. Several of them are missing RVs, so their full status is unclear until this velocity measurement is available. Of the 30 2D RW candidates, 21 candidates have RV measurements. Of these, ten candidates also trace back in 3D, three with RVs from *Gaia* DR2, seven with RVs from secondary literature sources.

Table 7.2 gives information about the candidates, that trace in all three dimensions, while Tables B.2 and B.3 in Appendix B provide the information for the 2D-candidates (and non-3D trace-backs), such as their 2D-velocity and RV in my chosen reference frame, flight time since ejection and an age estimate based on PARSEC isochrones. I have also searched for additional information in the literature about my sources and have found mass estimates and spectral types for several sources; however, only a few independent age estimates. I have not included these independent age estimates in the tables but comment where they are not within my age estimates from the PARSEC isochrones.

The brightest 3D RW candidate is HD 262042 (*Gaia* DR2 3326630811329448576), which is only a WW when considering its proper motion and turns into a RW due to its large RV (Kounkel et al. 2019). It has already reached the MS, which makes an age estimate with isochrones more difficult. The isochrone analysis provides a minimum

age of  $\sim 0.5$  Myr; however, Gontcharov (2012) suggested an age of 159 Myr. It has an estimate for its mass of  $\sim 2.3\text{--}2.9 M_{\odot}$  (Stassun et al. 2019; Anders et al. 2019) and a B2 spectral type (Cannon & Pickering 1993). The mass estimate is low for the given spectral type; however, the spectral type identification is quite old. This star also traces back to IRS 1 and IRS 2 and I discuss how I handle this issue in section 7.5.1.

The faintest 3D RW candidate is CSIMon-012143 (*Gaia* DR2 3326893624672681216) with an absolute magnitude of  $\sim 6.5$  mag and a mass estimate of  $\sim 0.7 M_{\odot}$  (Stassun et al. 2019; Anders et al. 2019). It is located underneath the 5 Myr isochrone but its error bars make it possibly young enough to have originated from S Mon. It is the second fastest 3D RW from this region and the fastest one only tracing back to S Mon with a space velocity of  $\sim 62 \text{ km s}^{-1}$  in my reference frame.

My fastest 3D RW candidate is *Gaia* DR2 3132380637509393920 with a space velocity of  $\sim 64 \text{ km s}^{-1}$ . It has an absolute magnitude of  $\sim 2.9$  mag, a mass estimate of  $\sim 1.1 M_{\odot}$  (Anders et al. 2019) and a G9 spectral type (Luo et al. 2019); however, this star also traces back to IRS 1/2.

Of the 65 2D WW candidates, only 18 candidates have RV measurements. 13 of these candidates trace back to S Mon in 3D, six with an RV measurement from secondary literature sources. I have three 3D WW candidates that have already reached the MS since S Mon formed.

The brightest of these three MS trace-backs is *Gaia* DR2 3327008867233046528 (HD 47662), which has an absolute magnitude of  $\sim -0.9$  mag, a mass estimate of  $\sim 3.3 M_{\odot}$  (Anders et al. 2019) and a spectral type of B5 (Voroshilov et al. 1985). My isochronal age analysis provides a minimum age estimate of  $\sim 0.7$  Myr; however, Tetzlaff et al. (2011) suggested an age of  $\sim 50$  Myr. The next brightest is BD+10 1222 (*Gaia* DR2 3326740693772293248) with an absolute magnitude of  $\sim 0.2$  mag, a mass estimate of  $\sim 2.3 M_{\odot}$  (Anders et al. 2019) and an A1 spectral type (Paunzen et al. 2001). The minimum age estimate from my isochrones is  $\sim 2.0$  Myr and Dahm et al. (2007) suggested a very similar age of  $\sim 2.1$  Myr. Finally, I have HD 261737 (*Gaia* DR2 3326713442204844160) with an absolute magnitude of  $\sim 0.4$  mag and an age estimate of  $\gtrsim 2.5$  Myr, a mass estimate of  $2.0 M_{\odot}$  (Stassun et al. 2019; Anders et al. 2019) and a A2/3 spectral type (Karlsson 1972). For this star, I have not found an independent age estimate. Of these three bright 3D WW candidates, only the brightest (*Gaia* DR2 3327008867233046528) also traces back to IRS 1 and 2, the other two are trace-backs to S Mon only.

The faintest 3D WW candidate is *Gaia* DR2 3326693857153492736 (CI\* NGC 2264 VAS 204) with an absolute magnitude of  $\sim 4.1$  mag. It has a mass estimate of  $0.8\text{--}1.3 M_{\odot}$

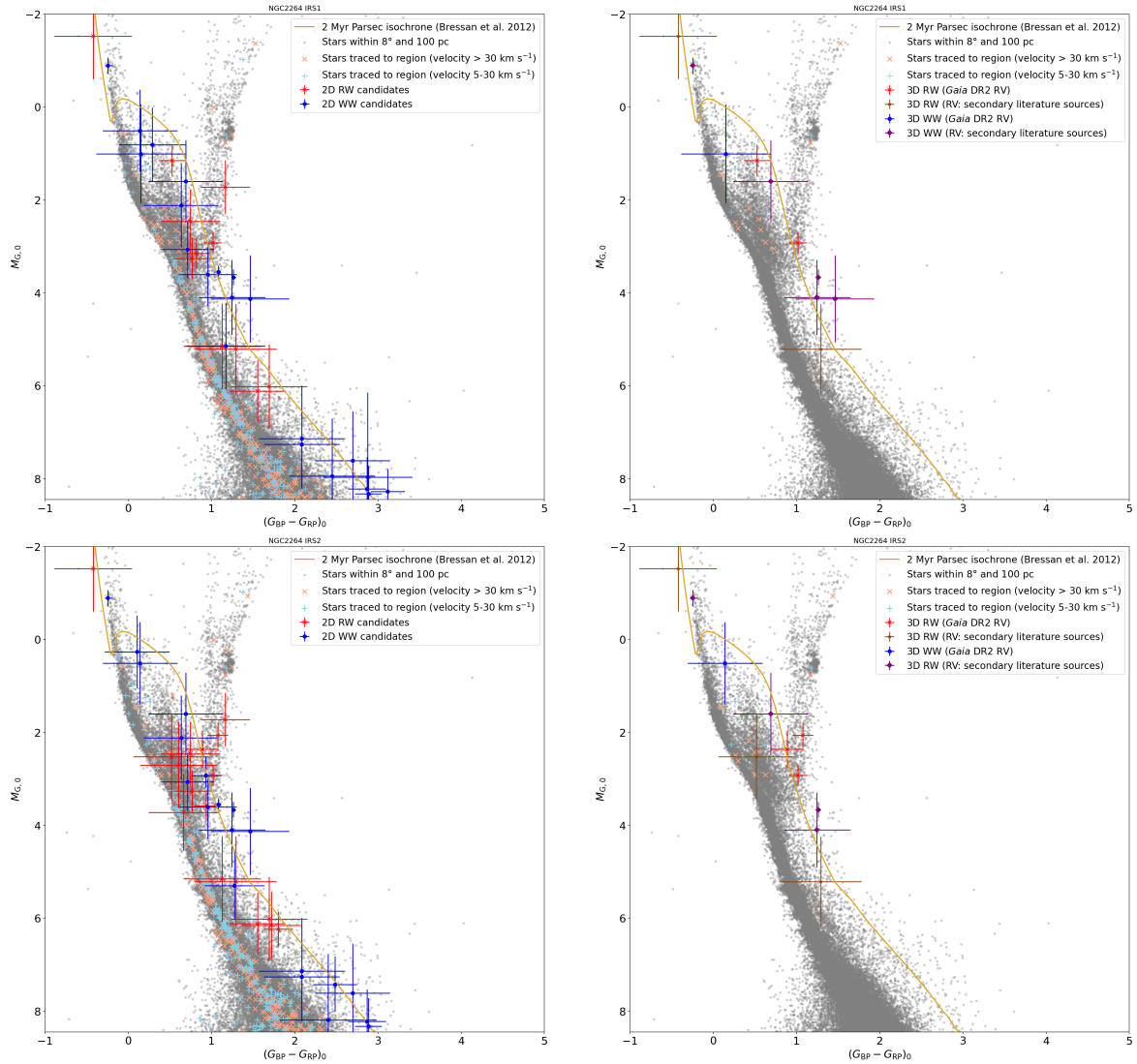


Figure 7.4: CAMDs showing all 2D (left) and 3D (right) RW ( $> 30 \text{ km s}^{-1}$ ) and WW ( $5\text{--}30 \text{ km s}^{-1}$ ) candidates that can be traced back to the cluster around IRS 1 (top) and IRS 2 (bottom) (RW: red “x”, WW: blue square). I magnitude-limit the diagram to  $-2 \text{ mag} < M_{G,0} < 8.5 \text{ mag}$ , which corresponds to a  $G$ -magnitude  $\approx 18 \text{ mag}$  at the fainter end. Around this apparent magnitude value, the typical uncertainties in the 5-parameter astrometry increase quickly. The CAMDs include a large number of stars that I have traced back to the cluster around IRS 1 or IRS 2 but that sit along the MS underneath the 2 Myr isochrone (RW-velocity: orange “x”, WW-velocity: lightblue “+”) even when considering their errors, which are not plotted here. These stars are therefore too old to have been born in IRS1/2. Many of my identified candidates have large errors in magnitude and colour, which are predominantly driven by the errors in the extinction and reddening. Some of these candidates sit below the 2 Myr isochrone but might be younger than their position suggests due to the large errors. In the right plot, I show the 3D-candidates with RV from *Gaia* DR2 (RW: red “x”, WW: blue square) and from secondary literature sources (RW: brown “x”, WW: purple dot).

(Venuti et al. 2018; Anders et al. 2019) but no spectral type information. This is one of the few stars for which I have been able to find an independent age estimate. My age estimate from PARSEC isochrones is  $\sim 1.3$  Myr, the age estimate from Venuti et al. (2018) is  $\sim 1.7$  Myr, which is within the uncertainties of my result.

I trace back a total of ten 3D RW candidates to S Mon; however, find five of these candidates also trace back to either IRS 1 or IRS 2, which could reduce the total number. At WW-velocities, I trace back 13 3D-candidates, one of which also traces back to IRS 1. I find two MS stars, one at RW-velocities the other at WW-velocities, which based on independent age estimates are too old to have originated in S Mon. I exclude these two stars from further analysis.

### 7.3.2 RW/WW stars from IRS 1 and IRS 2

Fig. 7.4 shows the CAMDs with all 2D- and 3D RW and WW candidates for the subclusters IRS 1 (first row) and IRS 2 (second row). On the left, I identify all 2D-candidates that I can trace back to either region, whereas on the right are the 3D-trace-backs, where RVs are considered as well. The magnitude range is the same as for S Mon. I also find further candidates with absolute magnitudes fainter than 8.5 mag; however, do not identify these here due to the very large uncertainties.

I find eleven 2D RW candidates that trace back to IRS 1 and 18 2D RW candidates that trace back to IRS 2. Several of these candidates trace back to both regions and I will comment on this in section 7.5.1. In 3D, I can trace back four RW-candidates to IRS 1, two of these with secondary RV. To IRS 2, I can trace back six 3D RW-candidates, three of these with secondary RV. Table 7.3 identifies all 3D RW and WW stars traced back to these subclusters and I also identify the subcluster to which these are traced back. The information for the 2D-candidates is located in Appendix B in Tables B.4 and B.5.

As mentioned in section 7.3.1, the brightest 3D RW candidate HD 262042 (*Gaia* DR2 3326630811329448576) traces back to all three regions and is a MS-star. The next brightest 3D RW candidate is TYC 747-2093-1 (*Gaia* DR2 3134140405870541568) and only traces back to IRS 1 with a flight time of  $\sim 1.7$  Myr. It has an absolute magnitude of  $\sim 1.2$  mag and an isochronal age estimate of  $\sim 3.5$  Myr, which is just at my upper age boundary of 2 Myr when considering the errors. Its spectral type is F6/7 based on information in Luo et al. (2019) and its mass estimate is  $\sim 1.2$ – $1.9 M_{\odot}$  (Stassun et al. 2019; Anders et al. 2019). The fastest 3D RW candidate from IRS 1 is the same as from S Mon, *Gaia* DR2 3132380637509393920 with a space velocity of  $\sim 64 \text{ km s}^{-1}$ .

The second brightest and fastest 3D RW candidate only tracing back to IRS 2 is



Table 7.3: IRS 1/2 3D RW and WW stars sorted by decreasing 3D-velocity. Column 2+3: velocity in respective IRS rest frame [rf]; Column 3: RV sources - <sup>a</sup>Gaia DR2, <sup>b</sup>Jackson et al. (2016), <sup>c</sup>Luo et al. (2019), <sup>d</sup>Kounkel et al. (2019), <sup>e</sup>Fehrenbach et al. (1992), <sup>f</sup>Duflot et al. (1995); Column 4: minimum flight time since ejection (crossing of search boundary); Column 5: age from PARSEC isochrones (Bressan et al. 2012); Column 6: Subcluster identification; Column 7–8: from literature sources - <sup>1</sup>Luo et al. (2019), <sup>2</sup>Venuti et al. (2015), <sup>3</sup>Venuti et al. (2017), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Kounkel et al. (2019), <sup>6</sup>Lamm et al. (2004), <sup>7</sup>Karlsson (1972), <sup>8</sup>Cannon & Pickering (1993), <sup>9</sup>Voroshilov et al. (1985), <sup>10</sup>Paunzen et al. (2001), <sup>11</sup>Rebull et al. (2002), <sup>12</sup>Stassun et al. (2019), <sup>13</sup>Anders et al. (2019), <sup>14</sup>Venuti et al. (2018).

Gaia DR2 source-id	2D-velocity rf (kms <sup>-1</sup> )	Radial velocity rf (kms <sup>-1</sup> )	Flight time (Myr)	Iso. age (Myr)	Subcluster	Mass (M <sub>⊙</sub> )	Spectral type
3D RW stars							
3132688088448708736	53.8 ± 1.2	69.6 ± 4.1 <sup>a</sup>	0.9	0.4 <sup>+0.5</sup> <sub>-0.3</sub>	IRS 2	1.1 <sup>13</sup>	K1 <sup>1</sup>
3132380637509393920	55.0 ± 0.8	-33.8 ± 6.3 <sup>a</sup>	0.9**	1.1 <sup>+1.0</sup> <sub>-0.6</sub>	IRS 2	1.1 <sup>13</sup>	G9 <sup>1</sup>
3132380637509393920	53.9 ± 0.9	-33.7 ± 6.3 <sup>a</sup>	0.9	1.1 <sup>+1.0</sup> <sub>-0.6</sub>	IRS 1	1.1 <sup>13</sup>	G9 <sup>1</sup>
3326630811329448576	9.2 ± 0.7	55.1 ± 15.3 <sup>d</sup>	0.4**	>0.5	IRS 2	2.3–2.9 <sup>12,13</sup>	B2 <sup>8</sup>
3326630811329448576	7.8 ± 0.9	55.1 ± 15.3 <sup>d</sup>	0.2	>0.5	IRS 1	2.3–2.9 <sup>12,13</sup>	B2 <sup>8</sup>
3134176728413264896	43.4 ± 0.8	8.3 ± 3.1 <sup>a</sup>	0.5	1.8 <sup>+3.2</sup> <sub>-1.3</sub>	IRS 2	1.1 <sup>13</sup>	-
3326632082639739264	29.1 ± 0.8	21.8 ± 2.9 <sup>b</sup>	0.1**	5.0 <sup>+35.0</sup> <sub>-4.3</sub>	IRS 2	0.9–1.1 <sup>12,13</sup>	-
3326654725707134464	10.9 ± 0.7	34.2 ± 7.4 <sup>c</sup>	0.7	12.0 <sup>+7.0</sup> <sub>-11.0</sub>	IRS 2	1.1–1.2 <sup>12,13</sup>	F9 <sup>1</sup>
3326632082639739264	27.7 ± 1.0	21.9 ± 2.9 <sup>b</sup>	0.1	5.0 <sup>+35.0</sup> <sub>-4.3</sub>	IRS 1	0.9–1.1 <sup>12,13</sup>	-
3134140405870541568	17.4 ± 1.1	-27.9 ± 20.0 <sup>a</sup>	1.7	3.5 ± 1.5	IRS 1	1.2–1.9 <sup>12,13</sup>	F6/7 <sup>1</sup>
3D WW stars							
3327008867233046528	7.0 ± 1.1	-24.6 ± 3.8 <sup>c</sup>	1.6*	>0.7	IRS 1	3.3 <sup>13</sup>	B5 <sup>9</sup>
3327008867233046528	6.0 ± 0.9	-24.6 ± 3.8 <sup>c</sup>	1.7*	>0.7	IRS 2	3.3 <sup>13</sup>	B5 <sup>9</sup>
3326938567209095936	9.7 ± 0.9	15.6 ± 3.0 <sup>b</sup>	0.6*	3.0 <sup>+4.0</sup> <sub>-2.7</sub>	IRS 1	2.1 <sup>12,13</sup>	B8 <sup>9</sup>
3326938567209095936	9.0 ± 0.8	15.6 ± 3.0 <sup>b</sup>	0.6*	3.0 <sup>+4.0</sup> <sub>-2.7</sub>	IRS 2	2.1 <sup>12,13</sup>	B8 <sup>9</sup>
3326685512032888320	9.5 ± 0.8	13.3 ± 10.4 <sup>a</sup>	0.4	3.0 <sup>+3.0</sup> <sub>-1.5</sub>	IRS 2	1.2–2.1 <sup>4,12,13</sup>	F5/G0 <sup>4,6</sup>
3326693857153492736	6.0 ± 0.7	11.8 ± 2.9 <sup>d</sup>	0.3**	1.3 <sup>+13.7</sup> <sub>-1.0</sub>	IRS 2	0.8–1.3 <sup>4,13,14</sup>	K6 <sup>4</sup>
3326704238089925120	10.5 ± 0.8	-7.9 ± 2.9 <sup>d</sup>	0.1	0.5 <sup>+3.0</sup> <sub>-0.4</sub>	IRS 1	0.6–1.0 <sup>4,13</sup>	M0 <sup>2,4</sup>
3326693857153492736	4.7 ± 0.8	11.9 ± 2.9 <sup>d</sup>	0.2	1.3 <sup>+13.7</sup> <sub>-1.0</sub>	IRS 1	0.8–1.3 <sup>4,13,14</sup>	K6 <sup>4</sup>
3326695991753074304	6.1 ± 1.2	5.0 ± 2.9 <sup>b</sup>	in cluster	0.6 ± 0.1	IRS 1	0.3 <sup>4</sup>	M3 <sup>4</sup>
3326695991753074304	4.9 ± 1.0	5.0 ± 2.9 <sup>b</sup>	in cluster	0.6 ± 0.1	IRS 2	0.3 <sup>4</sup>	M3 <sup>4</sup>
3326682552799138176	0.8 ± 0.9	-6.7 ± 11.5 <sup>a</sup>	in cluster	5.0 <sup>+5.0</sup> <sub>-3.1</sub>	IRS 1	-	-

\*More likely from S Mon, \*\*more likely from IRS 1, \*\*\*more likely from IRS 2

*Gaia* DR2 3132688088448708736. Its space velocity is  $\sim 88 \text{ km s}^{-1}$  and it has an absolute magnitude of  $\sim 2.1$  mag. Its estimated isochronal age is  $\sim 0.4$  Myr, its mass estimate is  $\sim 1.1 M_{\odot}$  (source) and the spectral type is K1 (Luo et al. 2019).

The faintest 3D RW candidate CSIMon-005775 (*Gaia* DR2 3326632082639739264) traces back to all three regions analysed here; however, its flight time to S Mon is  $\sim 0.2$  Myr longer than to IRS 1/2. Its absolute magnitude is  $\sim 5.2$  mag and its age estimate is  $\sim 5.0$  Myr with a large error putting it into the age range of IRS 1/2.

At WW-velocities, I trace back 21 2D-candidates to IRS 1 and 20 2D-candidates to IRS 2. Of these 2D-candidates, six trace back to IRS 1 in 3D, five of them with secondary RV measurements. Five trace back to IRS 2 in 3D, four of them with secondary RVs.

The brightest 3D WW candidate is the same for IRS 1 and IRS 2 as for S Mon, HD 47662 (*Gaia* DR2 3327008867233046528). This star already has reached the MS and it is possibly too old to have originated from NGC 2264 based on its age estimate of 50 Myr from Tetzlaff et al. (2011). The second brightest 3D WW candidate tracing only back to IRS 1 is NGC 2264 118 (*Gaia* DR2 3326682552799138176). Its absolute magnitude is  $\sim 1.0$  mag and its age estimate is  $\sim 5.0$  Myr and it is only young enough to possibly have originated from IRS 1 given its age errors. It is also one of the slowest 3D WW stars, with a proper motion far below the WW-velocity and also still located within the cluster, so could possibly still be bound to it when considering its large RV errors.

Only one 3D WW star traces solely back to IRS 2, *Gaia* DR2 3326685512032888320 (NGC 2264 189). It has an absolute magnitude of  $\sim 0.5$  mag and an estimated age of  $\sim 3.0$  Myr from the PARSEC isochrones with errors large enough to put it in the age range of IRS 2. I found information on the spectral type in two literature sources with it being either a F5 (Lamm et al. 2004) or a G0 spectral type (Venuti et al. 2014). In Venuti et al. (2014), I also find a mass estimate of  $\sim 2.1 M_{\odot}$  and an age estimate  $\sim 4.7$  Myr (Venuti et al. 2014), which is consistent with my age estimate when considering the errors. However, this independent age estimate puts this star outside the age range for IRS 2.

The faintest 3D WW candidate V\* V500 Mon (*Gaia* DR2 3326704238089925120) traces back in 2D to both IRS 1 and 2; however, in 3D it only traces back to IRS 1. It has an absolute magnitude of  $\sim 4.1$  mag and an age estimate from PARSEC isochrones of  $\sim 0.5$  Myr. This age is consistent with the estimate from Venuti et al. (2014) of  $\sim 0.7$  Myr.

I trace back a total of ten 3D RW candidates to IRS 1 and IRS 2; however, find three of these candidates trace back to both regions. At WW-velocities, I trace back eleven 3D-

candidates, four of which also trace back to S Mon and two common trace-backs between IRS 1 and 2. I find two MS stars, one at RW-velocities the other at WW-velocities, which based on independent age estimates are too old to have originated in IRS 1 or 2. I exclude these two stars from further analysis.

### 7.3.3 Confirming previously identified ejected stars

Maíz Apellániz et al. (2020) found two ejected stars from NGC 2264 and suggested they are more likely to originate from the northern region around S Mon. These two stars are *Gaia* DR2 3326734332924414976 and *Gaia* DR2 3326951215889632128.

I confirm that *Gaia* DR2 3326951215889632128 also traces back to S Mon in my analysis at a 2D WW-velocity of  $\sim 14 \text{ km s}^{-1}$ ; however, this star does not have a RV measurement, so cannot be confirmed in 3D. It is located very close to S Mon and is within the age range of this subcluster based on the isochrone analysis, so could either originate or possibly be a future visitor depending on the direction of its RV.

*Gaia* DR2 3326734332924414976 also traces back to S Mon in 2D at WW-velocities in my analysis, but it is again missing the RV. When plotted on the CAMD, it appears too old to have originated in S Mon as it is located right on the MS with very small photometry errors taken directly from *Gaia* DR2. The age estimate from its position on the MS puts it at an age of at least  $\sim 15$  Myr but a different method to determine its age might change my age estimate.

### 7.3.4 3D-candidates with a protoplanetary disc

In Schoettler & Parker (2021), I searched for evidence of the presence of protoplanetary discs around RWs and WWs (and past/future visitors) from the ONC identified in Schoettler et al. (2020). I showed that in addition to some of the protoplanetary discs surviving the ejection of their host stars from their birth clusters, these star-disc systems can encounter a second dense star-forming environment and possibly emerge with an intact disc from this encounter.

I repeat this analysis for NGC 2264 and search through the Simbad/VizieR databases (Wenger et al. 2000; Ochsenbein et al. 2000). In Appendix B, Table B.6, I show the older past visitors to the regions within the past 2 Myr (IRS 1 and IRS 2) and 5 Myr (S Mon) and in Tables B.7 and B.8 I show all future visitors to all subclusters of NGC 2264 up to a cluster age of 10 Myr. I have searched through the databases for any evidence of protoplanetary discs around the identified RW/WWs. I was successful in this endeavour

for several of my ejected 3D-trace-back stars from the region. Also, one of the future visitors to the region shows clear evidence of a disc.

*Gaia* DR2 3326704238089925120 is my faintest 3D WW trace-back to IRS 1. It appears in [Rebull et al. \(2002\)](#) with an IR-excess value  $\Delta(I_C-K) = 0.32$  mag. The authors used the IR-excess limit stated in [Hillenbrand et al. \(1998\)](#), where a value of  $\Delta(I_C-K) > 0.30$  mag is a clear indicator of a protoplanetary disc. This star also appears in [Cieza & Baliber \(2007\)](#) with a value of  $[3.6 \mu\text{m}] - [8 \mu\text{m}] > 1$  mag. A mid-IR colour index above 1 mag is suggested by [Rebull et al. \(2006\)](#) as an indicator of a disc.

*Gaia* DR2 3326695991753074304 traces back in 3D to both IRS 1 and 2 as a WW star and according to [Venuti et al. \(2018\)](#), who used IR- and UV-excess to identify disc-bearing, accreting objects, it is a disc-bearing star showing evidence of on-going accretion. It has also been previously identified as an accreting CTTS by [Venuti et al. \(2014\)](#). *Gaia* DR2 3326693857153492736 is a 3D WW tracing back to all three regions. It appeared in [Venuti et al. \(2018\)](#), but it was considered a non-accreting source there. However, according to [Venuti et al. \(2014\)](#) and [Venuti et al. \(2017\)](#) it was suggested to be a CTTS, which could be disc-bearing and accreting.

*Gaia* DR2 3326687878559500288 could be a future visitor to IRS 1 and 2 approaching these subclusters at a WW-velocity of  $\sim 10 \text{ km s}^{-1}$ . It is suggested to be an accreting CTTS in [Fűrész et al. \(2006, based on H \$\alpha\$ \)](#), [Venuti et al. \(2015\)](#) and [Venuti et al. \(2017\)](#). Its estimated age from the PARSEC isochrones is  $7.0 \pm 6.5$  Myr. Its maximum flight time since ejection from its birth cluster is therefore  $\sim 13.5$  Myr if I consider an ejection right after formation. Given its velocity and this maximum flight time, it could have travelled as far as  $\sim 130$  pc during this time. Its velocity and distance measurements show large margins of error, so I do not attempt to locate its possible origin. It has a K1.5 spectra type ([McGinnis et al. 2018](#)) and due to its close proximity to NGC 2264, it is typically associated with the cluster, even though it was not born there based on its kinematics.

*Gaia* DR2 3326689807000188032 is a visitor, which is currently passing through IRS 1 and IRS 2. It appears in [Venuti et al. \(2014\)](#) and [Venuti et al. \(2017\)](#) identified as a WTTS. As such it is no longer accreting, but with an age of  $\sim 2.1$ –10 Myr it would still be young enough to possibly have a debris disc.

## 7.4 RW and WW stars from $N$ -body simulations

I now use the  $N$ -body simulations to give me the predicted number and velocities of the RW and WW stars to compare to the observational data. I have fewer constraints on

the initial spatial and kinematic substructure of NGC 2264 compared to the ONC that was investigated in [Schoettler et al. \(2020\)](#).

[Fűrész et al. \(2006\)](#) suggested that the current subclustering observable in NGC 2264 is likely a remnant of spatial substructure that was initially present in this cluster. They also found that the kinematics of the region correlate with the spatial substructure. [Tobin et al. \(2015a\)](#) refined the results of [Fűrész et al. \(2006\)](#) by using further RV measurements. They found further velocity substructure, i.e. groups with different velocities and suggested that NGC 2264 is possibly a collection of star-forming clumps instead of a dense bound cluster. [Costado & Alfaro \(2018\)](#) studied the spatial and kinematic structure in the larger Monoceros region of which NGC 2264 is a part. They confirmed the previous results on the presence of kinematic substructure.

Based on these literature sources, I suspect that NGC 2264 started with some level of initial spatial and kinematic substructure as traces of both appear to be still present at current times. This substructure cannot be created by dynamical evolution, only erased (e.g. [Goodwin & Whitworth 2004](#); [Allison et al. 2010](#); [Jaehnig et al. 2015](#)). As I have no specific information on the exact level of substructure, I run a larger set of different initial conditions based on combinations of different initial density, spatial and kinematic substructure.

I use the same method for setting up the simulations as in Chapter 5. The general set-up of the simulations is described in detail chapter 2 and in the following section, I comment on parameters specific to the NGC 2264-like simulations.

### 7.4.1 Simulation set-up

In my  $N$ -body simulations, I use several different initial fractal dimensions, which are shown in Table 7.4. This table provides an overview of all the initial condition combinations used in my  $N$ -body simulations. The Simulation IDs are used to refer to my different sets of initial conditions and are a number combination of the fractal dimension, initial virial ratio and initial radius. I also provide initial median stellar densities for each of the initial condition combinations in Table 7.4.

A fractal dimension of  $D = 1.6$  represents a high amount of initial spatial substructure, whereas  $D = 3.0$  is representative of a completely smooth distribution. For lower initial density simulations with an initial radius of 5 pc, I use a value of  $D = 2.0$  to represent a moderate amount of substructure and do not use a fractal dimension of  $D = 3.0$ . The lower initial stellar density combined with a smooth initial spatial substructure represented by  $D = 3.0$  results in a very slow dynamical evolution of the simulated region

Table 7.4: Overview of all initial condition combinations used in my  $N$ -body simulations. Column 1 shows the Simulation ID, which is a number combination of the fractal dimension, initial virial ratio and initial radius; Column 2 shows the fractal dimension  $D$ ; Column 3 shows the initial virial ratio  $\alpha_{\text{vir}}$ ; Column 4 shows the initial radius of the simulated regions  $r_F$ ; Column 5 shows the initial median stellar density  $\tilde{\rho}$ ; Column 6 shows the initial (observable) median stellar surface density  $\tilde{\Sigma}$ .

Sim. ID	$D$	$\alpha_{\text{vir}}$	$r_F$ (pc)	$\tilde{\rho}$ ( $M_{\odot} \text{pc}^{-3}$ )	$\tilde{\Sigma}$ (stars $\text{pc}^{-2}$ )
16-03-1	1.6	0.3	1	10 000	3 000
30-03-1	3.0	0.3	1	150	400
16-05-1	1.6	0.5	1	10 000	3 000
30-05-1	3.0	0.5	1	150	400
16-03-5	1.6	0.3	5	70	100
20-03-5	2.0	0.3	5	10	40
16-05-5	1.6	0.5	5	70	100
20-05-5	2.0	0.5	5	10	40

and would only produce RWs at older cluster ages, if any are produced at all.

Fractals can also be used to set up the initial kinematic substructure. Stars that are located close to each other have correlated velocities, whereas stars at a larger distance can have very different velocities (e.g. Larson 1981; Goodwin & Whitworth 2004). In my simulations, the velocities are scaled in such a way that the regions are either initially subvirial with a virial ratio  $\alpha_{\text{vir}} = 0.3$  or initially virialized with a virial ratio  $\alpha_{\text{vir}} = 0.5$ .

More details about the construction of the fractals in the simulations can be found in Goodwin & Whitworth (2004), Parker et al. (2014) and Parker & Wright (2016). The set-up is also described in Chapter 2.

The number of systems in my simulations is much smaller than those in Schoettler et al. (2020) and is comparable to the number used in Schoettler et al. (2019). The number of members in NGC 2264 is suggested to be  $\sim 1400$ , distributed across the northern and southern regions (e.g. Teixeira et al. 2012). For simplicity, I assume that half of the stars are located in the northern region around S Mon and the other half in the southern regions around IRS 1 and IRS 2, which are subcluster with overlapping boundaries. For my simulations, I use a number of 725 systems per simulation with masses for these systems sampled randomly from a Maschberger (2013) IMF. I use a range of stellar masses between  $0.1 M_{\odot}$  and  $50 M_{\odot}$ . This upper mass limit is consistent with recent estimates for the primary star in S Mon ( $\sim 40$ - $50 M_{\odot}$ ; Tokovinin 2018; Maíz Apellániz 2019).

The binary fractions used in the simulations (see Chapter 2) result in an average total number of stars in my simulations of  $\sim 1000$  stars, average cluster masses of  $\sim 730 M_{\odot}$

Table 7.5: Ejected RW and WW stars from  $N$ -body simulations within the search radius of 100 pc at different times during the simulations. For all my initial condition combinations, I show averages from all 20 simulations and the maximum from a single simulation. I count ejected binary systems as one star when calculating averages and maxima. The uncertainties in my averages are the standard deviations. I show ejected stars with masses from 0.3–8  $M_{\odot}$ .

Mass $m$ ( $M_{\odot}$ )	Simulation ID							
RW stars	16-03-1	30-03-1	16-05-1	30-05-1	16-03-5	20-03-5	16-05-5	20-05-5
$0.3 \leq m < 8.0$								
- after 1 Myr	4.1±1.9 / 9	0.3±0.4 / 1	3.9±2.3 / 9	0.3±0.5 / 2	2.0±1.2 / 4	0.2±0.5 / 2	1.2±0.9 / 3	0.3±0.4 / 1
- after 2 Myr	2.9±1.6 / 6	0.4±0.6 / 2	3.2±2.0 / 9	0.6±0.9 / 4	1.5±1.1 / 4	0.2±0.5 / 2	1.0±0.9 / 3	0.2±0.4 / 1
- after 3 Myr	1.3±0.8 / 3	0.4±0.6 / 2	1.6±1.2 / 3	0.3±0.7 / 3	0.7±0.7 / 2	0 / 0	0.4±0.6 / 1	0.1±0.2 / 1
- after 4 Myr	0.5±0.7 / 2	0.4±0.5 / 1	0.7±0.8 / 3	0.2±0.5 / 2	0.2±0.4 / 1	0 / 0	0.1±0.3 / 1	0.1±0.2 / 1
- after 5 Myr	0.3±0.3 / 2	0.3±0.5 / 2	0.4±0.6 / 2	0.1±0.2 / 1	0.1±0.2 / 1	0 / 0	0.2±0.4 / 1	0 / 0
WW stars	16-03-1	30-03-1	16-05-1	30-05-1	16-03-5	20-03-5	16-05-5	20-05-5
$0.3 \leq m < 8.0$								
- after 1 Myr	27.6±7.0 / 43	2.0±1.4 / 5	25.3±5.2 / 36	1.1±1.1 / 4	9.0±3.1 / 17	1.9±1.6 / 6	9.0±2.0 / 14	2.3±1.2 / 5
- after 2 Myr	29.2±6.9 / 46	2.5±1.4 / 5	27.2±5.8 / 42	1.5±1.3 / 5	9.6±3.2 / 17	3.1±1.7 / 6	10.2±2.2 / 15	2.5±1.3 / 5
- after 3 Myr	30.3±6.9 / 47	3.1±1.6 / 7	29.9±6.2 / 45	1.8±1.4 / 5	10.1±3.8 / 19	3.4±1.8 / 6	10.3±2.4 / 16	2.8±1.3 / 5
- after 4 Myr	29.6±6.8 / 44	4.3±1.9 / 9	28.3±5.7 / 41	2.3±1.7 / 6	10.3±3.6 / 19	3.5±1.8 / 7	10.1±2.7 / 16	3.1±1.3 / 5
- after 5 Myr	28.1±6.4 / 42	4.4±2.1 / 9	27.0±5.6 / 38	2.4±1.9 / 7	10.0±3.8 / 19	3.8±1.9 / 7	10.0±3.1 / 16	3.2±1.6 / 8

and an escape velocity of  $\sim 2\text{--}3 \text{ km s}^{-1}$ , which confirms my choice of lower WW limit of  $5 \text{ km s}^{-1}$ . This cluster mass is similar to that suggested by [Teixeira et al. \(2012\)](#). However, I have considerably fewer stars contributing to this mass estimate, highlighting that the choice of an average mass of  $0.5 M_{\odot}$  in [Teixeira et al. \(2012\)](#) does not replicate results achieved with applying an IMF. This example illustrates the effect from sampling the mass of the secondary star from a flat mass ratio distribution, which leads to higher overall total masses for the star-forming regions compared to simulations where only single stars are present (see Chapter 4).

For the simulations, I use software from the **Starlab** environment: the  $N$ -body integrator **kira** and the stellar and binary evolution package **SeBa** (e.g. [Portegies Zwart et al. 1999, 2001](#)). My star-forming regions are evolved over a time period of 5 Myr, covering the age estimates for NGC 2264 used for this analysis and I take snapshots every 0.1 Myr. My initial radii are set at 1 pc (resulting in a higher initial stellar density,  $\sim 150\text{--}10\,000 M_{\odot} \text{ pc}^{-3}$ ) and 5 pc (resulting in a more moderate initial stellar density,  $\sim 10\text{--}70 M_{\odot} \text{ pc}^{-3}$ ), I have not applied any external tidal field.

The stellar systems in my simulations undergo both stellar and binary evolution and I see several SNe after 4 Myr. These SNe are the result of the stellar evolution of the highest mass stars in my simulations, all of which leave a black hole as the SN remnant during the analysis time of 5 Myr.

### 7.4.2 Numbers from the simulations

From my  $N$ -body simulations, I have predicted numbers for RW and WW stars across the full mass range ( $0.1\text{--}50 M_{\odot}$ ), which I show in the Appendix B in Table B.1. In the results from the *Gaia* DR2 observations I find the star (*Gaia* DR2 3326637442758920960) with the smallest mass amongst the RW/WWs has a mass of  $\sim 0.3 M_{\odot}$  and is an M3-star (Venuti et al. 2014).

Even though this star does not trace back to NGC 2264 in 3D, it gives me an indication of the lowest mass I am able to identify within my data set. My *Gaia* DR2 results show that there are no high-mass ( $> 8 M_{\odot}$ ) RWs or WWs from this region. When comparing my simulations to the *Gaia* DR2 observations, I therefore reduce the mass range to  $0.3\text{--}8 M_{\odot}$ . The number of ejected stars in the two velocity ranges from my simulations are shown in Table 7.5 for all my initial conditions within a radius of 100 pc. Appendix Table B.1 shows the numbers for the complete mass range separated into low/intermediate mass stars ( $0.1\text{--}8 M_{\odot}$ ) and high-mass stars ( $> 8 M_{\odot}$ ), also within 100 pc.

Schoettler et al. (2019) showed that I find more and faster RW/WW stars from simulations where the initial conditions are more spatially substructured and/or subvirial. I find the same trend in my simulations. Simulations with more initial spatial substructure produce a higher number of RW and WW stars (in particular simulations with IDs 16-03-1 and 16-05-1). At this high level of initial substructure (fractal dimension  $D = 1.6$ ), a change in the initial virial ratio, which sets the global bulk motion of the stars in my simulations, has very little effect on the number of RWs. Due to the initial spatial substructure, even globally initially virialized simulations are subvirial on local scales and undergo local collapse (e.g. Allison et al. 2010; Parker & Wright 2016), which causes ejection of stars at early times in the simulations.

The maximum number of RWs from a single simulation is ten, which I achieve in 16-03-1 ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.3$ , radius 1pc) between  $0.6\text{--}0.8$  Myr and in 16-05-1 ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.5$ , radius 1pc) between  $1.5\text{--}1.9$  Myr. These high numbers of RWs are only achieved at early ages and reduce quickly to a maximum of only two RWs after 5 Myr. High numbers of RWs occur only in initially highly substructured regions with an initial radius of 1 pc, regardless of initial virial ratio. For regions with an initial radius of 5 pc, none of my simulations reach these numbers of RWs and the highest number is four RWs achieved at early ages (up to 2 Myr).

The highest velocities are also reached at early simulation times. In Simulation ID 16-03-1, the fastest velocity is reached after 0.8 Myr with a value of 216 km/s by a



0.8  $M_{\odot}$  star. This is not only the fastest velocity in this set of simulations but across all of the simulations. At this velocity it takes this star only 0.4 Myr to leave the 100 pc region around its ejection site. The maximum velocities reached in the simulations are related to the initial conditions. Simulations that start more spatially substructured (lower fractal dimension  $D$ ) and in addition have a higher initial stellar density produce RWs with higher velocities. For these simulations the maximum velocities are above 100 km/s. If they are initially subvirial the maximum velocity reaches even higher values above 200 km/s. These main results from the simulations in this analysis are in line with what is shown in [Schoettler et al. \(2019\)](#) for a simpler set of initial conditions without primordial binaries.

I find high numbers of WWs (up to a maximum  $\sim 40$ – $50$ ) in two initial condition settings, both initially highly substructured and with an initial radius of 1 pc. I find lower numbers of WWs (up to a maximum  $\sim 15$ – $20$ ) from initially highly substructured simulations with an initial radius of 5 pc. Simulations with less or no initial substructure produce even fewer WWs ( $< 10$  WWs).

## 7.5 Discussion

### 7.5.1 3D-candidates tracing back to more than one subregion

In my results, I find several candidates that can be traced back to more than one subcluster as their velocity vector is oriented in such a way that an origin in more than one of the subclusters is possible. Five of the 3D RW-candidates that I trace back to S Mon can also be traced back to IRS 1 and/or IRS 2. Three 3D RW-candidates trace back to IRS 2 but also to IRS 1. Four 3D WW traces back to S Mon but also to IRS 1 and/or IRS 2 and two 3D WW trace back both to IRS 1 and IRS 2.

For these candidates, I check the flight times for each of these trajectories and consider the candidates to have originated from the subcluster which has the smallest flight time, i.e. time since ejection. My reasoning for this classification is as follows: if a star gets ejected from one region and then passes through another region it is possible that interactions in the second region can alter its trajectory. Once this interaction happens, it would be difficult/impossible to trace this star back to its origin region. I consider the alignment of a star's trajectory to a region "behind" another region to be a chance alignment and consider the first region a star traces back to as its origin region. However, it remains possible for these stars to have originated in the region further away from its current location. In my result tables I will still consider them a possible origin region,

but I note the more likely birth region.

For one of the 3D WW trace-backs (*Gaia* DR2 3326695991753074304) that traces back to IRS 1 and IRS 2, this approach fails. It is possibly still located within the overlap section of both subclusters and therefore could belong to either.

### 7.5.2 S Mon

When considering the approach for double trace-backs the number of successfully traced back 3D RW-candidates from S Mon reduces from nine to five stars. Of the 13 traced back 3D WW-stars, 11 stars remain after removing double trace-backs and the 3D WW MS star with an age estimate of 50 Myr. Two of these eleven have flight times that are larger than their estimated age, suggesting that they might have not originated from S Mon. This leaves me with nine WWs after also excluding these two.

I have two 3D WW candidates that only turn into WWs, when I consider their RVs. The 2D-velocity of one of these candidates (*Gaia* DR2 3326739933562218496) is so small that based on this it would still be considered bound to the cluster; it is also still located within the cluster (when the upper distance estimate is used). For this type of candidates it is important to make sure that their higher RV does not originate from binary motion. For NGC 2264 404 (*Gaia* DR2 3326739933562218496), there are only two RV measurements available on the Simbad database, a very old measurement in [Wilson \(1953\)](#) and the *Gaia* DR2 measurement. The older measurement has no error provided but the quality indicator “E” on Simbad means that the error is likely large. I consider these two measurements as being consistent with each other, so it does not suggest a binary origin. This star appeared in [Klagyivik et al. \(2013\)](#), where it is also not identified as a binary but as a YSO, so I keep it as part of my 3D WW list.

I now compare the S Mon 3D RW and WW stars from the *Gaia* DR2 observations to my simulations. The five RW stars found exceed the averages and maxima of most of the initial conditions sets shown in Table 7.5. This allows me to exclude several of these initial conditions as highly unlikely to be those of S Mon. However, I might have to revisit them again in the future, especially when a better restriction of the extent of the subcluster in the radial direction (i.e. depth) is available. Fig. 7.5 shows the maximum number of RWs at five different times for four selected initial conditions and I also plot the number of RWs I find in the observations. The maximum number of RWs decreases at later times in the simulations and I find that the highest number of RWs is achieved in two different initial condition sets (16-03-1 and 16-05-1). I also plot the observed numbers for S Mon (5 RWs) at its maximum age used for the PARSEC isochrones (5

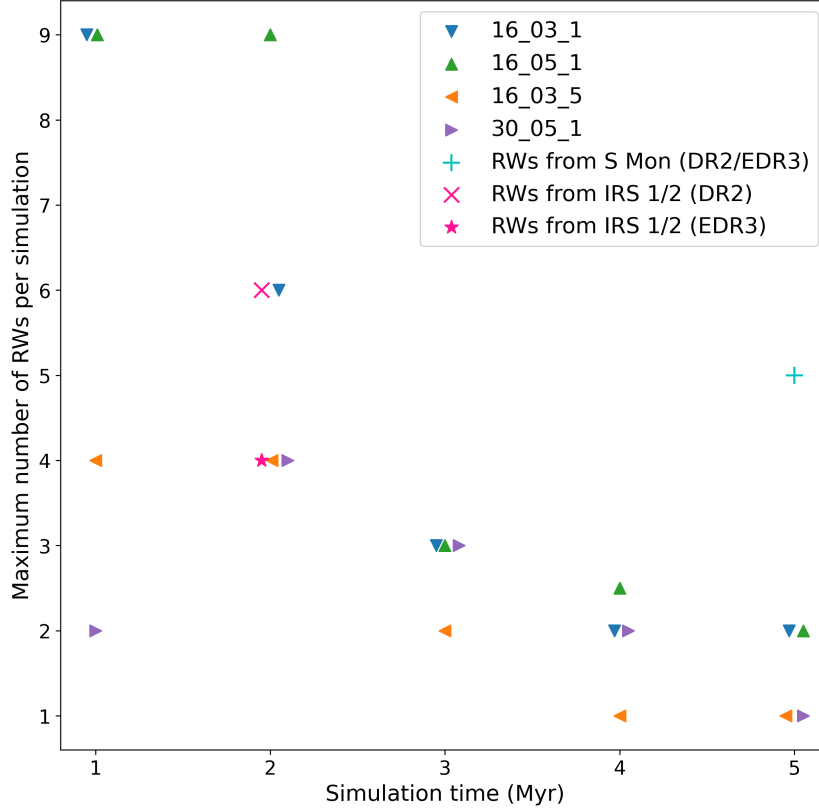


Figure 7.5: Maximum number of RWs for four selected initial conditions for simulation times from 1–5 Myr. The maximum number of RWs decreases at later times in the simulations and I find that the highest number of RWs is achieved in two different initial conditions (16-03-1 and 16-05-1). I also plot the observed numbers for S Mon (5 RWs) and IRS 1/2 (6/4 RWs) at their maximum ages used for the PARSEC isochrones (5/2 Myr).

Myr).

I find that only initial condition sets that feature a high amount of initial spatial substructure ( $D = 1.6$ ) can be fitted to the observations. Any simulation with a smooth initial spatial distribution, i.e.  $D = 3.0$ , or a moderate amount of initial spatial substructure, i.e.  $D = 2.0$  produces too few RWs to be consistent with the observations. For Simulation ID 16-03-1 ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.3$ , radius 1pc), I find a maximum of 5 RWs within 100 pc in my simulations up to an age of 2.5 Myr. When I consider the average number, the upper age at which this number of RWs is still present in the vicinity of S Mon reduces to 1.7 Myr. For Simulation ID 16-05-1 ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.5$ , radius 1pc), I find 5 RWs up to a maximum age of 2.6 Myr and when considering the average number, the maximum age of my simulations where the RWs/WWs are consistent with the observations reduces to 2.1 Myr.

The fastest RW that only traces back to S Mon (*Gaia* DR2 3326893624672681216) has a space velocity of  $\sim 62 \text{ km s}^{-1}$ . When I compare this velocity to the highest velocities reached in Simulation ID 16-03-1, the fastest star in my simulations has a similar velocity compared to this star between 2.5 Myr and 3.4 Myr. At 2 Myr, I find that the velocity of the fastest star within 100 pc in the simulations is double at  $\sim 124 \text{ km s}^{-1}$ . This  $\sim 1.0 M_{\odot}$  star leaves the 100 pc boundary shortly thereafter. After 3.4 Myr, a  $\sim 1.2 M_{\odot}$  star gets ejected at a velocity of  $\sim 87 \text{ km s}^{-1}$  and becomes the fastest star until it also leaves the 100 pc radius. At this point, the highest velocity across the simulations drops considerably to  $\sim 48 \text{ km s}^{-1}$ .

When comparing the fastest S Mon RW-star to the simulations with initial condition ID 16-05-1, I find that between 2.5–2.7 Myr, the velocity of the fastest star from all 20 simulations ( $\sim 66 \text{ km s}^{-1}$ ) is comparable to the observations. At most other times below 5 Myr, I find maximum velocities above this value (74–126  $\text{km s}^{-1}$ ) due to fast stars being continuously ejected and replacing those that leave the 100 pc region.

When I compare the averages of the highest velocities from all 20 simulations, I find much lower average-maximum velocities. For 16-03-1 at 1 Myr, the average of the highest velocities from all 20 simulations is  $\sim 76 \text{ km s}^{-1}$ , this average drops to  $\sim 53 \text{ km s}^{-1}$  at 2 Myr and to  $\sim 38 \text{ km s}^{-1}$  at 3 Myr. For 16-05-1 at 1 Myr, the average of the highest velocities is  $\sim 55 \text{ km s}^{-1}$ , after 2 Myr, this drops to  $\sim 44 \text{ km s}^{-1}$  and to  $\sim 35 \text{ km s}^{-1}$  after 3 Myr. Most of these velocities are below the velocity of the fastest star from S Mon.

Our upper age estimate for S Mon of less than 3 Myr is lower than that from most other studies. This lower age estimate is due to the high number of RWs that I can trace back to the region. I now briefly comment on the likelihood that all five RWs actually originated in S Mon. Two of the five candidates only trace back to the region when I consider the measurement errors. The first one is *Gaia* DR2 3131997187129420672, and it only traces back on the sky when considering the errors in proper motion. In addition it only becomes a 3D-candidate when considering the errors both in distance (error is  $\sim 25 \text{ pc}$ ) and RV. The second one is *Gaia* DR2 3132474680112352128, which once again only traces back on the sky when considering its proper motion errors. In 3D, it only traces back when considering either its distance or its RV error.

To confirm these candidates, I check the latest data from *Gaia* EDR3 with measurements covering a longer period than that in *Gaia* DR2 (Gaia Collaboration et al. 2021). This could possibly reduce the proper motion errors of these two stars and lead to them no longer tracing back to S Mon. In *Gaia* EDR3, there has been no update to the radial velocities. I use *Gaia* EDR3 to double-check all DR2 3D-candidates from my analysis.

Similar to my *Gaia* DR2 analysis, I do not use the parallaxes but Bayesian estimated photogeometric distances from [Bailer-Jones et al. \(2021\)](#) providing distances for  $\sim 1.35$  billion stars in *Gaia* EDR3. As *Gaia* EDR3 is missing extinction and reddening values for all stars, I cannot utilise *Gaia* EDR3 in the same way as DR2 to confirm stellar ages, so I have only used it to confirm the 3D trace-backs.

Both of the above mentioned error-only candidates still trace back to the region in 3D using *Gaia* EDR3. Also, the remaining three 3D RWs from S Mon continue to trace back to this region using the most recent kinematic data. So, after this check, I still have 5 RWs tracing back to S Mon, which fits to two sets of initial conditions in the simulations.

I have a fairly low number of traced back WWs from S Mon, considering that I have found five RWs. I compare these nine WWs to those predicted in simulations that fit the RW numbers (16-03-1 and 16-05-1). I find that the number of WWs is far below the average/maximum predicted by these simulations at all ages. However, of the 18 2D WW candidates with RV, over 70% trace back to S Mon in 3D. It is highly likely that when I measure RVs for the remaining 47 2D-candidates, I will be able to increase the number of traced back WWs.

Like for the RWs, I also have WW-candidates that only trace back on the sky given their velocity errors. These are *Gaia* DR2 3331597816450524288 and *Gaia* DR2 3327203588170236672. When considering the data from *Gaia* EDR3, I lose both of these stars, as well as *Gaia* DR2 3326740693772293248 and 3326713442204844160, which would leave me with five WWs. However, I can expect to identify additional 2D and possibly 3D WW candidates in *Gaia* EDR3 that do not trace back using *Gaia* DR2.

After comparing the *Gaia* DR2 RW/WWs to the simulations, I find that S Mon appears to be consistent with initial conditions that show a high level of initial substructure and either an initially subvirial ratio or are in virial equilibrium. In these two initial condition set-ups the initial stellar surface density would have been fairly high with  $\sim 3000$  stars  $\text{pc}^{-2}$  and an initial mass density of  $\sim 10\,000 M_{\odot} \text{pc}^{-3}$ , which would be similar to the initial density suggested for the ONC (see [Parker 2014](#), and references therein).

### 7.5.3 IRS 1 and IRS 2

While the centres of IRS 1 and IRS 2 are located around 10 pc from each other ([Kuhn et al. 2019](#)) in radial distance, the estimated ages for these subclusters are the same and my chosen boundaries of the regions overlap considerably. For the comparison to my simulations, I consider these two subclusters as one region. In my simulations, I find

several examples of a single star-forming region having evolved into two subclusters with centres located several pc from each other. I find a total of seven 3D RWs that I can trace back to either IRS 1 or IRS 2 and five 3D WWs, after I apply the approach for double trace-backs described in section 7.5.1. The number of 3D RWs further reduces to six RWs, when excluding the MS 3D RW, which has been suggested to have an age of 159 Myr (Gontcharov 2012).

Most of the RWs and WWs that trace back to both regions show very similar flight times, which is not surprising given their close proximity on the sky. This further highlights that the decision to treat IRS 1 and 2 as one initial star-forming region appears to be valid.

I use a PARSEC isochrone for IRS 1/2 with an age several Myr younger than S Mon (2 Myr instead of 5 Myr) and I trace back more RWs to IRS 1/2 than to S Mon. This result is fully in line with the predictions from simulations if these regions started from the same initial conditions. Due to the higher velocity of RWs, they leave my 100 pc search region much more quickly than WWs, so the older a region is the fewer RWs I can expect to find within a 100 pc radius.

My six RWs in the observations are only predicted in two of my initial conditions sets: ID 16-03-1 and 16-05-1. Both of these simulations start with a high amount of initial spatial substructure ( $D = 1.6$ ) and an initial radius of 1 pc. These regions can either be subvirial or in virial equilibrium.

For the initial conditions represented in 16-03-1 ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.3$ , radius 1pc), the maximum number of observed RWs is consistent with one of my simulations up to an age of 2.1 Myr. The average matches up to an age of 1.2 Myr. In the 16-05-1 ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.5$ , radius 1pc) simulations, the maximum number of RWs from the observations is consistent with one simulation up to an age of 2.4 Myr, which is above the age I chose for my isochrone (2 Myr). I can match the averages up to an age of 1.1 Myr.

The upper age estimates for these two initial conditions are higher than the upper end of the age estimates I find in literature and any reduction in RWs would further increase these age estimates. However, I can only match the RWs to a single simulation in each of the two initial condition sets at these high ages and also have individual simulations predicting this high number of RWs at minimum ages of 0.1 Myr for 16-03-1 and 0.2 Myr for 16-05-1.

The fastest star from IRS 1/2 is *Gaia* DR2 3132688088448708736 with a space velocity of  $\sim 88 \text{ km s}^{-1}$ . When comparing this velocity to the maximum reached in either of the simulations represented by 16-03-1 and 16-05-1 at early ages ( $< 2 \text{ Myr}$ ), the simulations

predict stars of even higher velocities within 100 pc. When comparing the average of the maximum velocities for all 20 simulations, for 16-03-1 I reach the highest value of  $\sim 86$   $\text{km s}^{-1}$  at simulation time 0.8 Myr. For 16-05-1, this average never reaches velocities over  $80 \text{ km s}^{-1}$ .

I now double-check the 3D RWs I find in *Gaia* DR2 against EDR3. Two of the 3D RW stars trace back to IRS 1/2 on the sky only when considering the proper motion errors. These two are *Gaia* DR2 3134140405870541568 (trace-back to IRS 1) and *Gaia* DR2 3132688088448708736 (trace-back to IRS 2). I lose the former of these stars when tracing back the stars with *Gaia* EDR3 but retain the latter. I also lose one further star *Gaia* DR2 3134176728413264896, reducing the total number of 3D RWs from six to four. In this *Gaia* EDR3 check, I also lose *Gaia* DR2 3326630811329448576, which is the MS-star with an age estimate of 159 Myr I previously already excluded.

Reducing the number of RWs from six to four increases the possible ages in Simulation IDs 16-03-1 and 16-05-1. For initial condition set 16-03-1, the maximum upper age from the simulation increases from 2.1 to 2.7 Myr and when comparing the average it goes from 1.2 Myr to 2.2 Myr. The lowest age, where I find a maximum of 4 RWs in these simulations remains at 0.1 Myr.

For initial condition ID 16-05-1, the upper age from the maximum increases from 2.4 to 2.9 Myr and the average from 1.1 Myr to 2.4 Myr. The lowest age, where I find a maximum of 4 RWs in these simulations is 0.1 Myr. This reduction of RWs to four also opens up two other possible initial conditions; however, only when fitting it to the maximum number of ejected stars as none of the averages fit to the observations.

Simulation ID 30-05-1 ( $D = 3.0$ ,  $\alpha_{\text{vir}} = 0.5$ , radius 1 pc) has one simulation with four RWs at ages of 1.1–2.1 Myr, whereas the average number of RWs in simulations is too low at any age. A second additional option for the initial conditions of IRS 1/2 is found in Simulation ID 16-03-5 ( $D = 1.6$ ,  $\alpha_{\text{vir}} = 0.3$ , radius 5 pc). Once again, I cannot match the average number of RWs to the observations. But I find that the maximum number from one of the simulations is consistent up to an age of 2.1 Myr and I find further simulations with this number of RWs down to an age of 0.1 Myr.

Fig. 7.5 shows the maximum number of RWs at five different times for these four possible initial conditions and I also plot the number of RWs I find in the observations. The maximum number of RWs decreases at later times in the simulations and I find that the highest number of RWs is achieved in two different initial condition sets (16-03-1 and 16-05-1). I also plot the observed numbers for IRS 1/2 (6 RWs) at their maximum age used for the PARSEC isochrones (2 Myr) and include the number of RWs after

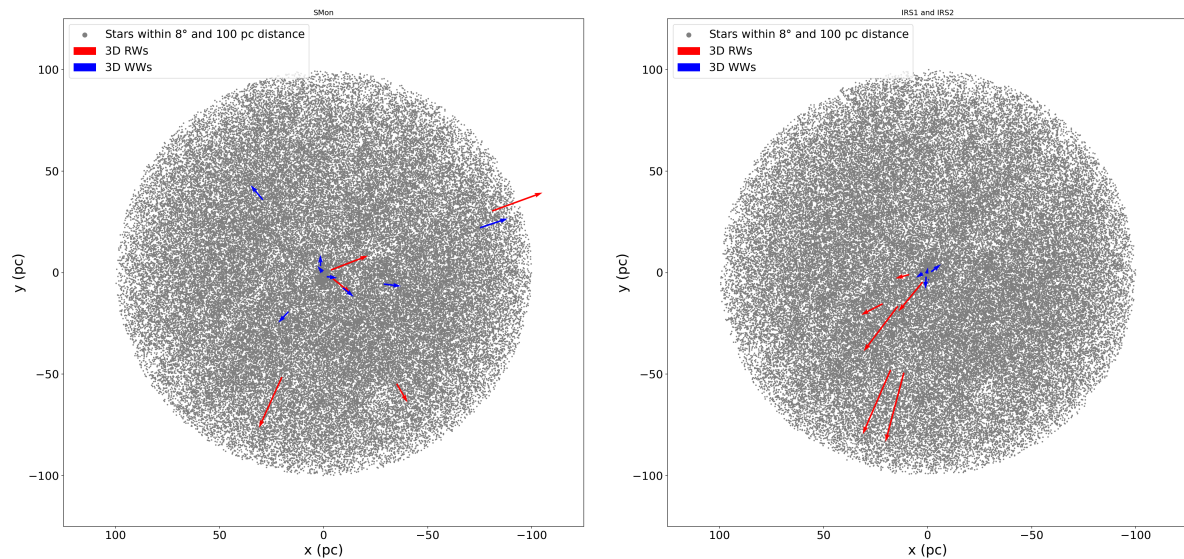


Figure 7.6: Location of the identified 3D RW and WW stars. Left: S Mon, right: IRS 1 and 2. The subclusters are located at the centres of the plot and extend to a radius of  $\sim 2$  pc. I invert the x-axis to replicate the orientation on the sky (i.e. decreasing right ascension from left to right). The RW stars are plotted in “red”, with the length of the arrows indicating their 2D-velocity in the respective rest frame, the WW stars are plotted in “blue”. Note: one 3D WW in each figure is not visible due to a very small 2D-velocity.

double-checking my results with *Gaia* EDR3 (4 RWs).

When considering the average age estimates with four RWs from 16-03-1 and 16-05-1, I find that these upper age estimates are far above those that other literature sources predict for these subclusters. However, I also have simulations predicting much lower ages, so both of these initial conditions are viable options. I now turn to the 3D WWs to evaluate if these provide any further insights.

I find five 3D WWs originating from IRS 1/2, which is much lower than the predicted numbers from simulations 16-03-1 and 16-05-1 at any age in my simulations. There are still a few 2D-candidates that are missing RVs; however, these are only 19 further candidates. If all of these 2D-candidates turned into 3D-candidates (i.e. 24 3D WWs), I would still not reach the maximum number of WWs predicted by these two initial conditions. However, I would be able to match the average number of WWs in both initial set-ups at all ages. So the number of WWs do not enable me to further constrain the initial conditions.

Using *Gaia* EDR3 to also check the 3D WWs, I lose one of these (*Gaia* DR2 3326693857153492736) taking me to four 3D WWs. I now match this number to the two additional initial conditions that open up after the *Gaia* EDR3 check of the 3D



RWs. For 30-05-1, this number of WWs only fits the maximum number at most ages; however, none of the averages. Any further WW discoveries in *Gaia* EDR3 or other sources would make this initial condition unsuitable for IRS 1/2. With initial condition 16-03-5, this number of WWs is still well below the maximum and below the averages; however, this gap would be much easier to close than for 16-03-1 and 16-05-1.

The number of ejected stars from S Mon and IRS 1/2 are consistent with both regions having formed from initial conditions represented by Simulation IDs 16-03-1 or 16-05-1. These simulations differ only by the initial virial ratio. If 16-05-1 was the true initial condition, IRS 1/2 could be a binary cluster (e.g. [Rozhavskii et al. 1976](#); [Pietrzynski & Udalski 2000](#); [de la Fuente Marcos & de la Fuente Marcos 2009](#); [Priyatikanto et al. 2016](#); [Zhong et al. 2019](#); [Bisht et al. 2021](#)), which is a common outcome in simulations with these initial conditions ([Arnold et al. 2017](#)). If 16-03-5 remained as a set of initial conditions consistent with the observations, then IRS 1/2 would have formed with a slightly different initial set-up than S Mon. Both regions would share their initially highly spatial substructure and possibly their initial subvirial radius with a different initial stellar density. The simulations with initial conditions represented by IDs 16-03-1 and 16-05-1 show initial stellar densities as high as  $\sim 3000$  stars  $\text{pc}^{-2}$  and a mass density of  $\sim 10\,000 M_{\odot} \text{pc}^{-3}$ , whereas simulations with initial condition ID 16-03-5 start with a lower initial stellar density of  $\sim 400$  stars  $\text{pc}^{-2}$  and a mass density of  $\sim 150 M_{\odot} \text{pc}^{-3}$ .

In past studies (e.g. [Fűrész et al. 2006](#); [Tobin et al. 2015a](#); [Costado & Alfaro 2018](#)), these authors found evidence of a clumpy (initial) spatial substructure and a correlation with the kinematic substructure. From my results presented here, I suggest that my results confirm these previous results using a different approach and provide avenues for further study. A study focused on the possible initial spatial substructure using alternative approaches such as mass segregation and the  $Q$ -parameter would allow this to be analysed further.

This analysis is presented in a companion study ([Parker & Schoettler 2022](#)). This study uses the same  $N$ -body simulations that I have used for my search for RW/WW stars and an observational mass census of  $\sim 750$  stars that I have compiled from literature sources. Applying the  $Q$ -parameter to the observational data to determine the spatial distribution shows that neither S Mon nor IRS 1/2 are substructured or centrally concentrated. There is a higher relative local surface density ratio  $\Sigma_{\text{LDR}}$  for the most massive stars in IRS 1/2. However, neither of the subclustered region exhibit mass segregation based on the  $\Lambda_{\text{MSR}}$  ratio. The analysis of the simulations in [Parker & Schoettler \(2022\)](#) shows that the initial conditions of NGC 2264 are consistent with the results from my

analysis using ejected stars, i.e. high initial density, substructured and subvirial.

Fig. 7.6 shows the location of the identified five 3D RWs and nine 3D WWs from S Mon on the left and the six 3D RWs and five 3D WWs from IRS 1/2, that I found based on my analysis of the *Gaia* DR2 data. I see clearly that the WWs from S Mon, which is older, have reached farther distances from the cluster than those ejected from IRS 1/2, which are still in close proximity of the cluster. Also there appear to be no ejected stars from IRS 1/2 that are on a NW-trajectory.

Future *Gaia* data releases and further RV measurements should provide me with more clarity thereby allowing me to better restrict the initial conditions of S Mon and IRS 1/2; however, this *Gaia* DR2 analysis has provided a framework of possible initial conditions for the subclusters that can be used for future analyses of new observational data.

## 7.6 Conclusions

In this chapter, I combine *Gaia* DR2 and EDR3 observations with predictions from  $N$ -body simulations to search for RW and WW stars from the three subclusters S Mon, IRS 1 and IRS 2 in NGC 2264 within a distance of 100 pc to constrain the initial conditions of these regions. The conclusions from my simulations and the search in *Gaia* DR2/EDR3 are summarised as follows:

- (i) I find five 3D RWs and nine 3D WWs within 100 pc of S Mon that I can trace back to this subcluster using *Gaia* DR2. All of these RWs and WWs are either low- or intermediate-mass stars. I find two 3D WWs, which have already reached the MS, all others are still pre-MS stars. When comparing the number of ejected stars to those predicted in  $N$ -body simulations, I find that S Mon appears to have evolved from initial conditions with an initial highly spatial substructure with a high stellar density ( $\sim 10\,000\text{ M}_{\odot}\text{ pc}^{-3}$ ) and either initially subvirial or virialized velocities.
- (ii) While I have searched separately for ejected stars from IRS 1 and IRS 2 in *Gaia* DR2, I treat these subclusters as one when comparing them to simulations, due to their similar age and considerable overlapping boundaries. For IRS 1/2, I trace back six 3D RWs and five 3D WWs still within 100 pc of these subclusters. As for S Mon, all of these stars are either low- or intermediate-mass stars. The higher number of 3D RWs suggests that this region is younger than S Mon, as the number

of RWs drops for the same initial condition with older ages. I find the same initial conditions fit to IRS 1/2 than to S Mon.

- (iii) I use *Gaia* DR2 for my analysis, as this release provides extinction and reddening values, which are not available in the more recent *Gaia* EDR3. This allowed me to use extinction/reddening corrected CAMDs to predict ages for my ejected stars. However, I use the more accurate *Gaia* EDR3 astrometry to check if my ejected stars still trace back given the more updated data. For S Mon, the number of 3D RWs does not change; however, two 3D RWs no longer trace back to IRS 1/2 in *Gaia* EDR3. With this lower number of four 3D RWs, two further sets of initial conditions become viable options for IRS 1/2. One set has smooth spatial substructure, low to moderate initial stellar density ( $\sim 150 M_{\odot} \text{pc}^{-3}$ ) and is initially virialized; the other is initially highly substructured, subvirial and with a low initial stellar density ( $\sim 70 M_{\odot} \text{pc}^{-3}$ ).
- (iv) For all subclusters, I only find a very low number of 3D WWs of nine or less, when some of the possible initial conditions predict 3–4 times that number. This is likely due to the lack of RVs available for my 2D-trace-backs and the number should increase with further RV availability.

## **Chapter 8**

### **Conclusions and future work**

In this thesis, I investigated if stars that were ejected during the early dynamical evolution of young star-forming regions can be used to constrain the initial density, spatial and kinematic substructure of their birth regions. My work has been published in four publications (see Appendix C).

## 8.1 Conclusions of the work presented in this thesis

### 8.1.1 Do unbound stars in $N$ -body simulations show differences depending on initial conditions?

In Chapter 4, I started with investigating if there are any differences in the number and velocities of unbound stars in  $N$ -body simulations of young star-forming regions. I set up simulations with a range of different initial spatial and velocity structures, which result in differences in the unbound population after 10 Myr. The unbound fractions of stars of different masses showed clear differences between the initial conditions, so did the numbers and velocities. The more subvirial a region was initially, the higher was the fraction of unbound high-mass stars compared to low/intermediate-mass stars in the vicinity of the region. High-mass stars that managed to escape their birth region often reached at least walkaway velocities. However, based on my simulations, not every young star-forming region would create a high-mass runaway or walkaway star. Most low/intermediate-mass stars left the regions at velocities just above the escape velocity, i.e. via evaporation, while others got ejected faster as a result of the DES.

In the simulations in Chapter 4, none of the stars got ejected via the BSS, as there was no stellar evolution and therefore no SNe to provide the initial velocity kick to the secondary companion. The fastest stars from my simulations were also low/intermediate-mass stars, which is in agreement with my analysis of the results from [Tetzlaff et al. \(2011\)](#) shown in Chapter 1.5. Low/intermediate-mass walkaway stars were ejected from the simulated star-forming regions in all initial condition sets. This number increased for regions that evolved more dynamically (more initial substructure and/or lower virial ratio). Based on these results, I expected to find at least a small number of these stars around virtually every young and high-density star-forming region.

### 8.1.2 Can ejected stars be used to constrain the initial conditions of the ONC?

In Chapter 5, I combined *Gaia* DR2 observations with predictions from  $N$ -body simulations to search for RW and WW stars from the ONC within a distance of 100 pc to constrain the region's initial conditions. For the simulations, I used initial conditions that were subvirial with a moderate amount of initial spatial substructure. I found a number of 3D RW ( $>30 \text{ km s}^{-1}$ ) and 3D WW ( $10\text{-}30 \text{ km s}^{-1}$ ) stars originating from the ONC in the low/intermediate-mass range ( $<8 M_{\odot}$ ). However, I did not find any high-mass stars ( $>8 M_{\odot}$ ) in either of the velocity ranges in all three dimensions. I traced back eight RWs that are still within my 100 pc search boundary.

My  $N$ -body simulations suggested that the older a star-forming region was, the fewer RWs were still found within this boundary. The number of RWs I found in my simulations agreed with those in the observations when the simulated regions had an age of  $\sim 1.6\text{-}2.5$  Myr. This age estimate for the ONC is in agreement with others from literature (e.g. [Da Rio et al. 2010](#); [Reggiani et al. 2011](#)). My results suggested that the ONC started initial subvirial with a moderate amount of substructure, which is in agreement with the results from other studies (e.g. [Allison et al. 2010](#); [Allison & Goodwin 2011](#)). My simulations predicted that all WWs were still to be found within the search region until at least 3 Myr and that the number of WWs increased up to this age. In *Gaia* DR2, I found 26 WWs all with masses between 0.3 and  $2.7 M_{\odot}$ . Twenty-four of those had been ejected within the past 2.5 Myr, which was the upper age estimate based on the number of RWs. This agreed with the average number of WWs at an age of  $\sim 1.5$  Myr from my simulations but was below the maximum from a single simulation at virtually any age.

### 8.1.3 Can star-disc systems experience more than one dense environment?

The work in Chapter 6 used the results from Chapter 5 to investigate if circumstellar discs could survive the ejection from young star-forming regions like the ONC. In my analysis, I used a list of ejected stars, past and future visitors to search for disc evidence found for these stars in other literature sources. I found that there were stars at RW and WW velocities that had been ejected from the ONC, which showed evidence of a disc. I also found a disc-bearing visitor from another star-forming region that was on approach to the ONC about to encounter a second higher density region. Finally, I found an older visitor that had just passed through the ONC and could possibly still have retained

some of its circumstellar disc. This suggested that a star-disc system could survive the ejection from their birth cluster and then possibly dynamically interact with stars in another dense star-forming region and emerge from this second region with an intact disc. Further work on this could provide more insights into the resilience of star-disc systems to external effects.

### 8.1.4 Can ejected stars be used to constrain the initial conditions of the NGC 2264?

In Chapter 7, I combined *Gaia* DR2 and also used EDR3 observations with predictions from  $N$ -body simulations to search for RW and WW stars from the three subclusters S Mon, IRS 1 and IRS 2 in NGC 2264. I used NGC 2264 as my second target for my work as there is less information from other literature sources available about its initial condition and so far, no confirmed RWs or WWs. For my simulations, I used eight different combinations of initial density, spatial and kinematic substructure. In *Gaia* DR2, I found five 3D-RWs and nine 3D-WWs within 100 pc of S Mon that I could trace back to this subcluster using *Gaia* DR2. All of these RWs and WWs were either low- or intermediate-mass stars. I found two 3D-WWs, which had already reached the MS, all others were still pre-MS stars. When comparing the number of ejected stars to those predicted in  $N$ -body simulations, I found that S Mon appeared to have evolved from initial conditions with an initial highly spatial substructure with a high stellar density ( $\sim 10\,000\ M_{\odot}\text{pc}^{-3}$ ) and either initially subvirial or virialized velocities. For IRS 1/2, I traced back six 3D-RWs and five 3D-WWs that were still within 100 pc of these subclusters. As for S Mon, all of these stars were either low- or intermediate-mass stars. The higher number of 3D-RWs suggested that this region was younger than S Mon, as the number of RWs dropped for the same initial condition with older ages, which was in agreement with results from other literature sources. I found the same initial conditions fit to IRS 1/2 than to S Mon.

I used the more accurate *Gaia* EDR3 astrometry to check if my ejected stars still traced back given the more updated data. For S Mon, the number of 3D-RWs did not change. However, two 3D-RWs no longer traced back to IRS 1/2 in *Gaia* EDR3. With this lower number of four 3D-RWs, two further sets of initial conditions became viable options for IRS 1/2. One set had smooth spatial substructure, low to moderate initial stellar density ( $\sim 150\ M_{\odot}\text{pc}^{-3}$ ) and was initially virialized; the other was initially highly substructured, subvirial and with a low initial stellar density ( $\sim 70\ M_{\odot}\text{pc}^{-3}$ ). For all subclusters, I only found a very low number of 3D-WWs of nine or less, when some of

the possible initial conditions predicted 3–4 times that number. This was likely due to the lack of RVs available for my 2D-trace-backs and the number should increase with further RV availability. I also searched for ejected star-disc systems and found several that appear to have originated in NGC 2264 and also one future visitor that still showed evidence of a disc.

### 8.1.5 Final remarks

The analysis presented in this thesis used ejected stars found in simulations and observations to constrain the initial conditions of young star-forming regions based on two example regions, the ONC and NGC 2264. The results show, that ejected stars can provide a method to constrain initial conditions that can be used alongside other methods, such as the  $Q$ -parameter and mass segregation measurements, as shown in [Parker & Schoettler \(2022\)](#) for NGC 2264. However, to use the RW method to its full extent requires further improvements in *Gaia* or other observations, e.g. more measurements of radial velocities in addition to proper motion; extinction and reddening values for a larger number of stars. For example, radial velocities for a large number of stars in a star-forming region can be measured using Integral Field Unit (IFU) techniques, where a large field of view is divided up into smaller units/cells/pixels. One of the IFU techniques used to gather radial velocities is Integral Field Spectroscopy, which is a combination of an IFU and a spectrograph (IFS)<sup>1</sup>. In IFS, a 2D-view of the sky (x-y or RA/Dec positions) is gathered in a datacube with spectra of each smaller unit (or pixel) providing information about the wavelength/velocity.

---

<sup>1</sup><https://www.eso.org/public/teles-instr/technology/ifu/> and <http://ifs.wikidot.com/what-is-ifs>



## 8.2 Outlook at potential future work

Binaries and higher order multiple systems are the dominant outcome of the star formation process and play a large role in the evolution of star-forming regions (e.g. [Duchêne & Kraus 2013](#)). What are the effects of different initial primordial binary populations on the formation of RW/WW stars? Do they enhance or suppress ejections compared to simulations that only contain dynamically formed binaries (see [Schoettler et al. 2019](#))? Are there differences in the ejected star populations when using different binary fractions to create the primordial binary population? What are the velocity limits of the runaway stars in each of the formation scenarios? Can the two standard scenarios convincingly explain faster hyper-runaways when taking different initial conditions for the origin cluster into account?

Star-forming regions do not exist in isolation but are subject to an external tidal field from the galaxies in which they reside. The regions are located at different distances to the galactic centres and experience different tidal forces. These tidal fields can accelerate the dissolution of the star-forming regions and cause more stars to escape from their birth environment. How much does an external tidal field affect the velocities of ejected stars? Does it increase the number of RW/WW stars ejected from a star-forming region? How much does the location of star-forming regions in a galaxy affect the ejected population?

*Gaia* DR3 (expected in the first half 2022) is expected to provide more accuracy for fainter sources (i.e. at larger distances), non-single stars and more radial velocity measurements amongst other improvements. My research has shown that we should find many low-mass ejected stars around star-forming regions ([Schoettler et al. 2019](#)). However, these low-mass stars will become fainter with higher uncertainties in the *Gaia* measurements the further away the regions are from Earth. Being able to use *Gaia* data for young star-forming regions that are at larger distances will allow us to increase the number of target regions, while not losing too many low-mass candidates at the fainter end of the magnitude spectrum.

# Bibliography

- ALMA Partnership et al., 2015, [ApJL](#), 808, L3
- Aarseth S. J., 1963, [MNRAS](#), 126, 223
- Aarseth S. J., 1999, [PASP](#), 111, 1333
- Aarseth S. J., 2003, *Gravitational N-Body Simulations*. Cambridge University Press
- Aarseth S. J., 2012, [MNRAS](#), 422, 841
- Adams F. C., Ruden S. P., Shu F. H., 1989, [ApJ](#), 347, 959
- Adams F. C., Hollenbach D., Laughlin G., Gorti U., 2004, [ApJ](#), 611, 360
- Adams F. C., Proszkow E. M., Fatuzzo M., Myers P. C., 2006, [ApJ](#), 641, 504
- Alexander R., Pascucci I., Andrews S., Armitage P., Cieza L., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, *Protostars and Planets VI*. p. 475, [doi:10.2458/azu\\_uapress\\_9780816531240-ch021](https://doi.org/10.2458/azu_uapress_9780816531240-ch021)
- Allen D. A., 1972, [ApJL](#), 172, L55
- Allison R. J., 2012, [MNRAS](#), 421, 3338
- Allison R. J., Goodwin S. P., 2011, [MNRAS](#), 415, 1967
- Allison R. J., Goodwin S. P., Parker R. J., Portegies Zwart S. F., de Grijs R., Kouwenhoven M. B. N., 2009, [MNRAS](#), 395, 1449
- Allison R. J., Goodwin S. P., Parker R. J., Portegies Zwart S. F., de Grijs R., 2010, [MNRAS](#), 407, 1098
- van Altena W. F., Lee J. T., Lee J. F., Lu P. K., Upgren A. R., 1988, [AJ](#), 95, 1744
- Anders P., Baumgardt H., Bissantz N., Portegies Zwart S. F., 2009, [MNRAS](#), 395, 2304
- Anders P., Baumgardt H., Gaburov E., Portegies Zwart S. F., 2012, [MNRAS](#), 421, 3557
- Anders F., et al., 2019, [A&A](#), 628, A94
- Andersson E. P., Renaud F., Agertz O., 2021, [MNRAS](#), 502, L29

- Andrae R., et al., 2018, [A&A](#), 616, A8
- Andre P., Ward-Thompson D., Barsony M., 1993, [ApJ](#), 406, 122
- Andre P., Ward-Thompson D., Barsony M., 2000, in Mannings V., Boss A. P., Russell S. S., eds, Protostars and Planets IV. p. 59 ([arXiv:astro-ph/9903284](#))
- André P., Di Francesco J., Ward-Thompson D., Inutsuka S. I., Pudritz R. E., Pineda J. E., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 27, [doi:10.2458/azu\\_uapress\\_9780816531240-ch002](#)
- Andrews S. M., 2020, [ARAA](#), 58, 483
- Andrews S. M., et al., 2018, [ApJL](#), 869, L41
- Arenou F., et al., 2018, [A&A](#), 616, A17
- Armitage P. J., Simon J. B., Martin R. G., 2013, [ApJL](#), 778, L14
- Arnold B., Goodwin S. P., Griffiths D. W., Parker R. J., 2017, [MNRAS](#), 471, 2498
- Athanassoula E., Fady E., Lambert J. C., Bosma A., 2000, [MNRAS](#), 314, 475
- Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, [AJ](#), 156, 58
- Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Demleitner M., Andrae R., 2021, [AJ](#), 161, 147
- Ballesteros-Paredes J., et al., 2020, [Space Science Reviews](#), 216, 76
- Bally J., 2016, [ARAA](#), 54, 491
- Bally J., Zinnecker H., 2005, [AJ](#), 129, 2281
- Bally J., Licht D., Smith N., Walawender J., 2006, [AJ](#), 131, 473
- Bally J., Youngblood A., Ginsburg A., 2012, [ApJ](#), 756, 137
- Bally J., Ginsburg A., Forbrich J., Vargas-González J., 2020, [ApJ](#), 889, 178
- Banerjee S., Kroupa P., 2015, [MNRAS](#), 447, 728
- Banerjee S., Kroupa P., Oh S., 2012, [MNRAS](#), 426, 1416
- Baratella M., et al., 2020, [A&A](#), 634, A34
- Barnes J., Hut P., 1986, [Nature](#), 324, 446
- Bastian N., Ercolano B., Gieles M., Rosolowsky E., Scheepmaker R. A., Gutermuth R., Efremov Y., 2007, [MNRAS](#), 379, 1302

- Bastian N., Gieles M., Ercolano B., Gutermuth R., 2009, [MNRAS](#), 392, 868
- Bastian N., Covey K. R., Meyer M. R., 2010, [ARAA](#), 48, 339
- Bate M. R., Bonnell I. A., Bromm V., 2002, [MNRAS](#), 332, L65
- Bate M. R., Bonnell I. A., Bromm V., 2003, [MNRAS](#), 339, 577
- Baxter E. J., Covey K. R., Muench A. A., Fűrész G., Rebull L., Szentgyorgyi A. H., 2009, [AJ](#), 138, 963
- Becklin E. E., Neugebauer G., 1967, [ApJ](#), 147, 799
- Bell C. P. M., Naylor T., Mayne N. J., Jeffries R. D., Littlefair S. P., 2013, [MNRAS](#), 434, 806
- Belokurov V., et al., 2020, [MNRAS](#), 496, 1922
- Benincasa S. M., et al., 2020, [MNRAS](#), 497, 3993
- Benisty M., et al., 2018, [A&A](#), 619, A171
- Biazzo K., Randich S., Palla F., 2011, [A&A](#), 525, A35
- Bigiel F., Leroy A., Walter F., Blitz L., Brinks E., de Blok W. J. G., Madore B., 2010, [AJ](#), 140, 1194
- Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press, Princeton, NJ USA
- Bischoff R., et al., 2020, [Astronomische Nachrichten](#), 341, 908
- Bisht D., Zhu Q., Yadav R. K. S., Ganesh S., Rangwal G., Durgapal A., Sariya D. P., Jiang I.-G., 2021, [MNRAS](#)
- Blaauw A., 1956, [PASP](#), 68, 495
- Blaauw A., 1961, Bulletin of the Astronomical Institutes of the Netherlands, 15, 265
- Blaauw A., 1964, [ARAA](#), 2, 213
- Blaauw A., 1993, in Cassinelli J. P., Churchwell E. B., eds, Astronomical Society of the Pacific Conference Series Vol. 35, Massive Stars: Their Lives in the Interstellar Medium. p. 207
- Blaauw A., Morgan W. W., 1953, Bulletin of the Astronomical Institutes of the Netherlands, 12, 76
- Blaauw A., Morgan W. W., 1954, [ApJ](#), 119, 625
- Bodensteiner J., Baade D., Greiner J., Langer N., 2018, [A&A](#), 618, A110

- Boersma J., 1961, Bulletin of the Astronomical Institutes of the Netherlands, [15](#), 291
- Boneberg D. M., Dale J. E., Girichidis P., Ercolano B., 2015, [MNRAS](#), 447, 1341
- Bonnell I. A., Bate M. R., 1994a, [MNRAS](#), 269, L45
- Bonnell I. A., Bate M. R., 1994b, [MNRAS](#), 271, 999
- Bonnell I. A., Bate M. R., 2002, [MNRAS](#), 336, 659
- Bonnell I. A., Bate M. R., 2005, [MNRAS](#), 362, 915
- Bonnell I. A., Bate M. R., Clarke C. J., Pringle J. E., 1997, [MNRAS](#), 285, 201
- Bonnell I. A., Bate M. R., Zinnecker H., 1998, [MNRAS](#), 298, 93
- Bonnell I. A., Smith K. W., Davies M. B., Horne K., 2001a, [MNRAS](#), 322, 859
- Bonnell I. A., Bate M. R., Clarke C. J., Pringle J. E., 2001b, [MNRAS](#), 323, 785
- Bonnell I. A., Clark P., Bate M. R., 2008, [MNRAS](#), 389, 1556
- Boss A. P., Keiser S. A., 2013, [ApJ](#), 764, 136
- Boubert D., Fraser M., Evans N. W., Green D. A., Izzard R. G., 2017, [A&A](#), 606, A14
- Boubert D., Guillochon J., Hawkins K., Ginsburg I., Evans N. W., Strader J., 2018, [MNRAS](#), 479, 2789
- Bragaglia A., Tosi M., 2006, [AJ](#), 131, 1544
- Bragança G. A., Daflon S., Cunha K., Bensby T., Oey M. S., Walth G., 2012, [AJ](#), 144, 130
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, [MNRAS](#), 427, 127
- Bressert E., et al., 2010, [MNRAS](#), 409, L54
- Bromley B. C., Kenyon S. J., Brown W. R., Geller M. J., 2009, The Astrophysical Journal, 706, 925
- Bromley B. C., Kenyon S. J., Brown W. R., Geller M. J., 2018, [ApJ](#), 868, 25
- Brown W. R., 2015, Annual Review of Astronomy and Astrophysics, 53
- Brown W. R., Geller M. J., Kenyon S. J., Kurtz M. J., 2005, [ApJL](#), 622, L33
- Brown W. R., Geller M. J., Kenyon S. J., Kurtz M. J., Bromley B. C., 2007a, [ApJ](#), 660, 311

- Brown W. R., Geller M. J., Kenyon S. J., Kurtz M. J., Bromley B. C., 2007b, *ApJ*, 671, 1708
- Brown W. R., Lattanzi M. G., Kenyon S. J., Geller M. J., 2018, *ApJ*, 866, 39
- Buckner A. S. M., et al., 2019, *A&A*, 622, A184
- Buckner A. S. M., et al., 2020, *A&A*, 636, A80
- Burgers J., 1939, *Trans. Roy. Neth. Acad. Sci. Amsterdam*, 17, 1
- Burkert A., Bodenheimer P., 1993, *MNRAS*, 264, 798
- Burrows A., Vartanyan D., 2021, *Nature*, 589, 29
- Busso G., et al., 2018, Gaia DR2 documentation Chapter 5 Photometry; Gaia Data Processing and Analysis Consortium, European Space Agency, <https://gea.esac.esa.int/archive/documentation/GDR2>
- Cannon A. J., Pickering E. C., 1993, VizieR Online Data Catalog, [p. III/135A](#)
- Cantat-Gaudin T., Anders F., 2020, *A&A*, 633, A99
- Cantat-Gaudin T., et al., 2018, *A&A*, 618, A93
- Carpenter J. M., Hodapp K. W., 2008, in Reipurth B., ed., ASP Monograph Publications Vol. 4, Handbook of Star Forming Regions, Volume I: The Northern Sky. Astronomical Society of the Pacific, p. 899
- Carrasco J. M., et al., 2016, *A&A*, 595, A7
- Carroll J. A., 1928, *MNRAS*, 88, 548
- Carroll J. A., 1933, *MNRAS*, 93, 478
- Cartwright A., Whitworth A. P., 2004, *MNRAS*, 348, 589
- Cartwright A., Whitworth A. P., 2009, *MNRAS*, 392, 341
- Castelaz M. W., Grasdalen G., 1988, *ApJ*, 335, 150
- Chabrier G., 2001, *ApJ*, 554, 1274
- Chabrier G., 2003a, *PASP*, 115, 763
- Chabrier G., 2003b, *ApJL*, 586, L133
- Chabrier G., 2005, in Corbelli E., Palla F., Zinnecker H., eds, Astrophysics and Space Science Library Vol. 327, The Initial Mass Function 50 Years Later. p. 41, [doi:10.1007/978-1-4020-3407-7\\_5](https://doi.org/10.1007/978-1-4020-3407-7_5)

- Chatterjee S., et al., 2005, [ApJL](#), 630, L61
- Chen H., Myers P. C., Ladd E. F., Wood D. O. S., 1995, [ApJ](#), 445, 377
- Cheng S., Yu H.-R., Inman D., Liao Q., Wu Q., Lin J., 2020, arXiv e-prints, p. [arXiv:2003.03931](#)
- Chevance M., et al., 2020a, [Space Science Reviews](#), 216, 50
- Chevance M., et al., 2020b, [MNRAS](#), 493, 2872
- Chick W. T., et al., 2020, [ApJS](#), 251, 29
- Cieza L., Baliber N., 2007, [ApJ](#), 671, 605
- Cieza L., et al., 2007, [ApJ](#), 667, 308
- Cieza L. A., et al., 2013, [ApJ](#), 762, 100
- Clark P. C., Bonnell I. A., Zinnecker H., Bate M. R., 2005, [MNRAS](#), 359, 809
- Clark P. C., Klessen R. S., Bonnell I. A., 2007, [MNRAS](#), 379, 57
- Clarke C. J., 2015, in Meyer M. R., Eyer L., Bell C. P., eds, Dynamics of Young Star Clusters and Associations: Saas-Fee Advanced Course, vol. 42. Swiss Society for Astrophysics and Astronomy. Springer, Berlin, Heidelberg, p. 73, [doi:10.1007/978-3-662-47290-3\\_6](#)
- Clarke C. J., Gendrin A., Sotomayor M., 2001, [MNRAS](#), 328, 485
- Cohn H., 1979, [ApJ](#), 234, 1036
- Cohn H., 1980, [ApJ](#), 242, 765
- Collins George W. I., Truax R. J., 1995, [ApJ](#), 439, 860
- Concha-Ramírez F., Wilhelm M. J. C., Portegies Zwart S. F., van Terwisga S. E., Hacar A., 2021, [MNRAS](#), 501, 1782
- Costado M. T., Alfaro E. J., 2018, [MNRAS](#), 476, 3160
- Cottaar M., et al., 2015, [ApJ](#), 807, 27
- Crowther P. A., Schnurr O., Hirschi R., Yusof N., Parker R. J., Goodwin S. P., Kassim H. A., 2010, [MNRAS](#), 408, 731
- Cruzalèbes P., et al., 2019, [MNRAS](#), 490, 3158
- Cuello N., et al., 2019, [MNRAS](#), 483, 4114
- Cuello N., et al., 2020, [MNRAS](#), 491, 504

- Da Rio N., Robberto M., Soderblom D. R., Panagia N., Hillenbrand L. A., Palla F., Stassun K. G., 2010, [ApJ](#), 722, 1092
- Da Rio N., Robberto M., Hillenbrand L. A., Henning T., Stassun K. G., 2012, [ApJ](#), 748, 14
- Da Rio N., et al., 2016, [ApJ](#), 818, 59
- Da Rio N., et al., 2017, [ApJ](#), 845, 105
- Daemgen S., Correia S., Petr-Gotzens M. G., 2012, [A&A](#), 540, A46
- Dahm S. E., 2008, in Reipurth B., ed., ASP Monograph Publications Vol. 4, Handbook of Star Forming Regions, Volume I: The Northern Sky. p. 966
- Dahm S. E., Simon T., Proszkow E. M., Patten B. M., 2007, [AJ](#), 134, 999
- Dale J. E., Ngoumou J., Ercolano B., Bonnell I. A., 2014, [MNRAS](#), 442, 694
- Das H., Deb S., Baruah A., 2021, [ApJ](#), 911, 83
- Dehnen W., 2001, [MNRAS](#), 324, 273
- Dehnen W., Read J. I., 2011, [European Physical Journal Plus](#), 126, 55
- Dias W. S., Alessi B. S., Moitinho A., Lépine J. R. D., 2002, [A&A](#), 389, 871
- Dias W. S., Alessi B. S., Moitinho A., Lepine J. R. D., 2012, VizieR Online Data Catalog, p. B/ocl
- Dorigo Jones J., Oey M. S., Pagneot K., Castro N., Moe M., 2020, [ApJ](#), 903, 43
- Drew J. E., Greimel R., Irwin M. J., Sale S. E., 2008, [MNRAS](#), 386, 1761
- Drew J. E., Herrero A., Mohr-Smith M., Monguió M., Wright N. J., Kupfer T., Napiwotzki R., 2018, [MNRAS](#), 480, 2109
- Duchêne G., Kraus A., 2013, [ARAA](#), 51, 269
- Duchêne G., Simon T., Eisloffel J., Bouvier J., 2001, [A&A](#), 379, 147
- Duchêne G., Lacour S., Moraux E., Goodwin S., Bouvier J., 2018, [MNRAS](#), 478, 1825
- Duflot M., Figon P., Meyssonier N., 1995, [A&A Suppl.](#), 114, 269
- Dunham M. M., et al., 2015, [ApJS](#), 220, 11
- Dyer C. C., Ip P. S. S., 1993, [ApJ](#), 409, 60
- Edelmann H., Napiwotzki R., Heber U., Christlieb N., Reimers D., 2005, [ApJL](#), 634, L181



- Eggleton P. P., Fitchett M. J., Tout C. A., 1989, [ApJ](#), 347, 998
- Eggleton P. P., Fitchett M. J., Tout C. A., 1990, [ApJ](#), 354, 387
- Eldridge J. J., Langer N., Tout C. A., 2011, [MNRAS](#), 414, 3501
- Elmegreen B. G., 2000, [ApJ](#), 530, 277
- Elmegreen B. G., 2007, [ApJ](#), 668, 1064
- Evans Neal J. I., et al., 2009, [ApJS](#), 181, 321
- Evans D. W., et al., 2018, [A&A](#), 616, A4
- Evans F. A., Renzo M., Rossi E. M., 2020, [MNRAS](#), 497, 5344
- Fűrész G., et al., 2006, [ApJ](#), 648, 1090
- Fűrész G., Hartmann L. W., Megeath S. T., Szentgyorgyi A. H., Hamden E. T., 2008, [ApJ](#), 676, 1109
- Fang M., Kim J. S., van Boekel R., Sicilia-Aguilar A., Henning T., Flaherty K., 2013, [ApJS](#), 207, 5
- Farias J. P., Tan J. C., 2018, [A&A](#), 612, L7
- Farias J. P., Tan J. C., Chatterjee S., 2019, [MNRAS](#), 483, 4999
- Farias J. P., Tan J. C., Eyer L., 2020, [ApJ](#), 900, 14
- Fehrenbach C., Burnage R., Figuiere J., 1992, *A&A Suppl.*, **95**, 541
- Fenske J., Arakawa J., Fallscheer C., Francesco J. D., 2021, [AJ](#), 161, 156
- Foster J. B., et al., 2015, [ApJ](#), 799, 136
- di Francesco J., Evans N. J. I., Caselli P., Myers P. C., Shirley Y., Aikawa Y., Tafalla M., 2007, in Reipurth B., Jewitt D., Keil K., eds, *Protostars and Planets V*. p. 17 ([arXiv:astro-ph/0602379](#))
- Fregeau J. M., Cheung P., Portegies Zwart S. F., Rasio F. A., 2004, [MNRAS](#), 352, 1
- de la Fuente Marcos R., de la Fuente Marcos C., 2005, [New Astronomy](#), 10, 551
- de la Fuente Marcos R., de la Fuente Marcos C., 2009, [A&A](#), 500, L13
- de la Fuente Marcos R., de la Fuente Marcos C., 2018, [Research Notes of the American Astronomical Society](#), 2, 35
- Fujii M. S., Portegies Zwart S., 2011, [Science](#), 334, 1380

- Funato Y., Makino J., Ebisuzaki T., 1992, Publications of the Astronomical Society of Japan, [44](#), 613
- Gaburov E., Harfst S., Portegies Zwart S., 2009, [New Astronomy](#), 14, 630
- Gagné J., Faherty J. K., Mamajek E. E., 2018, [ApJ](#), 865, 136
- Gagné J., David T. J., Mamajek E. E., Mann A. W., Faherty J. K., Bédard A., 2020, [ApJ](#), 903, 96
- Gaia Collaboration et al., 2016a, [A&A](#), 595, A1
- Gaia Collaboration et al., 2016b, [A&A](#), 595, A2
- Gaia Collaboration et al., 2018a, [A&A](#), 616, A1
- Gaia Collaboration et al., 2018b, [A&A](#), 616, A10
- Gaia Collaboration et al., 2021, [A&A](#), 649, A1
- Galli P. A. B., Bertout C., Teixeira R., Ducourant C., 2015, [A&A](#), 580, A26
- Gao Y., Li J., Jia S., 2019, [MNRAS](#), 487, 3178
- Getman K. V., et al., 2014, [ApJ](#), 787, 108
- Getman K. V., Feigelson E. D., Kuhn M. A., Garmire G. P., 2019, [MNRAS](#), 487, 2977
- Giersz M., 1998, [MNRAS](#), 298, 1239
- Glover S. C. O., Mac Low M. M., 2011, [MNRAS](#), 412, 337
- Goddi C., Humphreys E. M. L., Greenhill L. J., Chandler C. J., Matthews L. D., 2011, [ApJ](#), 728, 15
- Gomez M., Hartmann L., Kenyon S. J., Hewett R., 1993, [AJ](#), 105, 1927
- Gontcharov G. A., 2006, [Astronomy Letters](#), 32, 759
- Gontcharov G. A., 2012, [Astronomy Letters](#), 38, 694
- Goodwin S. P., 1997, [MNRAS](#), 284, 785
- Goodwin S. P., Bastian N., 2006, [MNRAS](#), 373, 752
- Goodwin S. P., Whitworth A. P., 2004, [A&A](#), 413, 929
- Goodwin S. P., Whitworth A. P., Ward-Thompson D., 2004, [A&A](#), 423, 169
- Gott J. Richard I., Gunn J. E., Ostriker J. P., 1970, [ApJL](#), 160, L91
- Gouliermis D. A., Hony S., Klessen R. S., 2014, [MNRAS](#), 439, 3775

- Gower J. C., Ross G. J. S., 1969, [Journal of the Royal Statistical Society: Series C \(Applied Statistics\)](#), 18, 54
- Greengard L., Rokhlin V., 1987, [Journal of Computational Physics](#), 73, 325
- Greengard L., Rokhlin V., 1997, [Acta Numerica](#), 6, 229
- Greenstein J. L., 1957, [AJ](#), 62, 16
- de Grijs R., 2010, [Philosophical Transactions of the Royal Society of London Series A](#), 368, 693
- Großschedl J. E., et al., 2018, [A&A](#), 619, A106
- Großschedl J. E., et al., 2019, [A&A](#), 622, A149
- Grudić M. Y., Guszejnov D., Hopkins P. F., Lamberts A., Boylan-Kolchin M., Murray N., Schmitz D., 2018, [MNRAS](#), 481, 688
- Guarcello M. G., et al., 2016, arXiv e-prints, p. [arXiv:1605.01773](#)
- Gutermuth R. A., Megeath S. T., Myers P. C., Allen L. E., Pipher J. L., Fazio G. G., 2009, [ApJS](#), 184, 18
- Guzmán-Díaz J., et al., 2021, [A&A](#), 650, A182
- Gvaramadze V. V., 2020, [Research Notes of the American Astronomical Society](#), 4, 213
- Gvaramadze V. V., Gualandris A., Portegies Zwart S., 2009, [MNRAS](#), 396, 570
- Gvaramadze V. V., Weidner C., Kroupa P., Pflamm-Altenburg J., 2012, [MNRAS](#), 424, 3037
- Gvaramadze V. V., Maryeva O. V., Kniazev A. Y., Alexashov D. B., Castro N., Langer N., Katkov I. Y., 2019, [MNRAS](#), 482, 4408
- Haisch Karl E. J., Lada E. A., Lada C. J., 2001, [ApJL](#), 553, L153
- Hambly N., et al., 2018, Gaia DR2 documentation Chapter 14: Datamodel description; Gaia Data Processing and Analysis Consortium, European Space Agency, <https://gea.esac.esa.int/archive/documentation/GDR2/>
- Hard R., Rose H., 2004, The Routledge Handbook of Greek Mythology: Based on H.J. Rose's "Handbook of Greek Mythology". Routledge, <https://books.google.co.uk/books?id=r1Y3xZWVlnIC>
- Hardy A., et al., 2015, [A&A](#), 583, A66
- Harfst S., Gualandris A., Merritt D., Spurzem R., Portegies Zwart S., Berczik P., 2007, [New Astronomy](#), 12, 357

- Hartmann L., Calvet N., Gullbring E., D'Alessio P., 1998, *ApJ*, 495, 385
- Hartmann L., Herczeg G., Calvet N., 2016, *ARAA*, 54, 135
- Haworth T. J., Clarke C. J., Rahman W., Winter A. J., Facchini S., 2018, *MNRAS*, 481, 452
- Hayashi C., 1961, Publications of the Astronomical Society of Japan, 13, 450
- Heber U., Edelman H., Napiwotzki R., Altmann M., Scholz R. D., 2008, *A&A*, 483, L21
- Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, *ApJ*, 591, 288
- Heggie D. C., 1975, *MNRAS*, 173, 729
- Heggie D. C., 1977, Comments on Astrophysics, 7, 43
- Heggie D., Hut P., 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics. Cambridge University Press
- Heggie D. C., Mathieu R. D., 1986, in Hut P., McMillan S. L. W., eds, Proceedings of a Workshop Held at the Institute for Advanced Study Vol. 267, The Use of Supercomputers in Stellar Dynamics. Springer Verlag, p. 233, doi:10.1007/BFb0116419
- Heggie D. C., Hut P., McMillan S. L. W., 1996, *ApJ*, 467, 359
- Heiter U., Soubiran C., Netopil M., Paunzen E., 2014, *A&A*, 561, A93
- Hennebelle P., Chabrier G., 2008, *ApJ*, 684, 395
- Hénon M. H., 1971, *Ap&SS*, 14, 151
- Hénon M., 1975, in Hayli A., ed., Proceedings from IAU Symposium Vol. 69, Dynamics of the Solar Systems. p. 133
- Herbig G. H., 1954, *ApJ*, 119, 483
- Herczeg G. J., Linsky J. L., Walter F. M., Gahm G. F., Johns-Krull C. M., 2006, *ApJS*, 165, 256
- Herpin F., et al., 2012, *A&A*, 542, A76
- Herschel W., 1803, Philosophical Transactions of the Royal Society of London Series I, 93, 339
- Hester J. J., et al., 1996, *AJ*, 111, 2349
- Hetem A., Gregorio-Hetem J., 2019, *MNRAS*, 490, 2521
- Hillenbrand L. A., 1997, *AJ*, 113, 1733

- Hillenbrand L. A., Hartmann L. W., 1998, [ApJ](#), 492, 540
- Hillenbrand L. A., Strom S. E., Calvet N., Merrill K. M., Gatley I., Makidon R. B., Meyer M. R., Skrutskie M. F., 1998, [AJ](#), 116, 1816
- Hills J. G., 1975a, [AJ](#), 80, 809
- Hills J. G., 1975b, [AJ](#), 80, 1075
- Hills J. G., 1988, [Nature](#), 331, 687
- Hintze J. L., Nelson R. D., 1998, [The American Statistician](#), 52, 181
- Hirsch H. A., Heber U., O'Toole S. J., Bresolin F., 2005, [A&A](#), 444, L61
- Hobbs G., Lorimer D. R., Lyne A. G., Kramer M., 2005, [MNRAS](#), 360, 974
- von Hoerner S., 1963, *Zeitschrift für Astrophysik*, [57](#), 47
- Hoffer J. B., 1983, [AJ](#), 88, 1420
- Höflich P., Kumar P., Wheeler J. C., 2011, *Cosmic Explosions in Three Dimensions*. Cambridge University Press
- Hohle M. M., Neuhäuser R., Schutz B. F., 2010, [Astronomische Nachrichten](#), 331, 349
- Hoogerwerf R., de Bruijne J. H. J., de Zeeuw P. T., 2001, [A&A](#), 365, 49
- Hopkins P. F., 2012, [MNRAS](#), 423, 2037
- Houk N., Smith-Moore M., 1988, *Michigan Catalogue of Two-dimensional Spectral Types for the HD Stars. Volume 4, Declinations -26°.0 to -12°.0..* Department of Astronomy, University of Michigan, Ann Arbor
- Houk N., Swift C., 1999, *Michigan Spectral Survey*, [5](#), 0
- Hsu W.-H., Hartmann L., Allen L., Hernández J., Megeath S. T., Mosby G., Tobin J. J., Espaillat C., 2012, [ApJ](#), 752, 59
- Hsu W.-H., Hartmann L., Allen L., Hernández J., Megeath S. T., Tobin J. J., Ingleby L., 2013, [ApJ](#), 764, 114
- Hueso R., Guillot T., 2005, [A&A](#), 442, 703
- Humason M. L., Zwicky F., 1947, [ApJ](#), [105](#), 85
- Hurley J. R., Pols O. R., Tout C. A., 2000, [MNRAS](#), 315, 543
- Hut P., 1983, [ApJ](#), 268, 342
- Hut P., 1993, [ApJ](#), 403, 256

- Hut P., 2003, [Symposium - International Astronomical Union](#), 208, 331
- Hut P., Bahcall J. N., 1983, [ApJ](#), 268, 319
- Ibeling D., Heger A., 2013, [ApJL](#), 765, L43
- Iben Icko J., Tutukov A. V., 1996, [ApJ](#), 456, 738
- Irrgang A., Kreuzer S., Heber U., 2018, [A&A](#), 620, A48
- Irrgang A., Dimpel M., Heber U., Raddi R., 2021, [A&A](#), 646, L4
- Isequilla N. L., Ortega M. E., Areal M. B., Paron S., 2021, [A&A](#), 649, A139
- Jackson R. J., et al., 2016, [A&A](#), 586, A52
- Jackson R. J., et al., 2020, [MNRAS](#), 496, 4701
- Jaehnig K. O., Da Rio N., Tan J. C., 2015, [ApJ](#), 798, 126
- Janson M., et al., 2012a, [ApJ](#), 754, 44
- Janson M., Jayawardhana R., Girard J. H., Lafrenière D., Bonavita M., Gizis J., Brandeker A., 2012b, [ApJL](#), 758, L2
- Jeans J. H., 1902, [Philosophical Transactions of the Royal Society of London Series A](#), 199, 1
- Jensen E. L. N., Akeson R., 2014, [Nature](#), 511, 567
- Jilinski E., Ortega V. G., Drake N. A., de la Reza R., 2010, [ApJ](#), 721, 469
- Johnston K. G., et al., 2015, [ApJL](#), 813, L19
- Johnstone D., Hollenbach D., Bally J., 1998, [ApJ](#), 499, 758
- Juhász A., et al., 2010, [ApJ](#), 721, 431
- Justham S., Wolf C., Podsiadlowski P., Han Z., 2009, [A&A](#), 493, 1081
- Karlsson B., 1972, [A&A Suppl.](#), 7, 35
- Katz D., et al., 2019, [A&A](#), 622, A205
- Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., Scholz R. D., 2005, [A&A](#), 438, 1163
- Kim K. H., et al., 2013, [ApJ](#), 769, 149
- Kim D., Lu J. R., Konopacky Q., Chu L., Toller E., Anderson J., Theissen C. A., Morris M. R., 2019, [AJ](#), 157, 109

- King J. R., Soderblom D. R., Fischer D., Jones B. F., 2000, [ApJ](#), 533, 944
- King R. R., Parker R. J., Patience J., Goodwin S. P., 2012, [MNRAS](#), 421, 2025
- Klagyivik P., et al., 2013, [ApJ](#), 773, 54
- Klassen M., Pudritz R. E., Kuiper R., Peters T., Banerjee R., 2016, [ApJ](#), 823, 28
- Köhler R., Petr-Gotzens M. G., McCaughrean M. J., Bouvier J., Duchêne G., Quirrenbach A., Zinnecker H., 2006, [A&A](#), 458, 461
- Kolmogorov A., 1941, *Akademiia Nauk SSSR Doklady*, **30**, 301
- Könyves V., et al., 2015, [A&A](#), 584, A91
- Korchagin V. I., Girard T. M., Borkova T. V., Dinescu D. I., van Altena W. F., 2003, [AJ](#), 126, 2896
- Kounkel M., Hartmann L., Tobin J. J., Mateo M., Bailey John I. I., Spencer M., 2016, [ApJ](#), 821, 8
- Kounkel M., et al., 2017, [ApJ](#), 834, 142
- Kounkel M., et al., 2018, [AJ](#), 156, 84
- Kounkel M., et al., 2019, [AJ](#), 157, 196
- Kouwenhoven M. B. N., Brown A. G. A., Portegies Zwart S. F., Kaper L., 2007, [A&A](#), 474, 77
- Kraus S., et al., 2007, [A&A](#), 466, 649
- Kraus A. L., Ireland M. J., Hillenbrand L. A., Martinache F., 2012, [ApJ](#), 745, 19
- Krause M. G. H., et al., 2020, [Space Science Reviews](#), 216, 64
- Kroupa P., 1995, [MNRAS](#), 277, 1491
- Kroupa P., 2001, [MNRAS](#), 322, 231
- Kroupa P., 2002, [Science](#), 295, 82
- Kroupa P., 2011, in *Stellar Clusters & Associations: A RIA Workshop on Gaia*. pp 17–27 ([arXiv:1111.5613](#))
- Kroupa P., Jeřábková T., 2019, [Nature Astronomy](#), 3, 482
- Kroupa P., Aarseth S., Hurley J., 2001, [MNRAS](#), 321, 699
- Kroupa P., Jeřábková T., Dinnbier F., Beccari G., Yan Z., 2018, [A&A](#), 612, A74
- Kruijssen J. M. D., 2012, [MNRAS](#), 426, 3008

- Kruijssen J. M. D., Bastian N., 2016, [MNRAS](#), 457, L24
- Krumholz M. R., 2015, in Vink J. S., ed., *Astrophysics and Space Science Library* Vol. 412, *Very Massive Stars in the Local Universe*. Springer, Cham, p. 43, [doi:10.1007/978-3-319-09596-7\\_3](https://doi.org/10.1007/978-3-319-09596-7_3)
- Kuhn M. A., Hillenbrand L. A., Sills A., Feigelson E. D., Getman K. V., 2019, [ApJ](#), 870, 32
- Kumar M. S. N., Schmeja S., 2007, [A&A](#), 471, L33
- Kunder A., et al., 2017, [AJ](#), 153, 75
- Lada C. J., 1987, in Peimbert M., Jugaku J., eds, *IAU Symposium* Vol. 115, *Star Forming Regions*. p. 1
- Lada C. J., Lada E. A., 2003, [ARAA](#), 41, 57
- Lada C. J., Margulis M., Dearborn D., 1984, [ApJ](#), 285, 141
- Lada E. A., Depoy D. L., Evans Neal J. I., Gatley I., 1991, [ApJ](#), 371, 171
- Lada C. J., Young E. T., Greene T. P., 1993, [ApJ](#), 408, 471
- Lai D., 2001, in Blaschke D., Glendenning N. K., Sedrakian A., eds, *Lecture Notes in Physics*, Vol. 578, *Physics of Neutron Star Interiors*. Springer, Berlin, Heidelberg, p. 424, [doi:10.1007/3-540-44578-1](https://doi.org/10.1007/3-540-44578-1)
- Lamm M. H., Bailer-Jones C. A. L., Mundt R., Herbst W., Scholz A., 2004, [A&A](#), 417, 557
- Larson R. B., 1981, [MNRAS](#), 194, 809
- Larson R. B., 1995, [MNRAS](#), 272, 213
- van Leeuwen F., et al., 2017, [A&A](#), 599, A32
- Lehmann H., Vitrichenko E., Bychkov V., Bychkova L., Klochkova V., 2010, [A&A](#), 514, A34
- Lennon D. J., et al., 2018, [A&A](#), 619, A78
- Leonard P. J. T., 1990, *J. R. Astron. Soc. Canada*, 84, 216
- Leonard P. J. T., 1991, [AJ](#), 101, 562
- Leonard P. J. T., 1995, [MNRAS](#), 277, 1080
- Leonard P. J. T., Duncan M. J., 1990, [AJ](#), 99, 608
- Li Y.-B., et al., 2021, [ApJS](#), 252, 3



- Lindgren L., 2018, Re-normalising the astrometric chi-square in Gaia DR2, GAIA-C3-TN-LU-LL-124, [http://www.rssd.esa.int/doc\\_fetch.php?id=3757412](http://www.rssd.esa.int/doc_fetch.php?id=3757412)
- Lindgren L., Bastian U., 2010, in EAS Publications Series. pp 109–114, [doi:10.1051/eas/1045018](https://doi.org/10.1051/eas/1045018)
- Lindgren L., et al., 2018, *A&A*, 616, A2
- Liu T., Wu Y., Wu J., Qin S.-L., Zhang H., 2013, *MNRAS*, 436, 1335
- Lodieu N., Smart R. L., Pérez-Garrido A., Silvotti R., 2019, *A&A*, 623, A35
- Lombardi M., Bouy H., Alves J., Lada C. J., 2014, *A&A*, 566, A45
- Longmore S. N., et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 291, [doi:10.2458/azu\\_uapress\\_9780816531240-ch013](https://doi.org/10.2458/azu_uapress_9780816531240-ch013)
- Lu C. X., Naoz S., 2019, *MNRAS*, 484, 1506
- Lucas W. E., Rybak M., Bonnell I. A., Gieles M., 2018, *MNRAS*, 474, 3582
- Luhman K. L., 2012, *ARAA*, 50, 65
- Luhman K. L., et al., 2005, *ApJL*, 631, L69
- Luo A. L., Zhao Y. H., Zhao G., et al. 2019, VizieR Online Data Catalog, p. V/164
- Luri X., et al., 2018, *A&A*, 616, A9
- Lux O., Neuhäuser R., Mugrauer M., Bischoff R., 2021, *Astronomische Nachrichten*, 342, 553
- Lynden-Bell D., 1967, *MNRAS*, 136, 101
- Lynden-Bell D., Pringle J. E., 1974, *MNRAS*, 168, 603
- Lyne A. G., Lorimer D. R., 1994, *Nature*, 369, 127
- Maíz Apellániz J., 2019, *A&A*, 630, A119
- Maíz Apellániz J., Weiler M., 2018, *A&A*, 619, A180
- Maíz Apellániz J., Pantaleoni González M., Barbá R. H., Simón-Díaz S., Negueruela I., Lennon D. J., Sota A., Trigueros Páez E., 2018, *A&A*, 616, A149
- Maíz Apellániz J., Crespo Bellido P., Barbá R. H., Fernández Aranda R., Sota A., 2020, *A&A*, 643, A138
- Makino J., 1991, Publications of the Astronomical Society of Japan, 43, 859
- Makino J., Aarseth S. J., 1992, Publications of the Astronomical Society of Japan, 44, 141

- Makino J., Taiji M., Ebisuzaki T., Sugimoto D., 1997, *ApJ*, 480, 432
- Mamajek E. E., 2009, in Usuda T., Tamura M., Ishii M., eds, American Institute of Physics Conference Series Vol. 1158, Exoplanets and Disks: Their Formation and Diversity. pp 3–10, [doi:10.1063/1.3215910](https://doi.org/10.1063/1.3215910)
- Manara C. F., Beccari G., Da Rio N., De Marchi G., Natta A., Ricci L., Robberto M., Testi L., 2013, *A&A*, 558, A114
- Marchetti T., Rossi E. M., Brown A. G. A., 2019, *MNRAS*, 490, 157
- Margulis M., Lada C. J., Young E. T., 1989, *ApJ*, 345, 906
- Mariñas N., Lada E. A., Teixeira P. S., Lada C. J., 2013, *ApJ*, 772, 81
- Marks M., Kroupa P., 2012, *A&A*, 543, A8
- Martell S. L., et al., 2017, *MNRAS*, 465, 3203
- Maschberger T., 2013, *MNRAS*, 429, 1725
- Maschberger T., Clarke C. J., 2011, *MNRAS*, 416, 541
- Mason B. D., Gies D. R., Hartkopf W. I., Bagnuolo William G. J., ten Brummelaar T., McAlister H. A., 1998, *AJ*, 115, 821
- Mathieu R. D., 1994, *ARAA*, 32, 465
- Maury A. J., et al., 2019, *A&A*, 621, A76
- Mayne N. J., Naylor T., 2008, *MNRAS*, 386, 261
- Mayor M., Duquennoy A., Halbwachs J. L., Mermilliod J. C., 1992, in McAlister H. A., Hartkopf W. I., eds, Astronomical Society of the Pacific Conference Series Vol. 32, IAU Colloq. 135: Complementary Approaches to Double and Multiple Star Research. p. 73
- McBride A., Kounkel M., 2019, *ApJ*, 884, 6
- McCuskey S. W., 1959, *ApJS*, 4, 23
- McEvoy C. M., Dufton P. L., Smoker J. V., Lambert D. L., Keenan F. P., Schneider F. R. N., de Wit W.-J., 2017, *ApJ*, 842, 32
- McGinnis P., Dougados C., Alencar S. H. P., Bouvier J., Cabrit S., 2018, *A&A*, 620, A87
- McKee C. F., Ostriker E. C., 2007, *ARAA*, 45, 565
- McKee C. F., Tan J. C., 2003, *ApJ*, 585, 850
- McMillan S. L. W., Vesperini E., Portegies Zwart S. F., 2007, *ApJ*, 655, L45

- Megeath S. T., et al., 2012, [AJ](#), 144, 192
- Menten K. M., Reid M. J., Forbrich J., Brunthaler A., 2007, [A&A](#), 474, 515
- Merritt D., 1996, [AJ](#), 111, 2462
- Michell J., 1767, Philosophical Transactions of the Royal Society of London Series I, [57](#), [234](#)
- Mikkola S., 1983, [MNRAS](#), 203, 1107
- Mikkola S., 1984a, [MNRAS](#), 207, 115
- Mikkola S., 1984b, [MNRAS](#), 208, 75
- Mikkola S., Aarseth S. J., 1998, [New Astronomy](#), 3, 309
- Miller G. E., Scalo J. M., 1979, [ApJS](#), 41, 513
- de Mink S. E., Brott I., Cantiello M., Izzard R. G., Langer N., Sana H., 2012, in Drissen L., Robert C., St-Louis N., Moffat A. F. J., eds, Astronomical Society of the Pacific Conference Series Vol. 465, Proceedings of a Scientific Meeting in Honor of Anthony F. J. Moffat. p. 65
- de Mink S. E., Sana H., Langer N., Izzard R. G., Schneider F. R. N., 2014, [ApJ](#), 782, 7
- Moeckel N., Bonnell I. A., 2009a, [MNRAS](#), 396, 1864
- Moeckel N., Bonnell I. A., 2009b, [MNRAS](#), 400, 657
- Moeckel N., Clarke C. J., 2011, [MNRAS](#), 410, 2799
- Moeckel N., Goddi C., 2012, [MNRAS](#), 419, 1390
- Moore G. E., 1965
- Moyano Loyola G. R. I., Hurley J. R., 2013, [MNRAS](#), 434, 2509
- Muench A. A., Lada C. J., Luhman K. L., Muzerolle J., Young E., 2007, [AJ](#), 134, 411
- Muench A., Getman K., Hillenbrand L., Preibisch T., 2008, in Reipurth B., ed., ASP Monograph Publications Vol. 4, Handbook of Star Forming Regions, Volume I: The Northern Sky. p. 483
- Mukherjee D., Zhu Q., Trac H., Rodriguez C. L., 2021, [ApJ](#), 916, 9
- Murphy S. J., Mamajek E. E., Bell C. P. M., 2018, [MNRAS](#), 476, 3290
- Myers P. C., Ladd E. F., 1993, [ApJL](#), [413](#), [L47](#)
- Naylor T., 2009, [MNRAS](#), 399, 432

- Németh P., Ziegerer E., Irrgang A., Geier S., Fürst F., Kupfer T., Heber U., 2016, [ApJL](#), 821, L13
- Nesterov V. V., Kuzmin A. V., Ashimbaeva N. T., Volchkov A. A., Röser S., Bastian U., 1995, *A&A Suppl.*, [110](#), 367
- Netopil M., Paunzen E., Heiter U., Soubiran C., 2016, [A&A](#), 585, A150
- Neunteufel P., 2020, [A&A](#), 641, A52
- Nicholson R. B., Parker R. J., Church R. P., Davies M. B., Fearon N. M., Walton S. R. J., 2019, [MNRAS](#), 485, 4893
- Nordhaus J., Brandt T. D., Burrows A., Almgren A., 2012, [MNRAS](#), 423, 1805
- O'Dell C. R., Poveda A., Allen C., Robberto M., 2005, [ApJL](#), 633, L45
- Ochsenbein F., Bauer P., Marcout J., 2000, [A&A Suppl.](#), 143, 23
- Offner S. S. R., Kratter K. M., Matzner C. D., Krumholz M. R., Klein R. I., 2010, [ApJ](#), 725, 1485
- Offner S. S. R., Clark P. C., Hennebelle P., Bastian N., Bate M. R., Hopkins P. F., Moraux E., Whitworth A. P., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, *Protostars and Planets VI*. p. 53, [doi:10.2458/azu\\_uapress\\_9780816531240-ch003](https://doi.org/10.2458/azu_uapress_9780816531240-ch003)
- Oh S., Kroupa P., 2016, [A&A](#), 590, A107
- Olczak C., Pfalzner S., Eckart A., 2008, [A&A](#), 488, 191
- Öpik E., 1924, *Publications of the Tartu Astrofizica Observatory*, [25](#), 1
- Padoan P., Nordlund Å., 2002, [ApJ](#), 576, 870
- Padoan P., Nordlund Å., 2004, [ApJ](#), 617, 559
- Palla F., Randich S., Pavlenko Y. V., Flaccomio E., Pallavicini R., 2007, [ApJL](#), 659, L41
- Panwar N., Pandey A. K., Samal M. R., Battinelli P., Ogura K., Ojha D. K., Chen W. P., Singh H. P., 2018, [AJ](#), 155, 44
- Park S.-M., Goodwin S. P., Kim S. S., 2018, [MNRAS](#), 478, 183
- Parker R. J., 2014, [MNRAS](#), 445, 4037
- Parker R. J., Dale J. E., 2017, [MNRAS](#), 470, 390
- Parker R. J., Goodwin S. P., 2007, [MNRAS](#), 380, 1271
- Parker R. J., Goodwin S. P., 2012, [MNRAS](#), 424, 272

- Parker R. J., Meyer M. R., 2014, [MNRAS](#), 442, 3722
- Parker R. J., Quanz S. P., 2012, [MNRAS](#), 419, 2448
- Parker R. J., Reggiani M. M., 2013, [MNRAS](#), 432, 2378
- Parker R. J., Schoettler C., 2022, [MNRAS](#), 510, 1136
- Parker R. J., Wright N. J., 2016, [MNRAS](#), 457, 3430
- Parker R. J., Wright N. J., 2018, [MNRAS](#), 481, 1679
- Parker R. J., Wright N. J., Goodwin S. P., Meyer M. R., 2014, [MNRAS](#), 438, 620
- Parker R. J., Goodwin S. P., Wright N. J., Meyer M. R., Quanz S. P., 2016, [MNRAS](#), 459, L119
- Paunzen E., Duffee B., Heiter U., Kuschnig R., Weiss W. W., 2001, [A&A](#), 373, 625
- Pavan L., Bordas P., Puhlhofer G., et al. 2016, in *Supernova Remnants: An Odyssey in Space after Stellar Death*. p. 48
- Pelkonen V. M., Padoan P., Haugbølle T., Nordlund Å., 2021, [MNRAS](#), 504, 1219
- Pelupessy F. I., van Elteren A., de Vries N., McMillan S. L. W., Drost N., Portegies Zwart S. F., 2013, [A&A](#), 557, A84
- Penoyre Z., Belokurov V., Wyn Evans N., Everall A., Koposov S. E., 2020, [MNRAS](#), 495, 321
- Perets H. B., 2009, [ApJ](#), 698, 1330
- Perets H. B., Šubr L., 2012, [ApJ](#), 751, 133
- Peretto N., André P., Belloche A., 2006, [A&A](#), 445, 979
- Peretto N., Hennebelle P., André P., 2007, [A&A](#), 464, 983
- Peri C. S., Benaglia P., Brookes D. P., Stevens I. R., Isequilla N. L., 2012, [A&A](#), 538, A108
- Perryman M. A. C., ESA, FAST Consortium, 1997, *The Hipparcos and Tycho catalogues: astrometric and photometric star catalogues derived from the ESA Hipparcos Space Astrometry Mission*. ESA Special Publication Vol. 1200, ESA Publications Division, Noordwijk, Netherlands
- Peter D., Feldt M., Henning T., Hormuth F., 2012, [A&A](#), 538, A74
- Peters T., Banerjee R., Klessen R. S., Mac Low M.-M., Galván-Madrid R., Keto E. R., 2010a, [ApJ](#), 711, 1017

- Peters T., Klessen R. S., Mac Low M.-M., Banerjee R., 2010b, *ApJ*, 725, 134
- Pflamm-Altenburg J., Kroupa P., 2010, *MNRAS*, 404, 1564
- Pietrzynski G., Udalski A., 2000, *Acta Astronomica*, 50, 355
- Piffl T., et al., 2014, *A&A*, 562, A91
- Pinte C., et al., 2020, *ApJL*, 890, L9
- Platais I., et al., 2020, *AJ*, 159, 272
- Plummer H. C., 1911, *MNRAS*, 71, 460
- Portegies Zwart S. F., 2000, *ApJ*, 544, 437
- Portegies Zwart S. F., McMillan S. L. W., 2018, *Astrophysical Recipes: The art of AMUSE*. IOP Publishing, doi:10.1088/978-0-7503-1320-9
- Portegies Zwart S. F., Verbunt F., 1996, *A&A*, 309, 179
- Portegies Zwart S. F., Makino J., McMillan S. L. W., Hut P., 1999, *A&A*, 348, 117
- Portegies Zwart S. F., McMillan S. L. W., Hut P., Makino J., 2001, *MNRAS*, 321, 199
- Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, *ARAA*, 48, 431
- Portegies Zwart S. F., McMillan S. L. W., van Elteren E., Pelupessy I., de Vries N., 2013, *Computer Physics Communications*, 184, 456
- Poveda A., Ruiz J., Allen C., 1967, *Boletín de los Observatorios Tonantzintla y Tacubaya*, 4, 86
- Poveda A., Allen C., Hernández-Alcántara A., 2005, *ApJ*, 627, L61
- Preibisch T., Balega Y., Hofmann K.-H., Weigelt G., Zinnecker H., 1999, *New Astronomy*, 4, 531
- Priyatikanto R., Kouwenhoven M. B. N., Arifyanto M. I., Wulandari H. R. T., Siregar S., 2016, *MNRAS*, 457, 1339
- Przybilla N., Fernanda Nieva M., Heber U., Butler K., 2008, *ApJL*, 684, L103
- Pudritz R. E., Kevlahan N. K. R., 2013, *Philosophical Transactions of the Royal Society of London Series A*, 371, 20120248
- Raddi R., Hollands M. A., Gänsicke B. T., Townsley D. M., Hermes J. J., Gentile Fusillo N. P., Koester D., 2018, *MNRAS*, 479, L96
- Raddi R., Irrgang A., Heber U., Schneider D., Kreuzer S., 2021, *A&A*, 645, A108

- Radhakrishnan V., Shukre C. S., 1985, in Srinivasan G., Radhakrishnan V., eds, *Supernovae, their Progenitors and Remnants*. p. 155
- Raghavan D., Farrington C. D., ten Brummelaar T. A., McAlister H. A., Ridgway S. T., Sturm L., Sturm J., Turner N. H., 2012, *ApJ*, 745, 24
- Rappenglück M. A., 2001, *Earth Moon and Planets*, 85, 391
- Rapson V. A., Pipher J. L., Gutermuth R. A., Megeath S. T., Allen T. S., Myers P. C., Allen L. E., 2014, *ApJ*, 794, 124
- Rate G., Crowther P. A., 2020, *MNRAS*, 493, 1512
- Rate G., Crowther P. A., Parker R. J., 2020, *MNRAS*, 495, 1209
- Rebull L. M., Hillenbrand L. A., Strom S. E., Duncan D. K., Patten B. M., Pavlovsky C. M., Makidon R., Adams M. T., 2000, *AJ*, 119, 3026
- Rebull L. M., et al., 2002, *AJ*, 123, 1528
- Rebull L. M., Stauffer J. R., Megeath S. T., Hora J. L., Hartmann L., 2006, *ApJ*, 646, 297
- Reggiani M. M., Meyer M. R., 2011, *ApJ*, 738, 60
- Reggiani M. M., Meyer M. R., 2013, *A&A*, 553, A124
- Reggiani M. M., Robberto M., Da Rio N., Meyer M., Soderblom D., Ricci L., 2011, *A&A*, 534, A83
- Reipurth B., Clarke C., 2001, *AJ*, 122, 432
- Reipurth B., Mikkola S., 2015, *AJ*, 149, 145
- Reipurth B., Guimarães M. M., Connelley M. S., Bally J., 2007, *AJ*, 134, 2272
- Renzo M., et al., 2019, *A&A*, 624, A66
- Rhode K. L., Herbst W., Mathieu R. D., 2001, *AJ*, 122, 3258
- Riaz R., Schleicher D. R. G., Vanaverbeke S., Klessen R. S., 2021, *Astronomische Nachrichten*, 342, 157
- Rice W. K. M., Armitage P. J., Bonnell I. A., Bate M. R., Jeffers S. V., Vine S. G., 2003, *MNRAS*, 346, L36
- Richert A. J. W., Getman K. V., Feigelson E. D., Kuhn M. A., Broos P. S., Povich M. S., Bate M. R., Garmire G. P., 2018, *MNRAS*, 477, 5191

- Ridpath I., 2018, A Dictionary of Astronomy. Oxford University Press, doi:10.1093/acref/9780191851193.001.0001, <https://www.oxfordreference.com/view/10.1093/acref/9780191851193.001.0001/acref-9780191851193>
- Rilinger A. M., Espaillat C. C., 2021, *ApJ*, 921, 182
- Rodríguez L. F., Poveda A., Lizano S., Allen C., 2005, *ApJL*, 627, L65
- Rodríguez C. L., Morscher M., Wang L., Chatterjee S., Rasio F. A., Spurzem R., 2016, *MNRAS*, 463, 2109
- Rodríguez C. L., Pattabiraman B., Chatterjee S., Choudhary A., Liao W.-k., Morscher M., Rasio F. A., 2018, *Computational Astrophysics and Cosmology*, 5, 5
- Rodríguez C. L., et al., 2022, *ApJS*, 258, 22
- de Rosa R. J., et al., 2012, *MNRAS*, 422, 2765
- de Rosa R. J., et al., 2014, *MNRAS*, 437, 1216
- Rosen A. L., Offner S. S. R., Sadavoy S. I., Bhandare A., Vázquez-Semadeni E., Ginsburg A., 2020, *Space Science Reviews*, 216, 62
- Rozhavskii F. G., Kuz'mina V. A., Vasilevskii A. E., 1976, *Astrophysics*, 12, 204
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sana H., et al., 2012, *Science*, 337, 444
- Sana H., et al., 2013, *A&A*, 550, A107
- Sandquist E. L., et al., 2021, *AJ*, 161, 59
- Santos N. C., Melo C., James D. J., Gameiro J. F., Bouvier J., Gomes J. I., 2008, *A&A*, 480, 889
- Sartoretti P., et al., 2018, *A&A*, 616, A6
- Scalo J. M., 1986, in De Loore C. W. H., Willis A. J., Laskarides P., eds, Proceedings of the Symposium Vol. 116, Luminous Stars and Associations in Galaxies. p. 451
- Scalo J., 2005, in Corbelli E., Palla F., Zinnecker H., eds, Astrophysics and Space Science Library Vol. 327, The Initial Mass Function 50 Years Later. Springer, Dordrecht, p. 23, doi:10.1007/978-1-4020-3407-7\_4
- Schmeja S., 2011, *Astronomische Nachrichten*, 332, 172
- Schneider F. R. N., Podsiadlowski P., Langer N., Castro N., Fossati L., 2016, *MNRAS*, 457, 2355
- Schneider F. R. N., et al., 2018, *Science*, 359, 69



- Schoettler C., Parker R. J., 2021, [MNRAS](#), 501, L12
- Schoettler C., Parker R. J., Arnold B., Grimmett L. P., de Bruijne J., Wright N. J., 2019, [MNRAS](#), 487, 4615
- Schoettler C., de Bruijne J., Vaher E., Parker R. J., 2020, [MNRAS](#), 495, 3104
- Schoettler C., Parker R. J., de Bruijne J., 2022, [MNRAS](#), 510, 3178
- Schönrich R., Binney J., Dehnen W., 2010, [MNRAS](#), 403, 1829
- Segura-Cox D. M., et al., 2018, [ApJ](#), 866, 161
- Sergison D. J., Mayne N. J., Naylor T., Jeffries R. D., Bell C. P. M., 2013, [MNRAS](#), 434, 966
- Shklovskii I. S., 1970, *Soviet Astronomy*, **13**, 562
- Shu F. H., Adams F. C., Lizano S., 1987, [ARAA](#), 25, 23
- Shukirgaliyev B., Parmentier G., Just A., Berczik P., 2018, [ApJ](#), 863, 171
- Sicilia-Aguilar A., et al., 2005, [AJ](#), 129, 363
- Sicilia-Aguilar A., et al., 2016, [PASA](#), 33, e059
- Sigalotti L. D. G., Cruz F., Gabbasov R., Klapp J., Ramírez-Velasquez J., 2018, [ApJ](#), 857, 40
- Sills A., Rieder S., Scora J., McCloskey J., Jaffa S., 2018, [MNRAS](#), 477, 1903
- Silva M. D. V., Napiwotzki R., 2011, [MNRAS](#), 411, 2596
- Silverberg S. M., et al., 2020, [ApJ](#), 890, 106
- Silverman B. W., 1986, *Density estimation for statistics and data analysis. Monographs on Statistics and Applied Probability Vol. 26*, Chapman and Hall, London
- Smartt S. J., 2009, [ARAA](#), 47, 63
- Smartt S. J., 2015, [PASA](#), 32, E016
- Sota A., Maíz Apellániz J., Walborn N. R., Alfaro E. J., Barbá R. H., Morrell N. I., Gamen R. C., Arias J. I., 2011, [ApJS](#), 193, 24
- Spera M., Capuzzo-Dolcetta R., 2017, [Ap&SS](#), 362, 233
- Spera M., Mapelli M., Bressan A., 2015, [MNRAS](#), 451, 4086
- Spina L., et al., 2017, [A&A](#), 601, A70
- Spurzem R., 1999, *Journal of Computational and Applied Mathematics*, 109, 407

- Spurzem R., Takahashi K., 1995, [MNRAS](#), 272, 772
- Stamatellos D., Hubber D. A., Whitworth A. P., 2007, [MNRAS](#), 382, L30
- Stamatellos D., Maury A., Whitworth A., André P., 2011a, [MNRAS](#), 413, 1787
- Stamatellos D., Whitworth A. P., Hubber D. A., 2011b, [ApJ](#), 730, 32
- Stammler S. M., Drażkowska J., Birnstiel T., Klahr H., Dullemond C. P., Andrews S. M., 2019, [ApJL](#), 884, L5
- Stassun K. G., et al., 2019, [AJ](#), 158, 138
- Stone R. C., 1979, [ApJ](#), 232, 520
- Stone R. C., 1991, [AJ](#), 102, 333
- Struve O., 1930, [ApJ](#), 72, 1
- Sung H., Stauffer J. R., Bessell M. S., 2009, [AJ](#), 138, 1116
- Takemura H., et al., 2021, [ApJL](#), 910, L6
- Tan J. C., 2004, [ApJL](#), 607, L47
- Tang S.-Y., Chen W. P., Chiang P. S., Jose J., Herczeg G. J., Goldman B., 2018, [ApJ](#), 862, 106
- Tauris T. M., 2015, [MNRAS](#), 448, L6
- Tauris T. M., Bailes M., 1996, [A&A](#), 315, 432
- Tauris T. M., Takens R. J., 1998, [A&A](#), 330, 1047
- Teixeira P. S., et al., 2006, [ApJL](#), 636, L45
- Teixeira P. S., Lada C. J., Marengo M., Lada E. A., 2012, [A&A](#), 540, A83
- Terebey S., Shu F. H., Cassen P., 1984, [ApJ](#), 286, 529
- Tetzlaff N., Neuhäuser R., Hohle M. M., Maciejewski G., 2010, [MNRAS](#), 402, 2369
- Tetzlaff N., Neuhäuser R., Hohle M. M., 2011, [MNRAS](#), 410, 190
- Thompson R. I., Corbin M. R., Young E., Schneider G., 1998, [ApJL](#), 492, L177
- Tobin J. J., Hartmann L., Furesz G., Mateo M., Megeath S. T., 2009, [ApJ](#), 697, 1103
- Tobin J. J., Hartmann L., Fűrész G., Hsu W.-H., Mateo M., 2015a, [AJ](#), 149, 119
- Tobin J. J., et al., 2015b, [ApJ](#), 805, 125

- Tobin J. J., et al., 2016a, *Nature*, 538, 483
- Tobin J. J., et al., 2016b, *ApJ*, 818, 73
- Tobin J. J., et al., 2020, *ApJ*, 890, 130
- Tohline J. E., 1981, *ApJ*, 248, 717
- Tohline J. E., 2002, *ARAA*, 40, 349
- Tokovinin A., 2018, *ApJS*, 235, 6
- Tomsick J. A., Bodaghee A., Rodriguez J., Chaty S., Camilo F., Fornasini F., Rahoui F., 2012, *ApJL*, 750, L39
- Tout C. A., Aarseth S. J., Pols O. R., Eggleton P. P., 1997, *MNRAS*, 291, 732
- Trapman L., Rosotti G., Bosman A. D., Hogerheijde M. R., van Dishoeck E. F., 2020, *A&A*, 640, A5
- Trenti M., Hut P., 2008, arXiv e-prints, p. [arXiv:0806.3950](https://arxiv.org/abs/0806.3950)
- Tsukamoto Y., Okuzumi S., Kataoka A., 2017, *ApJ*, 838, 151
- Tutukov A. V., 1978, *A&A*, 70, 57
- van Albada T. S., 1982, *MNRAS*, 201, 939
- Vaher E., 2020, *Research Notes of the American Astronomical Society*, 4, 116
- Vargas-Salazar I., Oey M. S., Barnes J. R., Chen X., Castro N., Kratter K. M., Faerber T. A., 2020, *ApJ*, 903, 42
- Venuti L., et al., 2014, *A&A*, 570, A82
- Venuti L., et al., 2015, *A&A*, 581, A66
- Venuti L., et al., 2017, *A&A*, 599, A23
- Venuti L., et al., 2018, *A&A*, 609, A10
- Vincke K., Pfalzner S., 2016, *ApJ*, 828, 48
- Vincke K., Pfalzner S., 2018, *ApJ*, 868, 1
- Vorobyov E. I., Basu S., 2010, *ApJ*, 719, 1896
- Voroshilov V. I., Guseva N. G., Kalandadze N. B., Kolesnik L. N., Kuznetsov V. I., Metreveli M. D., Shapovalov A. N., 1985, Catalogue of BV magnitudes and spectral classes for 6000 stars. Ukrainian Acad. Nauk, Kiev

- Wang J., Townsley L. K., Feigelson E. D., Broos P. S., Getman K. V., Román-Zúñiga C. G., Lada E., 2008, [ApJ](#), 675, 464
- Wang L., Spurzem R., Aarseth S., Nitadori K., Berczik P., Kouwenhoven M. B. N., Naab T., 2015, [MNRAS](#), 450, 4070
- Wang L., et al., 2016, [MNRAS](#), 458, 1450
- Wang L., Kroupa P., Jerabkova T., 2019, [MNRAS](#), 484, 1843
- Wareing C. J., Pittard J. M., Wright N. J., Falle S. A. E. G., 2018, [MNRAS](#), 475, 3598
- Waters L. B. F. M., Waelkens C., 1998, [ARAA](#), 36, 233
- Weidner C., Kroupa P., 2004, [MNRAS](#), 348, 187
- Weidner C., Bonnell I. A., Moeckel N., 2011, [MNRAS](#), 410, 1861
- Weiler M., 2018, [A&A](#), 617, A138
- Wenger M., et al., 2000, [A&A Suppl.](#), 143, 9
- Whitworth A. P., Zinnecker H., 2004, [A&A](#), 427, 299
- Wilhelm M. J. C., Portegies Zwart S., 2021, [MNRAS](#)
- Wilkin F. P., 1996, [ApJL](#), 459, L31
- Williams J. P., Cieza L. A., 2011, [ARAA](#), 49, 67
- Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017, [MNRAS](#), 468, 2359
- Wilson R. E., 1953, Carnegie Institute Washington D.C. Publication, p. 0
- Winter A. J., Clarke C. J., Rosotti G., Ih J., Facchini S., Haworth T. J., 2018, [MNRAS](#), 478, 2700
- de Wit W. J., Testi L., Palla F., Zinnecker H., 2005, [A&A](#), 437, 247
- Wolff S. C., Strom S. E., Dror D., Venn K., 2007, [AJ](#), 133, 1092
- Wootton B. A., Parker R. J., 2019, [MNRAS](#), 485, L48
- Wright N. J., 2020, [New Astronomy Reviews](#), 90, 101549
- Wright N. J., Drake J. J., Drew J. E., Vink J. S., 2010a, [ApJ](#), 713, 871
- Wright N. J., Drake J. J., Drew J. E., Vink J. S., 2010b, [ApJ](#), 713, 871
- Wright N. J., Parker R. J., Goodwin S. P., Drake J. J., 2014, [MNRAS](#), 438, 639

- Wright N. J., Bouy H., Drew J. E., Sarro L. M., Bertin E., Cuillandre J.-C., Barrado D., 2016, [MNRAS](#), 460, 2593
- Yorke H. W., Sonnhalter C., 2002, [ApJ](#), 569, 846
- Young C. H., Shirley Y. L., Evans Neal J. I., Rawlings J. M. C., 2003, [ApJS](#), 145, 111
- Zari E., Hashemi H., Brown A. G. A., Jardine K., de Zeeuw P. T., 2018, [A&A](#), 620, A172
- Zhang Y., et al., 2019, [ApJ](#), 873, 73
- Zhao H., Jiang B., Gao S., Li J., Sun M., 2018, [ApJ](#), 855, 12
- Zhao B., et al., 2020, [Space Science Reviews](#), 216, 43
- Zhong J., Chen L., Kouwenhoven M. B. N., Li L., Shao Z., Hou J., 2019, [A&A](#), 624, A34
- Zinnecker H., 2011, in Treyer M., Wyder T., Neill J., Seibert M., Lee J., eds, *Astronomical Society of the Pacific Conference Series Vol. 440, UP2010: Have Observations Revealed a Variable Upper End of the Initial Mass Function?*. p. 3
- Zinnecker H., Yorke H. W., 2007, [ARAA](#), 45, 481
- Zwicky F., 1957, *Morphological Astronomy*. Springer Berlin Heidelberg
- Zwintz K., et al., 2017, [A&A](#), 601, A101
- van den Heuvel E. P. J., van Paradijs J., 1997, [ApJ](#), 483, 399
- van den Heuvel E. P. J., Portegies Zwart S. F., Bhattacharya D., Kaper L., 2000, [A&A](#), 364, 563

# Appendix A

## Additional information for *Gaia* DR2 identified stars from the ONC

## A.1 Excluded candidates from ONC analysis

Table A.1 provides information on the identified 2D-candidates that have been excluded, either due to their RV pointing towards the ONC or because their required RV due to their radial position would be above  $|500| \text{ km s}^{-1}$ .

Table A.1: Excluded RW and WW star 2D-candidates. Column 2+3: velocity in ONC rest frame [rf]; Column 3: RV sources – <sup>a</sup>*Gaia* DR2, <sup>b</sup>Cottaar et al. (2015), <sup>c</sup>Gontcharov (2006); Column 4: Reason for exclusion; Column 5–7: from literature sources – <sup>1</sup>Da Rio et al. (2016), <sup>2</sup>Houk & Swift (1999), <sup>3</sup>Nesterov et al. (1995), <sup>4</sup>Houk & Smith-Moore (1988), <sup>5</sup>Tetzlaff et al. (2011).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Exclusion reason	Age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
Excluded RW candidates						
3011539434830106624	61.5 ± 0.5	-33.2 ± 6.9 <sup>a</sup>	RV points towards ONC	-	-	-
3016948241764201472	50.2 ± 0.8	-	RV required >  500  km s <sup>-1</sup>	-	-	-
3011184292574204672	43.5 ± 0.4	10.8 ± 6.6 <sup>a</sup>	RV points towards ONC	-	-	-
3017157630011568000	42.0 ± 0.4	-58.2 ± 6.6 <sup>b</sup>	RV points towards ONC	131 <sup>1</sup>	1.4 <sup>1</sup>	-
3017247618166246784	39.8 ± 0.4	85.9 ± 6.6 <sup>a</sup>	RV points towards ONC	2.0 <sup>1</sup>	2.2 <sup>1</sup>	-
3122079111207514496	39.2 ± 0.4	21.4 ± 6.6 <sup>a</sup>	RV points towards ONC	-	-	-
3206712880587397120	37.1 ± 0.4	-19.7 ± 6.6 <sup>a</sup>	RV points towards ONC	-	-	-
3187254518368160000	35.4 ± 0.5	-72.1 ± 6.6 <sup>a</sup>	RV points towards ONC	-	-	-
Excluded WW star candidates						
3209554744485606400	24.5 ± 0.7	-	RV required >  500  km s <sup>-1</sup>	-	-	-
3013314424554785024	23.9 ± 0.4	-17.2 ± 6.6 <sup>a</sup>	RV points towards ONC	-	-	G8/K0 <sup>2</sup>
2989308443587969664	18.5 ± 0.3	0.3 ± 6.6 <sup>a</sup>	RV points towards ONC	-	-	F6/7 <sup>4</sup>
3017240746218498176	17.9 ± 0.6	-2.3 ± 6.6 <sup>b</sup>	RV points towards ONC	-	-	-
3015045708692252672	17.4 ± 0.5	-11.4 ± 6.6 <sup>a</sup>	RV points towards ONC	-	-	-
3023538370864242688	17.3 ± 0.5	-19.2 ± 6.6 <sup>b</sup>	RV points towards ONC	-	-	A7 <sup>3</sup>
3019516254250648832	14.6 ± 0.4	-23.2 ± 6.7 <sup>a</sup>	RV points towards ONC	-	-	-
3215185309169853568	11.1 ± 0.6	8.1 ± 6.9 <sup>c</sup>	RV points towards ONC	0.2 <sup>5</sup>	7.9 <sup>5</sup>	B3 <sup>2</sup>
3017244216552060672	0.5 ± 0.4	15.6 ± 9.8 <sup>a</sup>	RV points towards ONC	5.7 <sup>1</sup>	1.3 <sup>1</sup>	-

## A.2 2D-candidates with multiple, varying RV measurements for the ONC

Table A.2 provides information on identified 2D-candidates that would not be identified as RWs or WWs only by their 2D-velocity but could be when considering their RV measurements. I have identified several such stars in Table 5.4, but also have stars with multiple RV measurements between different literature sources, which could also be indicating a binary system.



Table A.2: Slow 2D-candidates with multiple, varying RV measurements; Column 2: velocity in ONC rest frame [rf]; Column 3: RV sources of varying measurements - <sup>a</sup>*Gaia* DR2, <sup>b</sup>Cottaar et al. (2015), <sup>c</sup>Kounkel et al. (2018); Column 4: indication of 3D-candidate status; Column 5: minimum flight time since ejection (crossing of search boundary); Column 6: age from PARSEC isochrones (Bressan et al. 2012); Column 7–9: from literature sources - <sup>1</sup>Hillenbrand (1997), <sup>2</sup>Da Rio et al. (2016).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3209532135777678208	3.2 ± 0.4	a, b	?	in cluster	5.0 <sup>+5.0</sup> <sub>-3.0</sub>	5.6 <sup>2</sup>	2.1 <sup>2</sup>	-
3017270669252519680	2.9 ± 0.4	a, b	?	in cluster	1.0 <sup>+10.0</sup> <sub>-0.9</sub>	0.6 <sup>2</sup>	1.0 <sup>2</sup>	-
3209528012609165440	2.5 ± 0.3	a, b	?	in cluster	2.5 <sup>+20.0</sup> <sub>-2.2</sub>	1.4 <sup>2</sup>	0.9 <sup>2</sup>	-
3209521037582290304	1.7 ± 0.4	a, c	?	in cluster	0.3 ± 0.2	0.3 <sup>2</sup>	1.6 <sup>2</sup>	K3 <sup>1</sup>

### A.3 Past visitors to the ONC

Table A.3 provides information on sources that can be successfully traced back in 3D to the ONC in the past 2.4 Myr (upper age limit implied by the identified RW stars) but their position on the CMD identifies them as MS stars or older pre-MS stars.

Table A.3: Past visitors to the ONC. Column 2+3: velocity in ONC rest frame [rf]; Column 3: RV sources – <sup>a</sup>*Gaia* DR2; Column 4: age from PARSEC isochrones (Bressan et al. 2012); Column 5–7: from literature sources – <sup>1</sup>Da Rio et al. (2016).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Iso. age (Myr)	Age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
Visitors at RW velocities						
2986529565387616896	81.9 ± 0.5	-34.1 ± 6.8 <sup>a</sup>	~22–30	-	-	-
3240725452454418048	67.2 ± 0.5	0.2 ± 6.7 <sup>a</sup>	~18–20	-	-	-
3014832609594260224	62.2 ± 1.3	-19.8 ± 21.2 <sup>a</sup>	~15–20	-	-	-
3021115184676332288	59.1 ± 0.4	2.7 ± 6.7 <sup>a</sup>	~20	-	-	-
3024722888484450944	54.2 ± 0.4	17.5 ± 6.7 <sup>a</sup>	~17–20	-	-	-
2992509671692197632	50.7 ± 0.4	-47.0 ± 6.7 <sup>a</sup>	~25	-	-	-
3009308457018637824	50.6 ± 0.4	-72.3 ± 6.7 <sup>a</sup>	~25–30	-	-	-
2989584932106141184	48.4 ± 0.4	-30.2 ± 6.6 <sup>a</sup>	~20–25	-	-	-
3015908138125994112	44.1 ± 0.4	26.1 ± 6.8 <sup>a</sup>	~20	-	-	-
3240501873637049856	43.0 ± 0.4	-24.9 ± 6.7 <sup>a</sup>	~20	-	-	-
2971498824821941760	42.0 ± 0.4	-33.8 ± 6.7 <sup>a</sup>	~17–20	-	-	-
3005776138475075712	40.5 ± 1.4	3.8 ± 9.6 <sup>a</sup>	~30–35	-	-	-
2998151094058000128	39.3 ± 0.4	-41.0 ± 6.7 <sup>a</sup>	~28–30	-	-	-
3228935835946246272	34.0 ± 0.4	-2.2 ± 6.7 <sup>a</sup>	~12–15	-	-	-
3014762309569718272	33.5 ± 0.3	-47.8 ± 6.7 <sup>a</sup>	~30	-	-	-
2989747213149681408	31.9 ± 0.4	-8.6 ± 6.7 <sup>a</sup>	~25–30	-	-	-
3207687464502067456	29.5 ± 0.5	32.7 ± 6.7 <sup>a</sup>	~12–15	-	-	-
3011733017597107840	28.6 ± 0.4	-16.7 ± 6.8 <sup>a</sup>	~20–22	-	-	-
3017817268267372032	27.8 ± 0.6	-28.4 ± 7.8 <sup>a</sup>	~30–35	-	-	-
3022827296078444672	27.7 ± 0.4	27.3 ± 6.8 <sup>a</sup>	~17–20	-	-	-
3010392369324807808	26.5 ± 0.4	-15.2 ± 6.9 <sup>a</sup>	~30	-	-	-
3019025799048789504	24.6 ± 0.4	20.6 ± 6.6 <sup>a</sup>	~17–20	-	-	-
3017348803299280896	22.4 ± 0.4	-46.2 ± 6.7 <sup>a</sup>	~25–30	27.0 <sup>1</sup>	1.1 <sup>1</sup>	-
3210090515884826752	22.2 ± 0.4	22.8 ± 7.1 <sup>a</sup>	~25–30	-	-	-
3209536636903447936	15.6 ± 0.4	-38.7 ± 6.7 <sup>a</sup>	~12–17	16.2 <sup>1</sup>	1.4 <sup>1</sup>	-
Visitors at WW velocities						
3016578221743133952	24.2 ± 0.4	7.8 ± 6.9 <sup>a</sup>	~35–40	-	-	-
3219402241203000576	23.0 ± 0.5	-3.2 ± 7.3 <sup>a</sup>	~30	-	-	-
2996472071080530176	22.7 ± 0.5	-17.5 ± 6.9 <sup>a</sup>	~30	-	-	-
3010331032894313984	21.9 ± 0.4	-10.1 ± 6.7 <sup>a</sup>	~7–10	-	-	-
3010517434474967040	20.6 ± 1.1	4.5 ± 13.0 <sup>a</sup>	~5–20	-	-	-
3010652777484841344	20.5 ± 0.4	-3.0 ± 6.7 <sup>a</sup>	~28–30	-	-	-
3208716023269799808	19.8 ± 1.2	-0.9 ± 14.7 <sup>a</sup>	~30–32	-	-	-
3009639233922459008	17.2 ± 0.6	-22.6 ± 7.7 <sup>a</sup>	~30	-	-	-
3207885750255374336	16.6 ± 0.5	-18.7 ± 7.5 <sup>a</sup>	~40	-	-	-
3014826802798497920	16.6 ± 0.5	-9.3 ± 7.6 <sup>a</sup>	~30–32	-	-	-
3022823615290307840	15.4 ± 0.4	-24.8 ± 6.8 <sup>a</sup>	~20	-	-	-
3210977649969156224	13.9 ± 0.3	0.5 ± 6.6 <sup>a</sup>	~20–35	-	-	-

## A.4 Red giants visiting the ONC

Table B.9 provides information on sources that can be successfully traced forward in 2D and 3D to the ONC, however are located on or near the red giant branch and are therefore at the end of their stellar evolution. I state these here explicitly as some of these stars have large errors that could potentially turn them in to viable young candidates.

Table A.4: 2D and 3D tracebacks to the ONC that are located on near the red giant branch. Column 2+3: velocity in ONC rest frame [rf]; Column 3: RV sources – <sup>a</sup>*Gaia* DR2; Column 4: age from PARSEC isochrones (Bressan et al. 2012); Column 5–7: from literature sources – <sup>1</sup>Voroshilov et al. (1985), <sup>2</sup>McCuskey (1959), <sup>3</sup>Cannon & Pickering (1993).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )
2998537847270106240	42.7 ± 0.4	40.0 ± 6.6 <sup>a</sup>
3316474244466450432	24.8 ± 0.4	43.3 ± 6.6 <sup>a</sup>
2995865656058327552	46.7 ± 1.3	-
2989899774685582592	46.3 ± 0.4	-2.1 ± 6.6 <sup>a</sup>
3007636546509301376	17.0 ± 0.4	23.2 ± 6.6 <sup>a</sup>

## A.5 ONC high-velocity stars without a circumstellar disc

Table A.5 shows the high-velocity stars around the ONC that were found in surveys searching for circumstellar discs but show no evidence of any disc material indicators.

Table A.5: Stars ejected from the ONC without a circumstellar disc; Column 2+3: velocity in ONC rest frame [rf] from [Schoettler et al. \(2020\)](#); Column 3: RV sources - <sup>a</sup>*Gaia* DR2, <sup>b</sup>[Cottaar et al. \(2015\)](#), <sup>c</sup>[Kounkel et al. \(2018\)](#); Column 4: status identifier – ONC ejected, future or past visitor; Column 5–6: from literature sources - <sup>1</sup>[Van Altena et al. \(1988\)](#), <sup>2</sup>[Hillenbrand \(1997\)](#), <sup>3</sup>[Da Rio et al. \(2010\)](#), <sup>4</sup>[Da Rio et al. \(2012\)](#), <sup>5</sup>[Da Rio et al. \(2016\)](#); Column 7: Disc information from literature sources: [6] [Hillenbrand et al. \(1998\)](#), [7] [Rebull et al. \(2000\)](#), [8] [Rebull et al. \(2006\)](#), [9] [Fűrész et al. \(2008\)](#), [10] [Fang et al. \(2013\)](#).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Status identifier	Age (Myr)	Mass (M <sub>⊙</sub> )	Disc information source
3209624872711454976	18.1	14.7 <sup>a</sup>	ONC ejected	0.4 <sup>5</sup>	0.5 <sup>5</sup>	[7, 10]
3017166907140904320	17.2	5.3 <sup>b</sup>	ONC ejected	1.0 <sup>5</sup>	0.6 <sup>5</sup>	[7]
3017242051888552704	16.7	-5.5 <sup>b</sup>	ONC ejected	1.8 <sup>5</sup>	0.7 <sup>5</sup>	[7, 8]
3209424108758593408	16.4	8.3 <sup>b</sup>	ONC ejected	0.5-2.5 <sup>3,5</sup>	1.1-2.3 <sup>3,5</sup>	[7, 8]
3017402614955763200	14.8	-13.7 <sup>a</sup>	ONC ejected	-	-	[7]
3017260022031719040	10.6	8.7 <sup>c</sup>	ONC ejected	0.7-1.5 <sup>3,4</sup>	0.3-0.5 <sup>3,4</sup>	[7, 8]
3209529112120792320	3.0	12.1 <sup>b</sup>	ONC ejected	6.2 <sup>5</sup>	1.1 <sup>5</sup>	[7]
3209531650444835840	13.0	-4.2 <sup>a</sup>	ONC ejected	-	3.8 <sup>2</sup>	[7]
3017341385903759744	11.7	1.1 <sup>c</sup>	ONC ejected	0.7-0.9 <sup>3,5</sup>	0.5 <sup>3,5</sup>	[7]
3017252600328207104	9.3	-4.9 <sup>c</sup>	ONC ejected	0.1 <sup>5</sup>	0.3 <sup>5</sup>	[7]

## Appendix B

**Additional information for *Gaia* DR2  
identified stars from NGC 2264**

## **B.1 RW/WW predictions from simulations for the full mass range**

Table B.1 provides predicted numbers for the amount of RW/WW stars that we should expect to find around young star-forming regions based on simulations with different initial conditions as shown in Table 7.4.

Table B.1: Ejected RW and WW stars from  $N$ -body simulations within the search radius of 100 pc at different times during the simulations. For all our initial condition combinations, I show averages from all 20 simulations and the maximum from a single simulation. I count ejected binary systems as one star when calculating averages and maxima. The uncertainties in our averages are the standard deviations. I show ejected stars with masses from 0.1–50  $M_{\odot}$ .

Mass $m$ ( $M_{\odot}$ )	Simulation ID							
RW stars								
	16-03-1	30-03-1	16-05-1	30-05-1	16-03-5	20-03-5	16-05-5	20-05-5
$0.1 \leq m < 8.0$								
- after 1 Myr	7.8±2.3 / 10	0.6±0.7 / 2	8.6±4.4 / 19	0.4±0.7 / 2	3.9±1.6 / 7	0.4±0.8 / 3	3.4±1.6 / 6	0.4±0.7 / 2
- after 2 Myr	5.4±1.9 / 9	0.6±0.9 / 4	6.6±2.9 / 12	0.8±1.1 / 4	2.8±1.4 / 6	0.4±0.7 / 2	2.1±1.2 / 4	0.3±0.4 / 1
- after 3 Myr	2.0±1.2 / 5	0.5±1.1 / 5	2.8±1.9 / 8	0.5±0.9 / 3	0.8±0.9 / 3	0.2±0.4 / 1	0.8±1.0 / 3	0.1±0.3 / 1
- after 4 Myr	0.8±1.1 / 4	0.4±0.6 / 2	0.9±0.8 / 3	0.3±0.5 / 2	0.2±0.4 / 1	0 / 0	0.1±0.3 / 1	0.2±0.4 / 1
- after 5 Myr	0.3±0.6 / 2	0.3±0.6 / 2	0.4±0.6 / 2	0.2±0.4 / 1	0.1±0.2 / 1	0 / 0	0.2±0.4 / 1	0.1±0.3 / 1
$m \geq 8.0$								
- after 1 Myr	0.1±0.2 / 1	0 / 0	0.1±0.1 / 1	0 / 0	0.1±0.2 / 1	0.1±0.2 / 1	0.1±0.2 / 1	0 / 0
- after 2 Myr	0.1±0.3 / 1	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0.1±0.2 / 1	0 / 0
- after 3 Myr	0.1±0.2 / 1	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
- after 4 Myr	0.2±0.5 / 2	0.1±0.2 / 1	0.1±0.1 / 1	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
- after 5 Myr	0.1±0.2 / 1	0.1±0.2 / 1	0.1±0.1 / 1	0.1±0.4 / 2	0 / 0	0 / 0	0 / 0	0 / 0
WW stars								
	16-03-1	30-03-1	16-05-1	30-05-1	16-03-5	20-03-5	16-05-5	20-05-5
$0.1 \leq m < 8.0$								
- after 1 Myr	53.6±9.3 / 79	4.1±2.6 / 10	51.7±8.5 / 65	2.0±1.6 / 6	20.4±4.3 / 28	6.1±2.8 / 13	18.9±4.9 / 31	4.9±2.1 / 10
- after 2 Myr	55.6±10.2 / 72	5.0±2.8 / 10	54.1±9.1 / 72	2.9±2.0 / 7	21.7±4.2 / 28	6.7±2.7 / 12	20.9±5.4 / 34	5.5±1.9 / 10
- after 3 Myr	57.4±10.1 / 73	5.9±2.9 / 13	55.8±9.9 / 75	3.5±2.5 / 9	22.5±4.3 / 30	7.4±2.9 / 12	21.4±5.8 / 34	5.9±2.0 / 10
- after 4 Myr	55.5±9.5 / 70	7.4±2.9 / 14	54.2±9.3 / 71	4.1±2.8 / 9	22.1±4.2 / 29	7.5±2.37 / 12	20.6±5.7 / 32	6.3±2.1 / 11
- after 5 Myr	51.7±9.3 / 67	7.7±2.9 / 13	50.9±8.7 / 66	4.3±2.8 / 9	21.1±4.0 / 29	7.7±2.7 / 14	19.8±5.6 / 30	6.5±2.5 / 13
$m \geq 8.0$								
- after 1 Myr	1.0±0.7 / 3	0 / 0	0.7±1.0 / 3	0 / 0	0.2±0.3 / 1	0.2±0.6 / 2	0.3±0.6 / 2	0.2±0.6 / 2
- after 2 Myr	1.2±0.8 / 3	0.1±0.3 / 2	0.9±1.0 / 3	0 / 0	0.1±0.3 / 1	0.2±0.6 / 2	0.2±0.5 / 2	0.2±0.6 / 2
- after 3 Myr	1.3±0.7 / 3	0.1±0.4 / 2	1.2±1.0 / 3	0 / 0	0.2±0.4 / 1	0.2±0.6 / 2	0.2±0.5 / 2	0.2±0.6 / 2
- after 4 Myr	1.5±0.9 / 4	0.4±0.5 / 2	1.2±0.9 / 3	0.1±0.2 / 1	0.2±0.4 / 1	0.2±0.6 / 2	0.3±0.5 / 2	0.2±0.6 / 2
- after 5 Myr	1.6±0.9 / 4	0.4±0.5 / 2	1.3±1.1 / 4	0.3±0.5 / 2	0.3±0.6 / 2	0.2±0.6 / 2	0.4±0.8 / 3	0.2±0.6 / 2



## **B.2 2D-RW and WW candidates from S Mon and IRS 1/2**

Tables B.2, B.3, B.4 and B.5 provide information on the 2D-RW and WW candidates that can be traced back to either S Mon or IRS 1/2. The tables also include the non-3D trace-backs to both regions.

Table B.2: S Mon RW star 2D candidates sorted by decreasing 2D-velocity. Column 2+3: velocity in S Mon rest frame [rf]; Column 3: RV sources - <sup>a</sup>*Gaia* DR2, <sup>b</sup>Jackson et al. (2016), <sup>c</sup>Luo et al. (2019), <sup>d</sup>Kounkel et al. (2019), <sup>e</sup>Fehrenbach et al. (1992), <sup>f</sup>Duflot et al. (1995); Column 4: indication of 3D-candidate status; Column 5: minimum flight time since ejection (crossing of search boundary); Column 6: age from PARSEC isochrones (Bressan et al. 2012); Column 7–8: from literature sources - <sup>1</sup>Luo et al. (2019), <sup>2</sup>Venuti et al. (2015), <sup>3</sup>Venuti et al. (2017), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Kounkel et al. (2019), <sup>6</sup>Lamm et al. (2004), <sup>7</sup>Karlssohn (1972), <sup>8</sup>Cannon & Pickering (1993), <sup>9</sup>Voroshilov et al. (1985), <sup>10</sup>Paunzen et al. (2001), <sup>11</sup>Rebull et al. (2002), <sup>12</sup>Stassun et al. (2019), <sup>13</sup>Anders et al. (2019), <sup>14</sup>Venuti et al. (2018).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3134372235323909248	74.9 ±0.5	-58.2 ±4.7 <sup>c</sup>	no	-	2.3 <sup>+7.7</sup> <sub>-1.5</sub>	0.9–1.0 <sup>12,13</sup>	G5 <sup>1</sup>
3326642764222483328	72.8 ±0.8	-	-	0.1	9.0 <sup>+41.0</sup> <sub>-7.0</sub>	0.6–0.8 <sup>12,13</sup>	-
3353807577672866432	68.4 ±0.6	86.2 ±3.7 <sup>a</sup>	no	-	1.5 <sup>+1.5</sup> <sub>-1.0</sub>	1.1 <sup>13</sup>	K1 <sup>1</sup>
3326673142526903296	58.7 ±0.6	-	-	0.1	20.0 <sup>+15.0</sup> <sub>-18.0</sub>	1.0 <sup>12,13</sup>	-
3157444730917606656	46.3 ±2.6	-	-	0.9	20.0 <sup>+12.0</sup> <sub>-15.0</sub>	0.6 <sup>12</sup>	-
3326639882300319744	46.1 ±1.0	-17.1 ±2.9 <sup>b</sup>	no	-	4.0 <sup>+16.0</sup> <sub>-3.0</sub>	0.9–1.1 <sup>12,13</sup>	-
3331426086479602944	41.9 ±0.5	-	-	1.1	8.0 ±5.0	0.9–1.0 <sup>12,13</sup>	-
3134341650859166208	39.2 ±0.9	-	-	0.4	6.0 <sup>+25.0</sup> <sub>-4.0</sub>	0.9 <sup>12,13</sup>	-
3326519142179922688	37.7 ±0.8	-	-	0.3	12.0 <sup>+28.0</sup> <sub>-11.0</sub>	0.9–1.0 <sup>12,13</sup>	-
3326992103978048512	36.8 ±0.5	-16.3 ±3.9 <sup>a</sup>	no	-	5.0 <sup>+5.0</sup> <sub>-3.0</sub>	1.3–1.7 <sup>12,13</sup>	-
3328110440447873024	36.6 ±8.8	-	-	0.8	100.0 <sup>+100.0</sup> <sub>-98.5</sub>	-	-
3326736570603687168	33.6 ±0.6	105.7 ±2.9 <sup>b</sup>	no	-	30.0 <sup>+20.0</sup> <sub>-26.0</sub>	0.8–1.0 <sup>12,13</sup>	G7 <sup>6</sup>
3326595352079413888	32.8 ±0.5	-73.0 ±10.0 <sup>a</sup>	no	-	10.0 <sup>+5.0</sup> <sub>-8.5</sub>	1.1–1.2 <sup>12,13</sup>	F2 <sup>1</sup>
3357507846618517248	31.8 ±0.8	-	-	1.9	7.0 ±3.0	0.6–0.9 <sup>12,13</sup>	-
3355894515164624000	30.4 ±4.1	-	-	1.8	15.0 <sup>+20.0</sup> <sub>-12.0</sub>	0.4 <sup>12</sup>	-
3318797443817943040	27.8 ±0.6	-28.0 ±3.2 <sup>a</sup>	no	-	1.5 <sup>+2.5</sup> <sub>-1.0</sub>	1.2 <sup>13</sup>	-
3132210251867043200	23.9 ±0.7	-38.9 ±7.0 <sup>c</sup>	no	-	5.0 <sup>+7.0</sup> <sub>-3.0</sub>	1.0 <sup>13</sup>	G5 <sup>1</sup>
3133279939241456000	21.6 ±0.6	-50.7 ±5.3 <sup>a</sup>	no	-	9.0 <sup>+6.0</sup> <sub>-7.0</sub>	1.0 <sup>12,13</sup>	-
3328003654677416192	20.9 ±0.6	29.0 ±3.2 <sup>a</sup>	no	-	0.2 <sup>+0.8</sup> <sub>-0.1</sub>	1.1 <sup>13</sup>	-
3131085447170896128	13.0 ±0.6	34.9 ±3.4 <sup>a</sup>	no	-	5.0 ±2.0	1.1–1.5 <sup>12,13</sup>	-

Table B.3: S Mon WW star 2D candidates sorted by decreasing 2D-velocity. Column 2+3: velocity in S Mon rest frame [rf]; Column 3: RV sources - <sup>a</sup>*Gaia* DR2, <sup>b</sup>Jackson et al. (2016), <sup>c</sup>Luo et al. (2019), <sup>d</sup>Kounkel et al. (2019), <sup>e</sup>Fehrenbach et al. (1992), <sup>f</sup>Duflot et al. (1995); Column 4: indication of 3D-candidate status; Column 5: minimum flight time since ejection (crossing of search boundary); Column 6: age from PARSEC isochrones (Bressan et al. 2012); Column 7–8: from literature sources - <sup>1</sup>Luo et al. (2019), <sup>2</sup>Venuti et al. (2015), <sup>3</sup>Venuti et al. (2017), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Kounkel et al. (2019), <sup>6</sup>Lamm et al. (2004), <sup>7</sup>Karlssohn (1972), <sup>8</sup>Cannon & Pickering (1993), <sup>9</sup>Voroshilov et al. (1985), <sup>10</sup>Paunzen et al. (2001), <sup>11</sup>Rebull et al. (2002), <sup>12</sup>Stassun et al. (2019), <sup>13</sup>Anders et al. (2019), <sup>14</sup>Venuti et al. (2018).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3326731554082039296	27.3 ±0.6	-	-	0.2	20.0 <sup>+10.0</sup> <sub>-16.0</sub>	1.0 <sup>12,13</sup>	-
3326610298565496320	27.1 ±1.3	-	-	0.2	20.0 <sup>+25.0</sup> <sub>-17.0</sub>	0.7–0.9 <sup>12,13</sup>	-
3326670389452200832	23.0 ±0.9	-	-	0.2	50.0 <sup>+20.0</sup> <sub>-46.0</sub>	0.7–0.8 <sup>12,13</sup>	-
3331136468244152832	22.2 ±1.5	-	-	1.9	5.0 <sup>+25.0</sup> <sub>-3.0</sub>	0.7–1.0 <sup>12,13</sup>	-
3130519576640345344	21.5 ±1.5	-	-	3.7	1.5 <sup>+2.5</sup> <sub>-1.0</sub>	-	-
3326628230053022080	21.0 ±2.9	-	-	0.4	50.0 <sup>+50.0</sup> <sub>-46.0</sub>	0.6 <sup>12</sup>	-
3158160611766998400	20.7 ±0.7	-	-	1.9	4.8 <sup>+2.0</sup> <sub>-1.5</sub>	1.6–1.7 <sup>12,13</sup>	F0 <sup>1</sup>
3327753756299891328	20.0 ±0.6	-	-	0.7	20.0 <sup>+5.0</sup> <sub>-15.0</sub>	0.7–0.9 <sup>12,13</sup>	-
3161262853763313024	18.9 ±0.9	22.6 ±6.7 <sup>a</sup>	no	-	6.0 ±2.0	1.4–1.6 <sup>12,13</sup>	-
3134469026706163456	18.6 ±0.6	-	-	0.6	10.0 <sup>+20.0</sup> <sub>-7.5</sub>	1.1 <sup>13</sup>	-
3326641462848241408	17.6 ±0.6	-	-	0.4	20.0 <sup>+5.0</sup> <sub>-16.0</sub>	1.0–1.1 <sup>12,13</sup>	-
3326535012082791424	17.5 ±1.1	-	-	0.5	12.0 <sup>+50.0</sup> <sub>-10.0</sub>	0.5–0.7 <sup>12,13</sup>	-
3326641462848241792	17.0 ±0.6	-	-	0.3	7.0 <sup>+25.0</sup> <sub>-5.5</sub>	0.8–1.0 <sup>12,13</sup>	-
3350754336961977984	15.1 ±0.6	-	-	0.4	25.0 <sup>+10.0</sup> <sub>-11.0</sub>	0.9–1.0 <sup>12,13</sup>	-
3326170730136908032	15.1 ±1.8	-	-	1.2	2.0 ±0.2	2.2 <sup>13</sup>	A0/1 <sup>7</sup>
3350679806397436928	14.7 ±0.6	-	-	0.9	9.0 <sup>+3.0</sup> <sub>-6.0</sub>	1.0–1.2 <sup>12,13</sup>	-
3356837591202942976	14.3 ±0.6	-	-	4.9	5.0 <sup>+5.0</sup> <sub>-2.5</sub>	0.7–0.9 <sup>12,13</sup>	-
3350847554932397056	13.9 ±0.9	-	-	1.5	2.0 <sup>+2.0</sup> <sub>-1.0</sub>	0.5 <sup>12</sup>	-
3350758013453747328	13.8 ±0.5	-	-	0.8	20.0 <sup>+10.0</sup> <sub>-16.0</sub>	1.0 <sup>12,13</sup>	-
3326951215889632128	13.7 ±0.5	-	-	0.2	15.0 <sup>+5.0</sup> <sub>-12.0</sub>	1.2–1.5 <sup>12,13</sup>	-
3326626344563404672	12.9 ±0.7	-	-	0.6	32.0 <sup>+8.0</sup> <sub>-30.0</sub>	0.8 <sup>12,13</sup>	-
3326737189078970880	12.6 ±1.5	-	-	-0.1	2.5 <sup>+17.5</sup> <sub>-2.0</sub>	-	-
3327689606667717248	11.5 ±2.2	-	-	2.3	2.0 <sup>+5.0</sup> <sub>-1.8</sub>	0.7 <sup>12</sup>	-
3327310343876193536	11.5 ±0.7	-	-	2.7	4.0 <sup>+6.0</sup> <sub>-3.6</sub>	1.2–1.7 <sup>12,13</sup>	-
3326644138611656832	11.1 ±0.5	18.6 ±3.1 <sup>a</sup>	no	-	5.0 <sup>+6.0</sup> <sub>-4.6</sub>	1.1 <sup>13</sup>	-
3352137969269081472	10.6 ±0.6	8.8 ±5.8 <sup>c</sup>	no	-	3.0 <sup>+3.0</sup> <sub>-1.5</sub>	0.9 <sup>13</sup>	K7 <sup>1</sup>
3350774471768004736	10.0 ±0.5	-	-	0.7	1.0 <sup>+8.0</sup> <sub>-0.6</sub>	1.0 <sup>13</sup>	-
3159113274168209152	9.9 ±0.7	-	-	3.9	4.0 ±0.5	2.0–2.1 <sup>12,13</sup>	A2 <sup>8</sup>
3327841717229545856	9.9 ±0.5	16.9 ±6.4 <sup>a</sup>	no	-	7.0 ±3.0	1.5–1.9 <sup>12,13</sup>	F0 <sup>9</sup>
3355872628012737152	9.7 ±2.9	-	-	4.8	20.0 <sup>+25.0</sup> <sub>-16.5</sub>	0.5 <sup>12</sup>	-
3326637442758920960	9.7 ±0.6	-1.4 ±2.9 <sup>b</sup>	no	-	2.0 ±0.5	0.3 <sup>4</sup>	M3 <sup>2,4</sup>
3326495017346142592	8.2 ±1.5	-	-	1.7	0.4 <sup>+1.5</sup> <sub>-0.2</sub>	-	-
3133869616772191232	7.8 ±0.9	-	-	2.8	2.7 <sup>+2.4</sup> <sub>-1.4</sub>	0.6 <sup>12</sup>	-
3351891609942897664	7.7 ±0.6	-	-	2.6*	1.5 <sup>+0.8</sup> <sub>-0.5</sub>	1.0 <sup>13</sup>	-
3327882841540580224	7.6 ±1.9	-	-	4.4	0.5 <sup>+1.6</sup> <sub>-0.4</sub>	-	-

Table B.3 - continued

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Mass (M <sub>⊙</sub> )	Spectral type
3326894591041117696	6.8 ±1.1	-	-	0.7	2.0 <sup>+6.0</sup> <sub>-1.3</sub>	1.8 <sup>13</sup>	-
3326994406080500608	6.8 ±0.6	-	-	1.3	3.0 <sup>+18.0</sup> <sub>-1.3</sub>	0.9 <sup>13</sup>	-
3326492028048885120	6.6 ±2.2	-	-	1.9	2.0 <sup>+2.0</sup> <sub>-1.5</sub>	0.3 <sup>12</sup>	-
3326589025592180864	6.2 ±1.7	-	-	0.9	1.8 <sup>+6.2</sup> <sub>-1.6</sub>	-	-
3327008935952524032	6.1 ±0.6	-	-	0.9	3.5 <sup>+3.5</sup> <sub>-1.9</sub>	1.5–1.8 <sup>12,13</sup>	-
3326909670670846976	6.0 ±1.9	-	-	0.1	30.0 <sup>+120.0</sup> <sub>-26.0</sub>	0.6 <sup>12</sup>	-
3326707154372788480	5.9 ±1.9	-	-	0.3	3.5 <sup>+26.0</sup> <sub>-3.0</sub>	0.4 <sup>12</sup>	-
3351060477934502016	5.8 ±1.8	-	-	1.7	2.0 <sup>+2.0</sup> <sub>-1.6</sub>	-	-
3134320416539722368	5.8 ±1.7	-	-	2.7	2.0 <sup>+2.0</sup> <sub>-1.6</sub>	0.7 <sup>12</sup>	-
3326525975471926656	5.8 ±1.8	-	-	1.7	3.5 <sup>+17.0</sup> <sub>-2.1</sub>	0.4 <sup>12</sup>	-
3326629814897031040	5.6 ±3.6	-	-	1.4	1.2 <sup>+8.8</sup> <sub>-1.1</sub>	-	-
3326581191571739648	5.5 ±1.0	-	-	1.5	1.0 <sup>+4.3</sup> <sub>-0.7</sub>	-	-
3327852815425166592	5.4 ±0.7	-	-	3.2	9.0 <sup>+11.0</sup> <sub>-7.0</sub>	-	-
3326690661698141056	5.4 ±3.3	-	-	0.4	15.0 <sup>+85.0</sup> <sub>-14.0</sub>	0.8 <sup>12</sup>	-
3134443016384346496	5.2 ±1.4	-	-	2.9	2.0 <sup>+4.0</sup> <sub>-1.0</sub>	0.5 <sup>12</sup>	-
3326641428488508928	5.1 ±1.1	-	-	1.2	7.5 <sup>+22.5</sup> <sub>-4.5</sub>	0.2–0.5 <sup>4,12</sup>	M4 <sup>4</sup>
3326703688334117120	5.0 ±2.0	-	-	0.3	20.0 <sup>+75.0</sup> <sub>-18.0</sub>	0.6	-

Table B.4: IRS 1/2 RW star 2D candidates sorted by decreasing 2D-velocity. Column 2+3: velocity in respective IRS rest frame [rf]; Column 3: RV sources - <sup>a</sup>*Gaia* DR2, <sup>b</sup>Jackson et al. (2016), <sup>c</sup>Luo et al. (2019), <sup>d</sup>Kounkel et al. (2019), <sup>e</sup>Fehrenbach et al. (1992), <sup>f</sup>Dufflot et al. (1995); Column 4: indication of 3D-candidate status; Column 5: minimum flight time since ejection (crossing of search boundary); Column 6: age from PARSEC isochrones (Bressan et al. 2012); Column 7: Subcluster identification; Column 8–9: from literature sources - <sup>1</sup>Luo et al. (2019), <sup>2</sup>Venuti et al. (2015), <sup>3</sup>Venuti et al. (2017), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Kounkel et al. (2019), <sup>6</sup>Lamm et al. (2004), <sup>7</sup>Karlsson (1972), <sup>8</sup>Cannon & Pickering (1993), <sup>9</sup>Voroshilov et al. (1985), <sup>10</sup>Paunzen et al. (2001), <sup>11</sup>Rebull et al. (2002), <sup>12</sup>Stassun et al. (2019), <sup>13</sup>Anders et al. (2019), <sup>14</sup>Venuti et al. (2018).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Subcluster	Mass (M <sub>⊙</sub> )	Spectral type
3326689188524905344	64.2 ± 0.8	-1.3 ± 3.1 <sup>a</sup>	no	-	4.0 <sup>+8.0</sup> <sub>-3.5</sub>	IRS 1	1.0 <sup>13</sup>	G7 <sup>1</sup>
3326689188524905344	63.4 ± 0.7	-1.3 ± 3.1 <sup>a</sup>	no	-	4.0 <sup>+8.0</sup> <sub>-3.5</sub>	IRS 2	1.0 <sup>13</sup>	G <sup>1</sup>
3326639882300319744	47.2 ± 1.1	-16.8 ± 2.9 <sup>b</sup>	no	-	4.0 <sup>+16.0</sup> <sub>-3.0</sub>	IRS 2	0.9–1.1 <sup>12,13</sup>	-
3134341650859166208	40.5 ± 1.1	-	-	0.3	6.0 <sup>+25.0</sup> <sub>-4.0</sub>	IRS 2	0.9 <sup>13</sup>	-
3134341650859166208	39.1 ± 1.2	-	-	0.3	6.0 <sup>+25.0</sup> <sub>-4.0</sub>	IRS 1	0.9 <sup>13</sup>	-
3326519142179922688	39.0 ± 1.0	-	-	0.2	12.0 <sup>+28.0</sup> <sub>-11.0</sub>	IRS 2	0.9–1.0 <sup>12,13</sup>	-
3326626928678945408	38.1 ± 0.8	-	-	0.1	3.8 <sup>+26.2</sup> <sub>-3.0</sub>	IRS 2	0.7–0.9 <sup>12,13</sup>	-
3326519142179922688	37.6 ± 1.1	-	-	0.1	12.0 <sup>+28.0</sup> <sub>-11.0</sub>	IRS 1	0.9–1.0 <sup>12,13</sup>	-
3355468591847181440	37.1 ± 0.8	-	-	1.3	3.5 <sup>+2.5</sup> <sub>-1.5</sub>	IRS 2	1.0 <sup>13</sup>	-
3326626928678945408	36.9 ± 0.8	-	-	0.1	3.8 <sup>+26.2</sup> <sub>-3.0</sub>	IRS 1	0.7–0.9 <sup>12,13</sup>	-
3134452465312268032	32.9 ± 0.8	-	-	0.4	20.0 <sup>+10.0</sup> <sub>-18.0</sub>	IRS 2	1.1 <sup>12,13</sup>	-
3355801022316951552	32.7 ± 3.3	-	-	1.7	100.0 <sup>+100.0</sup> <sub>-98.0</sub>	IRS 1	0.6 <sup>12</sup>	-
3355801022316951552	31.6 ± 3.2	-	-	1.7	100.0 <sup>+100.0</sup> <sub>-98.0</sub>	IRS 2	0.6 <sup>12</sup>	-
3352821109587150592	26.4 ± 0.8	37.5 ± 7.0 <sup>a</sup>	no	1.8	4.5 <sup>+7.5</sup> <sub>-2.5</sub>	IRS 2	0.9–1.0 <sup>12,13</sup>	-
3133279939241456000	23.8 ± 0.7	-50.6 ± 5.3 <sup>a</sup>	no	-	9.0 <sup>+6.0</sup> <sub>-7.0</sub>	IRS 2	1.0 <sup>12,13</sup>	-
3133279939241456000	22.6 ± 0.9	-50.5 ± 5.3 <sup>a</sup>	no	-	9.0 <sup>+6.0</sup> <sub>-7.0</sub>	IRS 1	1.0 <sup>12,13</sup>	-
3328003654677416192	20.2 ± 0.9	29.1 ± 3.2 <sup>a</sup>	no	-	0.2 <sup>+0.8</sup> <sub>-0.1</sub>	IRS 1	1.1 <sup>13</sup>	-
3328003654677416192	18.9 ± 0.7	29.1 ± 3.2 <sup>a</sup>	no	-	0.2 <sup>+0.8</sup> <sub>-0.1</sub>	IRS 2	1.1 <sup>13</sup>	-
3326698289559319936	6.1 ± 0.7	-31.0 ± 11.7 <sup>a</sup>	no	-	10.0 <sup>+10.0</sup> <sub>-9.0</sub>	IRS 2	1.0–1.7 <sup>4,12,13</sup>	G5 <sup>2,4</sup>

Table B.5: IRS 1/2 WW star 2D candidates sorted by decreasing 2D-velocity. Column 2+3: velocity in respective IRS rest frame [rf]; Column 3: RV sources - <sup>a</sup>*Gaia* DR2, <sup>b</sup>Jackson et al. (2016), <sup>c</sup>Luo et al. (2019), <sup>d</sup>Kounkel et al. (2019), <sup>e</sup>Fehrenbach et al. (1992), <sup>f</sup>Duflot et al. (1995); Column 4: indication of 3D-candidate status; Column 5: minimum flight time since ejection (crossing of search boundary); Column 6: age from PARSEC isochrones (Bressan et al. 2012); Column 7: Subcluster identification; Column 8–9: from literature sources - <sup>1</sup>Luo et al. (2019), <sup>2</sup>Venuti et al. (2015), <sup>3</sup>Venuti et al. (2017), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Kounkel et al. (2019), <sup>6</sup>Lamm et al. (2004), <sup>7</sup>Karlsson (1972), <sup>8</sup>Cannon & Pickering (1993), <sup>9</sup>Voroshilov et al. (1985), <sup>10</sup>Paunzen et al. (2001), <sup>11</sup>Rebull et al. (2002), <sup>12</sup>Stassun et al. (2019), <sup>13</sup>Anders et al. (2019), <sup>14</sup>Venuti et al. (2018).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	3D-cand.	Flight time (Myr)	Iso. age (Myr)	Subcluster (M <sub>⊙</sub> )	Mass	Spectral type
3326691726849639168	23.8 ±2.0	-	-	0.0	6.0 <sup>+100.0</sup> <sub>-5.0</sub>	IRS 2	0.8 <sup>13</sup>	-
3326691726849639168	22.5 ±2.0	-	-	0.0	6.0 <sup>+100.0</sup> <sub>-5.0</sub>	IRS 1	0.8 <sup>13</sup>	-
3326641462848241792	18.6 ±0.7	-	-	0.2	7.0 <sup>+25.0</sup> <sub>-5.5</sub>	IRS 2	0.8–1.0 <sup>12,13</sup>	-
3350980149161701888	17.1 ±0.9	-16.6 ±6.4 <sup>a</sup>	no	-	10.0 <sup>+5.0</sup> <sub>-8.5</sub>	IRS 1	1.5 <sup>12,13</sup>	F0 <sup>1</sup>
3350980149161701888	16.4 ±0.7	-16.6 ±6.4 <sup>a</sup>	no	-	10.0 <sup>+5.0</sup> <sub>-8.5</sub>	IRS 2	1.5 <sup>12,13</sup>	-
3134369658343616640	15.4 ±0.8	-	-	1.2	4.0 <sup>+16.0</sup> <sub>-3.5</sub>	IRS 2	1.0–1.1 <sup>12,13</sup>	-
3134369658343616640	13.9 ±0.9	-	-	1.2	4.0 <sup>+16.0</sup> <sub>-3.5</sub>	IRS 1	1.0–1.1 <sup>12,13</sup>	-
3326644138611656832	12.6 ±0.7	18.7 ±3.1 <sup>a</sup>	no	-	5.0 <sup>+6.0</sup> <sub>-4.6</sub>	IRS 2	1.1 <sup>13</sup>	-
3326644138611656832	11.4 ±0.8	18.8 ±3.1 <sup>a</sup>	no	-	5.0 <sup>+6.0</sup> <sub>-4.6</sub>	IRS 1	1.1 <sup>13</sup>	-
3326637442758920960	10.9 ±0.8	-1.3 ±2.9 <sup>b</sup>	no	-	2.0 ±0.5	IRS 2	0.3 <sup>4</sup>	M3 <sup>2,4</sup>
3326637442758920960	9.8 ±0.9	-1.2 ±2.9 <sup>b</sup>	no	-	2.0 ±0.5	IRS 1	0.3 <sup>4</sup>	M3 <sup>2,4</sup>
3326704238089925120	9.3 ±0.6	-7.9 ±2.9 <sup>d</sup>	no	-	0.5 <sup>+3.0</sup> <sub>-0.4</sub>	IRS 2	0.6–1.0 <sup>4,13</sup>	M0 <sup>2,4</sup>
3326685512032888320	8.5 ±1.0	13.4 ±10.4 <sup>a</sup>	no	-	3.0 <sup>+3.0</sup> <sub>-1.5</sub>	IRS 1	1.2–2.1 <sup>4,12,13</sup>	F5/G0 <sup>4,6</sup>
3326480865429886720	8.5 ±2.1	-	-	1.4	2.0 <sup>+3.0</sup> <sub>-1.6</sub>	IRS 2	-	-
3326480865429886720	7.7 ±2.2	-	-	1.3	2.0 <sup>+3.0</sup> <sub>-1.6</sub>	IRS 1	-	-
3326495017346142592	6.9 ±1.6	-	-	1.8	0.4 <sup>+1.5</sup> <sub>-0.2</sub>	IRS 1	-	-
3326634212943546496	6.7 ±0.9	-18.5 ±5.0 <sup>b</sup>	no	-	2.0 <sup>+3.5</sup> <sub>-1.3</sub>	IRS 2	1.8–2.5 <sup>12,13</sup>	-
3327008935952524032	6.6 ±0.8	-	-	1.7	3.5 <sup>+3.5</sup> <sub>-1.9</sub>	IRS 1	1.5–1.8 <sup>12,13</sup>	-
3326629814897031040	6.6 ±3.7	-	-	0.5	1.2 <sup>+8.8</sup> <sub>-1.1</sub>	IRS 1	-	-
3134443016384346496	6.4 ±1.5	-	-	1.6	2.0 <sup>+4.0</sup> <sub>-1.0</sub>	IRS 2	0.5 <sup>12</sup>	-
3326589025592180864	6.0 ±1.8	-	-	0.5	1.8 <sup>+6.2</sup> <sub>-1.6</sub>	IRS 2	-	-
3326692620203367168	6.0 ±2.9	-	-	in cluster	6.0 <sup>+46.0</sup> <sub>-5.5</sub>	IRS 1	0.6 <sup>12</sup>	-
3326684549960233728	5.9 ±1.0	12.7 ±2.9 <sup>b</sup>	no	-	10.0 <sup>+25.0</sup> <sub>-9.0</sub>	IRS 1	0.9–1.1 <sup>12,13</sup>	-
3326492028048885120	5.7 ±2.2	-	-	1.5	2.0 <sup>+2.0</sup> <sub>-1.5</sub>	IRS 2	0.3 <sup>12</sup>	-
3326690661698141056	5.7 ±3.3	-	-	in cluster	15.0 <sup>+85.0</sup> <sub>-14.0</sub>	IRS 2	0.8 <sup>12</sup>	-
3350961560543556352	5.6 ±1.7	-	-	1.7	10.0 <sup>+38.0</sup> <sub>-8.5</sub>	IRS 1	-	-
3350961560543556352	5.1 ±1.6	-	-	1.6	10.0 <sup>+38.0</sup> <sub>-8.5</sub>	IRS 2	-	-
3326492028048885120	5.1 ±2.3	-	-	2.0	2.0 <sup>+2.0</sup> <sub>-1.5</sub>	IRS 1	0.3 <sup>12</sup>	-
3326589025592180864	5.1 ±1.9	-	-	0.4	1.8 <sup>+6.2</sup> <sub>-1.6</sub>	IRS 1	-	-
3326678193408399232	3.6 ±0.7	-16.2 ±4.5 <sup>a</sup>	no	1.5	2.8 <sup>+2.2</sup> <sub>-1.5</sub>	IRS 2	1.9	F0 <sup>1</sup>

### B.3 Past visitors to NGC 2264

Table B.6 provides information on sources that can be successfully traced back in 3D to NGC 2264 in the past 2/5 Myr (upper age limit for IRS 1/2 and S Mon respectively) but their position on the CAMD identifies them as MS stars or older pre-MS stars.

*Gaia* DR2 3327203588170236672 is a past visitor both in IRS 1 and IRS 2 but an ejected WW from S Mon. It is therefore older than IRS 1/2, but young enough to have originated from S Mon. The same applies to *Gaia* DR2 3134179335455713408, which is a past visitor to IRS 1, but is shown to have been ejected from S Mon.

Table B.6: Past visitors to NGC 2264 (3D trace-backs). Column 2+3: velocity in the respective NGC 2264 rest frame [rf]; Column 3: RV sources – <sup>a</sup>*Gaia* DR2; Column 4: age from PARSEC isochrones (Bressan et al. 2012); Column 5: from literature sources – <sup>1</sup>Voroshilov et al. (1985), <sup>2</sup>McCuskey (1959), <sup>3</sup>Cannon & Pickering (1993), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Karlssohn (1972); Column 6: Subcluster identification.

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Iso. age (Myr)	Spectral type	Subcluster
Visitors at RW velocities					
S Mon					
3128924155208294656	50.7 ±0.7	-36.3 ±4.7 <sup>a</sup>	~10-15	-	-
3326608919880182144	46.5 ±0.6	34.5 ±3.7 <sup>a</sup>	~6-12	B3/5 <sup>5</sup>	-
3127637722304864000	44.7 ±0.7	9.4 ±4.2 <sup>a</sup>	~8-12	-	-
3133406275699512320	37.8 ±0.6	22.4 ±3.4 <sup>a</sup>	~6-11	-	-
3129614889029245696	35.0 ±0.9	2.2 ±6.3 <sup>a</sup>	~?8-12	-	-
3129675293448445440	31.0 ±0.7	-14.9 ±5.3 <sup>a</sup>	~6-12	-	-
3157158583015446016	30.8 ±0.7	-16.7 ±7.7 <sup>a</sup>	~13-14	-	-
3131005736878997248	25.8 ±0.5	-27.4 ±3.0 <sup>a</sup>	~10-12	-	-
3352025681642829952	23.2 ±0.6	36.8 ±3.4 <sup>a</sup>	~8-11	-	-
3159544931266074880	21.2 ±0.9	21.3 ±6.3 <sup>a</sup>	~13-14	-	-
3355756320296166528	12.9 ±0.7	-29.7 ±7.9 <sup>a</sup>	~7-10	-	-
IRS 1 / IRS 2					
3128924155208294656	52.0 ±0.9	-35.9 ±4.7 <sup>a</sup>	~10-12	-	IRS 2
3128924155208294656	50.8 ±1.1	-35.7 ±4.7 <sup>a</sup>	~10-12	-	IRS 1
3357475892062058752	46.1 ±1.2	-17.9 ±6.3 <sup>a</sup>	~10-12	-	IRS 1
3357475892062058752	45.5 ±1.0	-17.9 ±6.3 <sup>a</sup>	~10-12	-	IRS 2
3127637722304864000	45.1 ±0.9	9.7 ±4.2 <sup>a</sup>	~8-12	-	IRS 2
3127637722304864000	44.1 ±1.1	9.8 ±4.2 <sup>a</sup>	~8-12	-	IRS 1
3133406275699512320	37.6 ±0.9	22.6 ±3.4 <sup>a</sup>	~6-11	-	IRS 2
3132201314034932864	37.3 ±1.0	-34.2 ±5.4 <sup>a</sup>	~3-11	-	IRS 1
3133406275699512320	36.6 ±1.1	22.7 ±3.4 <sup>a</sup>	~6-11	-	IRS 1
3132362529927264128	34.6 ±0.9	1.0 ±3.5 <sup>a</sup>	~6-10	-	IRS 2
3132362529927264128	33.4 ±1.1	1.1 ±3.5 <sup>a</sup>	~6-10	-	IRS 1
3129675293448445440	32.7 ±0.9	-14.7 ±5.3 <sup>a</sup>	~6-11	-	IRS 2
3129675293448445440	31.5 ±1.1	-14.6 ±5.3 <sup>a</sup>	~6-11	-	IRS 1
3129273692526220544	31.1 ±0.9	-50.0 ±3.7 <sup>a</sup>	~10-13	-	IRS 2
3352775922231300992	27.6 ±1.1	19.3 ±6.9 <sup>a</sup>	~8-12	-	IRS 2
3352025681642829952	24.4 ±1.1	36.8 ±3.4 <sup>a</sup>	~8-11	-	IRS 1



Table B.6 - continued

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Iso. age (Myr)	Spectral type	Subcluster
Visitors at WW velocities					
S Mon					
3326203616704192896	25.1 ±0.6	13.3 ±3.3 <sup>a</sup>	~5–11	-	-
3354418768701989248	18.8 ±0.7	-5.6 ±5.0 <sup>a</sup>	~8–12	-	-
3331208765431910528	18.6 ±0.8	22.9 ±5.5 <sup>a</sup>	~5–12	-	-
3160409491002912512	16.2 ±0.8	-9.5 ±5.1 <sup>a</sup>	~11	-	-
3158315471107784960	12.6 ±0.5	-27.2 ±3.2 <sup>a</sup>	~4–11	-	-
3325585416291697408	11.0 ±1.3	-8.3 ±16.4 <sup>a</sup>	~8–12	-	-
IRS 1 / IRS 2					
3326616410302784256	20.5 ±0.9	17.2 ±6.5 <sup>a</sup>	~10–12	-	IRS 2
3326616410302784256	19.5 ±1.1	17.3 ±6.5 <sup>a</sup>	~10–12	-	IRS 1
3134179335455713408	13.3 ±1.0	26.6 ±6.6 <sup>a</sup>	~3–10	-	IRS 2
3134179335455713408	11.9 ±1.2	26.7 ±6.6 <sup>a</sup>	~3–10	-	IRS 1
3327203588170236672	11.1 ±0.9	-7.1 ±5.8 <sup>a</sup>	~4–11	-	IRS 2
3326689807000188032	0.7 ±0.8	8.7 ±5.7 <sup>a</sup>	~2–10	F5 <sup>4</sup>	IRS 1
3326689807000188032	0.6 ±0.7	8.6 ±5.7 <sup>a</sup>	~2–10	F5 <sup>4</sup>	IRS 2

## B.4 Future visitors to NGC 2264

Tables B.7 and B.8 provide information on sources that can be successfully traced forward in 3D to NGC 2264 in the future 5–8 Myr (these values are based on the assumption of a lifetime of 10 Myr for each subcluster).

Table B.7: Future 3D-RW visitors to NGC 2264. Column 2+3: velocity in the respective NGC 2264 rest frame [rf]; Column 3: RV sources – <sup>a</sup>*Gaia* DR2; Column 4: age from PARSEC isochrones (Bressan et al. 2012); Column 5: from literature sources – <sup>1</sup>Voroshilov et al. (1985), <sup>2</sup>McCuskey (1959), <sup>3</sup>Cannon & Pickering (1993), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Cruzalèbes et al. (2019).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Iso. age (Myr)	Spectral type
Visitors at RW velocities				
S Mon				
3129232293342932864	41.4 ±0.7	30.0 ±5.0 <sup>a</sup>	~6–11	-
3352221841389893888	30.9 ±0.8	-18.5 ±9.4 <sup>a</sup>	~1	-
3158670342784515584	30.2 ±0.5	-2.3 ±4.8 <sup>a</sup>	~1–6	-
3157467988160982912	26.6 ±0.5	19.9 ±3.4 <sup>a</sup>	~13	-
3327867585815573376	26.3 ±0.5	54.6 ±3.2 <sup>a</sup>	~1–10	-
3131017483608422016	21.8 ±0.8	-41.5 ±4.6 <sup>a</sup>	~10	-
3328077695617951360	17.0 ±0.5	34.6 ±3.8 <sup>a</sup>	~2–10	-
3352147521276456320	16.1 ±0.6	28.1 ±7.0 <sup>a</sup>	~12–13	-
3133971085379278080	12.5 ±0.6	-29.6 ±5.0 <sup>a</sup>	~1–4	-
3324391273644856448	9.7 ±1.1	-33.3 ±14.3 <sup>a</sup>	~7–11	-
IRS 1				
3129990986425585152	49.7 ±0.8	-4.1 ±3.4 <sup>a</sup>	~1–5	-
3129232293342932864	42.1 ±0.9	29.9 ±5.0 <sup>a</sup>	~6–11	-
3327331883137624832	41.9 ±0.8	71.3 ±2.9 <sup>a</sup>	~0.5–1	-
3157240462272160128	37.3 ±0.9	8.5 ±6.6 <sup>a</sup>	~10–12	-
3132766016335294976	31.4 ±0.8	7.7 ±3.6 <sup>a</sup>	~13	-
3327867585815573376	26.3 ±0.8	54.9 ±3.2 <sup>a</sup>	~1–10	-
3356469564045105664	25.1 ±0.9	-18.9 ±3.1 <sup>a</sup>	~7–13	-
3131017483608422016	22.6 ±1.1	-41.5 ±4.6 <sup>a</sup>	~10	-
3324293520190299904	22.4 ±1.0	22.1 ±7.2 <sup>a</sup>	~10–12	-
3157132675772801280	19.4 ±0.9	33.1 ±5.9 <sup>a</sup>	~12–13	-
3134392335771085568	18.2 ±0.9	-31.2 ±3.1 <sup>a</sup>	~2–10	-
3351325357155320320	15.8 ±0.9	30.8 ±10.9 <sup>a</sup>	~10–14	-
3352147521276456320	15.6 ±0.9	28.4 ±7.0 <sup>a</sup>	~12–13	-
3326892250283336448	15.2 ±0.8	26.7 ±4.7 <sup>a</sup>	~4–5	F5/7 <sup>5</sup>
3351375835908945024	14.6 ±1.1	26.5 ±12.3 <sup>a</sup>	~9–11	-
3324391273644856448	11.3 ±1.3	-33.2 ±14.3 <sup>a</sup>	~7–11	-
IRS 2				
3129990986425585152	48.4 ±0.7	-4.1 ±3.4 <sup>a</sup>	~1–5	-
3129232293342932864	40.9 ±0.8	29.9 ±5.0 <sup>a</sup>	~6–11	-
3157240462272160128	36.2 ±0.8	8.5 ±6.6 <sup>a</sup>	~10–12	-
3132766016335294976	30.4 ±0.7	7.7 ±3.6 <sup>a</sup>	~13	-
3327867585815573376	27.5 ±0.6	54.8 ±3.2 <sup>a</sup>	~1–10	-
3134392335771085568	17.0 ±0.8	-31.2 ±3.1 <sup>a</sup>	~2–10	-
3352147521276456320	16.8 ±0.8	28.3 ±7.0 <sup>a</sup>	~12–13	-
3324391273644856448	11.4 ±1.2	-33.3 ±14.3 <sup>a</sup>	~7–11	-

Table B.8: Future WW visitors to NGC 2264. Column 2+3: velocity in the respective NGC 2264 rest frame [rf]; Column 3: RV sources – <sup>a</sup>*Gaia* DR2; Column 4: age from PARSEC isochrones (Bressan et al. 2012); Column 5: from literature sources – <sup>1</sup>Voroshilov et al. (1985), <sup>2</sup>McCuskey (1959), <sup>3</sup>Cannon & Pickering (1993), <sup>4</sup>Venuti et al. (2014), <sup>5</sup>Cruzalèbes et al. (2019), <sup>6</sup>McGinnis et al. (2018).

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Iso. age (Myr)	Spectral type
Visitors at WW velocities				
S Mon				
3330828020872121472	23.2 ±0.7	-3.4 ±6.1 <sup>a</sup>	~7–10	-
3324278745502788992	22.8 ±0.7	6.9 ±6.6 <sup>a</sup>	~9–11	-
3355407985568310912	21.7 ±1.5	-6.6 ±19.7 <sup>a</sup>	~11–13	-
3130189418211411712	20.9 ±0.8	-9.1 ±4.6 <sup>a</sup>	~9–11	-
3327020240307406464	19.9 ±0.5	-7.7 ±5.2 <sup>a</sup>	~10–11	-
3133552068369881984	15.7 ±0.5	17.8 ±3.6 <sup>a</sup>	~1–10	-
3327620642374686848	13.9 ±0.6	-23.8 ±6.8 <sup>a</sup>	~6–11	B9 <sup>1</sup>
3325784565337792896	11.3 ±0.5	9.9 ±4.7 <sup>a</sup>	~20	-
3351090405265460736	9.2 ±0.7	-7.5 ±14.5 <sup>a</sup>	~1–10	-
3158447073205530240	7.6 ±0.5	-4.2 ±4.6 <sup>a</sup>	~13	-
3327865528528223872	7.0 ±0.7	10.2 ±5.3 <sup>a</sup>	~1–5	-
IRS 1				
3157987649142772480	25.0 ±0.9	10.3 ±3.3 <sup>a</sup>	~5–10	-
3153460646235577728	24.9 ±1.0	-3.0 ±4.2 <sup>a</sup>	~5–12	-
3130189418211411712	21.8 ±1.0	-9.1 ±4.6 <sup>a</sup>	~9–11	-
3327020240307406464	19.9 ±0.8	-7.4 ±5.2 <sup>a</sup>	~10–11	-
3358899312943684992	19.2 ±0.9	6.1 ±3.2 <sup>a</sup>	~0.3–2	-
3158958277392687744	19.0 ±0.9	-9.7 ±5.4 <sup>a</sup>	~6–9	-
3355361462483283968	18.4 ±0.8	15.6 ±4.2 <sup>a</sup>	~10–11	-
3353006308573860736	18.0 ±0.8	-7.8 ±3.2 <sup>a</sup>	~0.3–2	-
3327106792487405696	17.2 ±0.9	12.5 ±6.9 <sup>a</sup>	~5–12	-
3133552068369881984	16.5 ±0.8	17.8 ±3.6 <sup>a</sup>	~1–10	-
3133854739004375040	14.6 ±0.8	2.1 ±10.4 <sup>a</sup>	~5–6	A5/7 <sup>5</sup>
3327963415126752896	12.9 ±0.9	17.7 ±3.5 <sup>a</sup>	~9–11	-
3330780157758632448	12.0 ±1.4	1.0 ±12.1 <sup>a</sup>	~11–13	-
3134016607734184832	11.5 ±1.1	-1.9 ±18.0 <sup>a</sup>	~11–13	-
3356244778336418176	9.1 ±2.1	-10.9 ±19.6 <sup>a</sup>	~5–10	-
3131647714228692224	8.7 ±1.6	-25.0 ±14.1 <sup>a</sup>	~10–11	A2 <sup>1</sup>
3129542694918227072	8.1 ±1.4	18.9 ±12.2 <sup>a</sup>	~10–12	-
3153947145772918016	8.1 ±1.1	-5.7 ±7.1 <sup>a</sup>	~14	-
3132900332851633024	7.6 ±0.9	-2.0 ±3.3 <sup>a</sup>	~10–11	-
3331424505931639680	7.1 ±1.1	-11.2 ±7.0 <sup>a</sup>	~5–8	-
3351516195439957632	7.1 ±0.9	2.6 ±4.3 <sup>a</sup>	~4–10	-
3133726581481375872	5.5 ±0.8	11.4 ±3.1 <sup>a</sup>	~2–5	-
3330679453658587904	5.5 ±1.0	2.6 ±7.4 <sup>a</sup>	~10–12	-
3326687878559500288	2.0 ±0.9	9.8 ±13.4 <sup>a</sup>	~0.5–13	K0 <sup>6</sup>

**Table B.8** - continued

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )	Iso. age (Myr)	Spectral type
IRS 2				
3324278745502788992	24.3 ±0.8	7.0 ±6.6 <sup>a</sup>	~9–11	-
3157987649142772480	24.2 ±0.8	10.3 ±3.3 <sup>a</sup>	~5–10	-
3153460646235577728	23.7 ±0.9	-3.0 ±4.2 <sup>a</sup>	~5–12	-
3327020240307406464	21.2 ±0.7	-7.5 ±5.2 <sup>a</sup>	~10–11	-
3358899312943684992	20.2 ±0.8	6.0 ±3.2 <sup>a</sup>	~0.5–2	-
3355361462483283968	19.4 ±0.7	15.5 ±4.2 <sup>a</sup>	~10–11	-
3133854739004375040	13.8 ±0.8	2.1 ±10.4 <sup>a</sup>	~5–6	-
3325784565337792896	13.5 ±0.7	10.0 ±4.7 <sup>a</sup>	~20	-
3330780157758632448	13.1 ±1.3	1.0 ±12.1 <sup>a</sup>	~11–13	-
3130581905206339200	10.6 ±1.1	17.7 ±8.7 <sup>a</sup>	~10	A0 <sup>5</sup>
3134016607734184832	10.3 ±1.0	-1.9 ±18.0 <sup>a</sup>	~11–13	-
3356244778336418176	10.2 ±2.0	-11.0 ±19.6 <sup>a</sup>	~5–10	-
3331411036914202368	8.6 ±1.1	-1.6 ±10.7 <sup>a</sup>	~9–11	-
3331424505931639680	8.5 ±0.9	-11.2 ±7.5 <sup>a</sup>	~5–8	-
3326046042943212544	8.0 ±0.9	0.5 ±10.4 <sup>a</sup>	~6–11	-
3351516195439957632	7.0 ±0.7	2.5 ±4.3 <sup>a</sup>	~4–10	-
3132900332851633024	6.3 ±0.8	-2.1 ±3.3 <sup>a</sup>	~10–11	-
3158447073205530240	5.2 ±0.7	-4.1 ±4.6 <sup>a</sup>	~13	-
3326687878559500288	2.7 ±0.7	9.7 ±13.4 <sup>a</sup>	~0.5–13	-

## B.5 Red giants visiting NGC 2264

Table B.9 provides information on sources that can be successfully traced to NGC 2264 in 2D and/or 3D, however, are located on or near the red giant branch and are therefore at the end of their stellar evolution. I exclude these stars from the main CAMDs in our analysis, but state these here explicitly as some of these stars show large errors that could potentially turn them into viable younger candidates.

Table B.9: 2D trace-backs to NGC 2264 that are located on or near the red giant branch. Column 2+3: velocity in the respective NGC 2264 rest frame [rf]; Column 3: RV sources – <sup>a</sup> *Gaia* DR2.

<i>Gaia</i> DR2 source-id	2D-velocity rf (km s <sup>-1</sup> )	Radial velocity rf (km s <sup>-1</sup> )
S Mon		
3128637079593202048	34.0 ±0.6	179.9 ±2.9 <sup>a</sup>
3133974899309720960	28.0 ±0.5	63.5 ±3.5 <sup>a</sup>
3352423666192288896	18.2 ±0.6	-37.7 ±4.8 <sup>a</sup>
3326929874196650496	11.3 ±0.5	3.9 ±4.2 <sup>a</sup>
IRS 1		
3324925468797373184	73.3 ±0.8	36.7 ±2.9 <sup>a</sup>
3331183854621309696	41.2 ±0.9	26.7 ±2.9 <sup>a</sup>
3128637079593202048	33.5 ±0.8	180.2 ±2.9 <sup>a</sup>
3131775150199250176	29.8 ±0.8	26.9 ±2.9 <sup>a</sup>
3134350554328161152	17.0 ±1.1	-52.4 ±2.9 <sup>a</sup>
3133698303416819712	16.5 ±1.0	-5.6 ±2.9 <sup>a</sup>
3351991940380416000	14.9 ±1.2	4.2 ±2.9 <sup>a</sup>
3326646788607684480	10.8 ±0.9	-8.6 ±2.9 <sup>a</sup>
IRS 2		
3324925468797373184	73.0 ±0.7	36.7 ±2.9 <sup>a</sup>
3128637079593202048	35.1 ±0.8	180.1 ±2.9 <sup>a</sup>
3132441802143684864	49.0 ±0.7	-
3133974899309720960	28.5 ±0.8	63.7 ±3.5 <sup>a</sup>
3134350554328161152	17.8 ±0.9	-52.5 ±2.9 <sup>a</sup>
3133698303416819712	17.1 ±0.9	-5.6 ±2.9 <sup>a</sup>
3351991940380416000	14.3 ±1.0	4.2 ±2.9 <sup>a</sup>
3326646788607684480	11.3 ±0.8	-8.7 ±2.9 <sup>a</sup>

# Appendix C

## List of Publications

1. **Schoettler, C.**, Parker, R.J., Arnold, B., Grimmett, L.P., de Bruijne, J. and Wright, N.J., 2019. Dynamical evolution of star-forming regions: III. Unbound stars and predictions for Gaia. *Monthly Notices of the Royal Astronomical Society*, 487(4), pp.4615-4630
2. **Schoettler, C.**, de Bruijne, J., Vaher, E. and Parker, R.J., 2020. Runaway and walkaway stars from the ONC with Gaia DR2. *Monthly Notices of the Royal Astronomical Society*, 495(3), pp.3104-3123
3. **Schoettler, C.** and Parker, R.J., 2021. Double trouble: Gaia reveals (proto)-planetary systems that may experience more than one dense star-forming environment. *Monthly Notices of the Royal Astronomical Society:Letters*, 501(1), pp.L12–L17
4. Parker, R.J. and **Schoettler, C.**, 2022. Constraints on star formation in NGC 2264. *Monthly Notices of the Royal Astronomical Society*, 510(1), pp.1136-1147
5. **Schoettler, C.**, Parker, R.J. and de Bruijne, J., 2022. Constraining the initial conditions of NGC 2264 using ejected stars found in Gaia DR2. *Monthly Notices of the Royal Astronomical Society*, 510(3), pp.3178-3206