Multiplicity among the cool supergiants in the Magellanic Clouds

R. Dorda[1,2]⋆ and L. R. Patrick [1,2,3]

[1]*Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[2]*Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain*
[3]*Departamento de Física Aplicada, Facultad de Ciencias, Universidad de Alicante, E-03690 San Vicente del Raspeig, Alicante, Spain*

ABSTRACT
The characterization of multiplicity of high-mass stars is of fundamental importance to understand their evolution, the diversity of observed core-collapse supernovae and the formation of gravitational wave progenitor systems. Despite that, until recently, one of the final phases of massive star evolution – the cool supergiant phase – has received comparatively little attention. In this study, we aim to explore the multiplicity among the cool supergiant (CSG) population in the Large and Small Magellanic Clouds (LMC and SMC, respectively). To do this we compile extensive archival radial velocity (RV) measurements for over 1000 CSGs from the LMC and SMC, spanning a baseline of over 40 yr. By statistically correcting the RV measurements of each stellar catalogue to the *Gaia* DR2 reference frame we are able to effectively compare these diverse observations. We identify 45 CSGs where RV variations cannot be explained through intrinsic variability, and are hence considered binary systems. We obtain a minimum binary fraction of 15 ± 4 per cent for the SMC and of 14 ± 5 per cent for the LMC, restricting our sample to objects with at least 6 and 5 observational epochs, respectively. Combining these results, we determine a minimum binary fraction of 15 ± 3 per cent for CSGs. These results are in good agreement with previous results which apply a correction to account for observational biases. These results add strength to the hypothesis that the binary fraction of CSGs is significantly lower than their main-sequence counterparts. Going forward, we stress the need for long-baseline multi-epoch spectroscopic surveys to cover the full parameter space of CSG binary systems.

Key words: binaries: general – stars: massive – supergiants – Magellanic Clouds.

1 INTRODUCTION

Multiplicity among high-mass stars[1] has been studied observationally by many authors, and the general conclusion is that almost all of these stars are born with one or more companions. For example Sana et al. ([2012](#)) estimate that 75 per cent of O-type stars are born with a companion, and most of these will interact at some point during their lives (see also Kobulnicky et al. [2014](#); Sana et al. [2014](#)). Indeed taking into long period systems and B-type stars Moe & Di Stefano ([2017](#)) argue that 80–90 per cent of all high-mass stars will interact with a companion. This interaction may have profound effects on their evolution (De Mink et al. [2013](#)), the nature of their subsequent supernova explosions (Podsiadlowski, Joss & Hsu [1992](#); De Marco & Izzard [2017](#); Hirai et al. [2020](#)), the formation of mergers and compact-object binaries (Marchant et al. [2017](#)), and potentially in the generation of gravitational waves (Belczynski, Kalogera & Bulik [2002](#)). Stellar mergers and mass transfer may also have an important role in shaping the observed colour–magnitude diagrams of young clusters (Beasor et al. [2019](#); Britavskiy et al. [2019](#); Wang et al. [2020](#)).

Evolved high-mass stars such as cool supergiants[2] (CSGs) are typically considered as a key constraint for single-star evolutionary models (Ekström et al. [2013](#)). In addition, CSGs are the immediate progenitors of most core-collapse supernovae and consequently, compact objects. Given their importance to stellar feedback and gravitational wave progenitors (Langer et al. [2020](#)), understanding the multiplicity properties of this stage of evolution is vital to comprehend the observed distribution of core-collapse supernovae and gravitational wave progenitors. Moreover, the comparison of the multiplicity in the main sequence and in the CSG phase is important to characterize the different evolutionary pathways when binary interaction is taken into account (Wang et al. [2020](#)).

Despite the key role of CSGs in these topics, their multiplicity properties have been comparatively neglected, compared with OB-type stars. Burki & Mayor ([1983](#)), compiled multi-epoch radial velocities (RVs) for 181 F–M supergiants in the Milky Way (of which only 25 are M-type, and not all of them are high-mass stars) and estimated a binary fraction of 31–38 per cent. Neugent, Levesque & Massey ([2018](#)) and Neugent et al. ([2019](#)) explored the single epoch spectra of 749 CSGs from the MCs, M31, and M33 and found evidence for the presence of a B-type spectrum in 87 CSGs of their

---

⋆ E-mail: ri.dorda@gmail.com
[1]Those that are massive enough to end their lives as core-collapse supernovae, which require initial masses higher than 8–$10\,M_\odot$ depending on metallicity (Doherty et al. [2015](#)).

[2]The term red supergiant (RSG) is commonly used because most of these stars in the Galaxy are M-type. However, in lower metallicity galaxies their types are significantly earlier (late-G or early-K; Dorda et al. [2016](#)). Thus, the more general term cool supergiants (CSGs) is adopted here.

sample. Later, Neugent et al. (2020) used the photometry of the LMC RSGs with a B-type component to train a kNN-algorithm. By applying this algorithm to their CSG photometric sample, they estimated a binary fraction of $19.5^{+7.6}_{-6.7}$ per cent. Patrick et al. (2019, 2020) studied multi-epoch RVs of CSGs in both the 30 Doradus region of the LMC (Patrick et al. 2019), 14 stars, and in the cluster NGC 330 of the SMC (Patrick et al. 2020), 9 stars. They obtained a binary fraction of 30 ± 10 per cent in both samples, in reasonable agreement with Burki & Mayor (1983), but significantly lower than is found in main-sequence high-mass stars. Patrick et al. (2019, 2020) argued that this difference is expected, as binary interactions would have prevented the evolution of their components towards regular CSGs in many systems. However, their results are based on the study of a small number of CSGs and they must be tested on larger samples.

Multi-epoch RV studies that aim to identify binary motion of CSGs are challenging as the sizes of these stars (several hundreds of solar radii) impose lower limits to the orbital periods (P_{orb}) of several hundreds–thousands of days, depending on temperature and luminosity. Furthermore, they may exhibit intrinsic RV variations as large as $10 \, \mathrm{km \, s^{-1}}$ (Josselin & Plez 2007; Gray 2008; Kravchenko et al. 2019), due to their large-scale convective cells (Schwarzschild 1975; Chiavassa et al. 2009; Stothers 2010) that complicate the detection of orbital velocities. In this paper, we explore the largest multi-epoch sample of CSGs to date, almost 1000 stars from both MCs, taking advantage of the large spectroscopic catalogues of CSGs available in the literature over the last 40 yr. Although the observation frequency is low, observations are separated by years or even decades, allowing the study of long-period CSGs.

In Section 2, we present the catalogues used, their characteristics and how we combined them into a homogeneous sample. The method for the identification of binary systems is presented in Section 3, and we discuss the results and their implications in Section 4.

2 THE COOL SUPERGIANT SAMPLE

2.1 A cool supergiant reference catalogue

To build our sample, we searched the literature for large catalogues focused on MC CSGs. We compile measurements from six large catalogues, described below, because to correct RVs from each catalogue of systematic effects requires large samples (see Section 2.3).

We collected all the data from each catalogue, specifically including both CSGs and other stars, to improve the homogenization statistics (see Section 2.3). Then we cross-matched all the targets, creating a list of unique stars. We tagged the targets as CSGs or non-CSGs using the spectral type (SpTs) classifications provided by the catalogues. In some catalogues, a photometric selection criteria was used, filtered by the measured RV, rather than a spectral classification. Therefore, whenever a discrepancy was found between these catalogues and others in which spectral classification was performed, we used the identification of the latter. In addition, we used the information available in SIMBAD database (from sources as OGLE; Soszyński et al. 2009; Soszyński et al. 2011) to identify Cepheids and AGB-type stars and to tag them as non-CSGs. The final catalogue has 1 200 unique stars confirmed as CSGs.

2.1.1 Prevot et al. (1985) and Mauron et al. (1987)

These two papers are part of a series of 8 articles that studied different samples of stars from the Southern hemisphere with the photoelectric scanner CORAVEL. Prevot et al. (1985) provide RV measurements for 404 F- to M-type supergiants in the direction of the LMC, while Maurice et al. (1987) did the same for 233 F- to M-type stars in the direction of the SMC (and the surrounding region). These observations were performed over a baseline of several years, from 1981 to 1984 in the case of the LMC, and to 1985 in the case of the SMC data. Additionally, these authors provided more than one epoch for a number of targets. Both Prevot et al. (1985) and Maurice et al. (1987) provided SpTs for their respective samples and used their RV measurements to confirm membership. The average RV errors are 1.5 and $1.3 \, \mathrm{km \, s^{-1}}$, for the LMC and SMC, respectively.

2.1.2 Massey & Olsen (2003)

Massey & Olsen (2003) presented spectroscopic observations for a large number of red stars in the direction of both MCs; 118 towards the SMC and 167 towards the LMC. The targets were observed during three nights (4–6 October) of 2001, where only one epoch is provided for all targets. These authors identified CSGs (the vast majority of targets) through spectral classification and RV confirmation. An average error of $0.3 \, \mathrm{km \, s^{-1}}$ is provided for all observations in this sample.

2.1.3 Neugent et al. (2012, 2013)

In these two papers, Neugent et al. observed large samples of candidate CSGs from the LMC (close to 2000 targets; 2012) and yellow supergiants from the SMC ($\sim$500 targets; 2010). The SMC data were observed during one run of five nights in 2009 October and LMC targets were observed during one run of eight nights in 2011 January. Thus, only one epoch is provided for the stars of each galaxy. In both studies, these authors do not provide SpTs for their targets, but determined their nature indirectly, using RV measurements to determine membership. Each target has different a quality flag that determines the RV errors. The average value for these errors are $2.7 \, \mathrm{km \, s^{-1}}$ for the SMC and $4.2 \, \mathrm{km \, s^{-1}}$ for the LMC.

2.1.4 González-Fernández et al. (2015)

González-Fernández et al. (2015, GDN15 thereafter) presented multi-epoch observations between 2010 and 2013 for 544 unique stars in the SMC (315 CSGs) and 289 in the LMC (229 CSGs). Initially, observations for both galaxies were taken in one three-night run in 2010 August, and SMC targets were observed again in a two-night run in 2011 July. An additional epoch for each galaxy (this time including a large amount of candidate CSGs in addition) was observed in 2012 July and 2013 November for the SMC and LMC, respectively. These authors identified CSGs through SpT classification, and confirmed their nature by RV measurements. A 1σ average value of $1 \, \mathrm{km \, s^{-1}}$ is provided for all targets.

2.2 Generic catalogues

We cross-matched the list previously obtained from the CSGs catalogues, including both, CSGs and non-CSGs, with two large generic (i.e. not focused on CSGs) catalogues:

2.2.1 RAVE-DR5

RAVE-DR5 observed 454 of our stars (108 CSGs), and a significant number of them over multiple epochs, between 2004 December and 2012 November. They provided individual RV measurements for
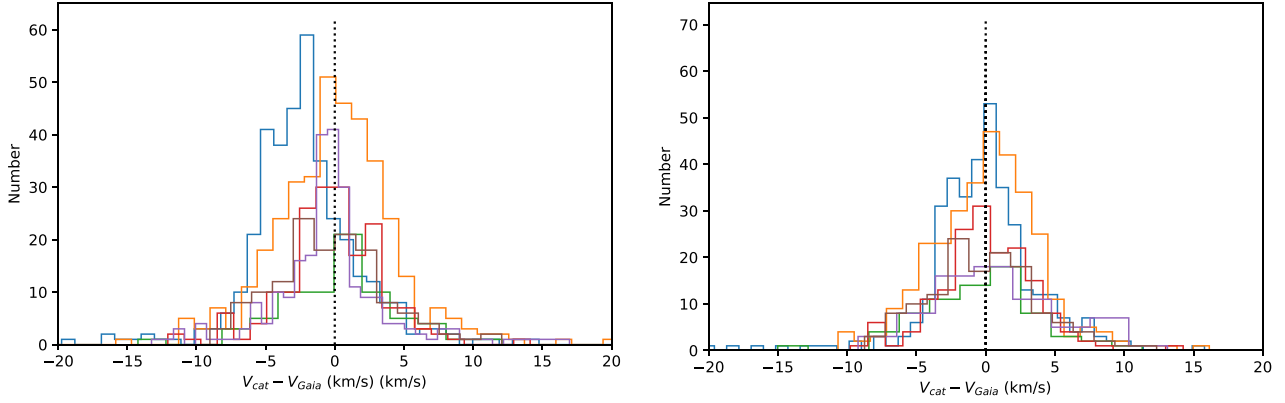
Figure 1. Distribution of differences in RV between the sources of each catalogue and their counterparts in *Gaia* DR2 for the LMC sample. The colour indicates the catalogue: orange for Prevot et al. (1985), brown for Massey & Olsen (2003), purple for RAVE-DR5, blue for Neugent et al. (2012), green and red for GDN15 data from 2010 and 2013, respectively. *Left*: The whole catalogues (including CSGs and non-CSGs), before the correction. *Right:* The CSGs from each catalogue, after the correction has been applied.

each epoch and individual errors for the RV measurements, with an average value of $1.8 \, \mathrm{km \, s^{-1}}$.

2.2.2 Gaia-DR2

Gaia DR2 (Gaia Collaboration 2018) provided RV measurements for 1436 stars from our sample, including 916 of our CSGs (92 per cent of the CSGs in our sample). The current data release (DR2) provide only one RV value for each star, which is the averaged value of all the measurements obtained during the almost 2 yr from 2014 July, until 2016 May. The RV uncertainty of any binary may be greatly overestimated (Katz et al. 2019), as the dispersion of the measures is increased by the RV variations due to orbital motion and/or intrinsic variations. In fact, Katz et al. (2019) showed that the distribution of RV uncertainties for *Gaia* (their fig. 10) peaks between 1 and $2 \, \mathrm{km \, s^{-1}}$. Above $2 \, \mathrm{km \, s^{-1}}$ there is a long tail composed by 'a mix of stars with insufficient signal to be processed in *Gaia* DR2, large amplitude variables and undetected binary or multiple systems'.

In our sample we expect both effects (intrinsic variations and orbital motions). Thus, we may expect overestimated uncertainties for our sample. In fact, 11 per cent of our CSGs have RV uncertainties in the range between 2 and $12 \, \mathrm{km \, s^{-1}}$, with an average of $3.1 \, \mathrm{km \, s^{-1}}$. This will affect to our determination of what is a significant RV variation in Section 3.1. In order to not overestimate the error of those CSGs with large RV variations, we truncate the uncertainties for our sample. The value chosen is the third quartile for the distribution of RV uncertainties for *Gaia*: $2.08 \, \mathrm{km \, s^{-1}}$, which delimits the beginning of the tail of overestimated uncertainties. This change only affects to 90 of our CSGs, which is only the upper decile.

2.3 Homogenization of the radial velocities

Combining RVs from different sources has the problem that there may be systematic differences between individual catalogues, as result of the instruments and methods used in each one. Therefore, it is necessary to homogenize the RV measurements to compare the different epochs of each CSG. To do this, we use *Gaia* DR2 as reference framework. For each catalogue (including also RAVE-DR5) and each galaxy, we calculated the RV differences between their targets and their corresponding counterparts in *Gaia* DR2.

As explained in Section 3, CSG have intrinsic RV variations. However, as these variations are random and we use a large sample, the average offset should be zero in absence of systematic differences. This statistical method works better with a larger sample size, as the error of the median decreases with $1/\sqrt{n}$. As a consequence, we decided to include the non-CSGs in each catalogue to calculate this correction. Most of these stars are evolved intermediate-mass stars from the MCs (AGBs or carbon stars), or foreground red stars (Galactic dwarfs or halo giants). We show the differences of the stars from each catalogue with respect to their *Gaia* counterparts in the left-hand panels of Figs 1 and 2.

The median value of each distribution of RV differences indicates the systematic deviation of that catalogue with respect to *Gaia*. Thus, we applied the median value of each distribution as a correction factor for all the sources in the corresponding catalogue and galaxy. These values are provided in Table 1.

Once the correction is applied, we remove all the non-CSGs from our samples, as well as those CSGs with only one epoch. We show the corrected samples of CSGs in the right-hand panels of Figs 1 and 2. The result is a multi-epoch-RV sample of almost 1000 stars confirmed as CSGs, 303 in the SMC, and 693 in the LMC.

3 ANALYSIS

In this section, we analyse the constructed sample with the aim of detecting binary motion. To do this, we used two complementary methods to maximize the information obtained from our data. First, we study the largest RV variations of each star, looking for those that can unambiguously be identified as caused by binary motions and not by intrinsic variations. Secondly, exploiting our long-baseline observations, we analyse radial velocity curves (RVCs) for out targets with the largest number of epochs, where we aim to identify RV variability that cannot be explained by intrinsic variability. Finally, we also cross-matched our catalogue with previous works where CSGs with spectral features suggesting they have an OB companion were reported.

3.1 Dynamical binary detection

To identify significant RV variation as a result of binary motion we use similar criteria adopted in several studies (Sana et al. 2013;
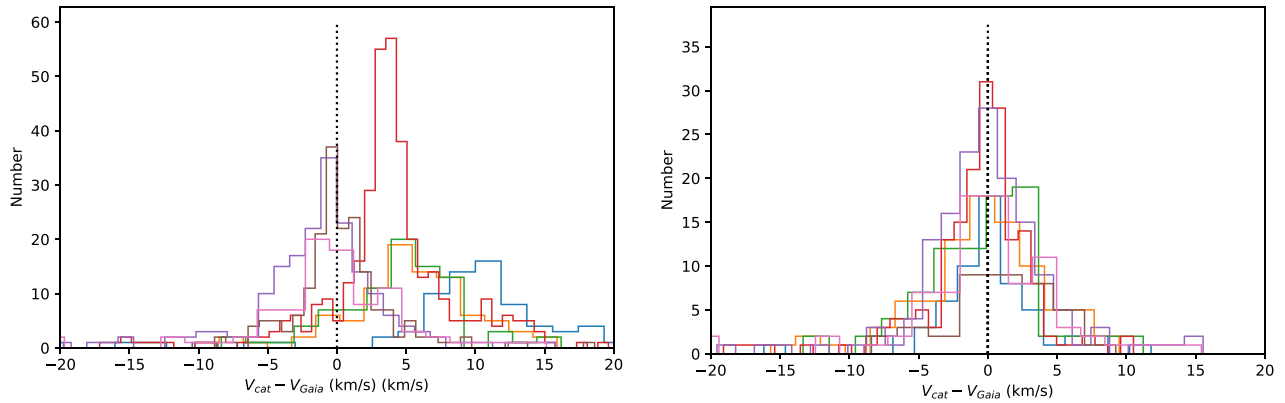
Figure 2. Distribution of differences in RV between the sources of each catalogue and their counterparts in *Gaia* DR2 for the SMC sample. The colour indicates the catalogue: purple for Maurice et al. (1987), pink for Massey & Olsen (2003), brown for RAVE-DR5, blue for Neugent et al. (2010), orange, green and red for GDN15 data from 2010, 2011, and 2012, respectively. *Left*: The whole catalogues (including CSGs and non-CSGs), before the correction. *Right*: The CSGs from each catalogue, after the correction has been applied.

Table 1. Correction factor for the RVs of each catalogue, with its corresponding error (calculated using $1.253 \cdot \sigma/\sqrt{n}$, as the correction factor is a median). We also show the sample size in each case (stars of each catalogue in common with *Gaia*-DR2).

Catalogue	Galaxy	Correction ($km\,s^{-1}$)	Median standard error ($km\,s^{-1}$)	Number of stars used
Maurice et al. (1987)	SMC	− 0.8	±0.5	186
Massey & Olsen (2003)	SMC	− 0.3	±0.7	88
RAVE-DR5	SMC	− 0.0	±0.3	188
Neugent et al. (2010)	SMC	10.1	±1.8	77
GDN15 2010 survey	SMC	5.4	±0.6	90
GDN15 2011 survey	SMC	5.0	±0.6	87
GDN15 2012 survey	SMC	3.9	±0.4	369
Prevot et al. (1985)	LMC	0.1	±0.3	375
Massey & Olsen (2003)	LMC	− 0.3	±0.4	137
RAVE-DR5	LMC	− 0.5	±0.3	248
Neugent et al. (2012)	LMC	− 2.4	±0.7	383
GDN15 2010 survey	LMC	0.8	±0.6	77
GDN15 2013 survey	LMC	− 0.1	±0.3	181

Dunstall et al. 2015; Patrick et al. 2019), which aimed at identifying binary motion in high-mass stars at different evolutionary stages. These authors use a two-staged approach in which both criteria must be simultaneously satisfied: the first (equation 1) determines the minimum change in RV that can be considered as a result of binarity; the second (equation 2) assesses the significance (at the 3σ level) of such a variation. Mathematically

$$|v_i - v_j| > \Delta V_{\rm lim}, \qquad (1)$$

$$\frac{|v_i - v_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}} > 3, \qquad (2)$$

where v_i and v_j are the RVs measured in any two epochs, and σ_i^2 and σ_j^2 are their respective uncertainties. $\Delta V_{\rm lim}$ is the limit imposed on the absolute velocity difference of any two measurements. By solely applying criterion 2, we find that 37 per cent (112) and 26 per cent (180) of CSGs have significant RV variation in the SMC and LMC, respectively. The results are displayed in Fig. 3, which shows the fraction of systems that simultaneously meet both variability criteria as a function of $\Delta V_{\rm lim}$.

To determine a valid $\Delta V_{\rm lim}$, intrinsic RV variations and uncorrected instrumental effects must be considered. It is well known that CSGs present intrinsic RV variations, due to the large convective cells that appear in their atmospheres (Schwarzschild 1975; Chiavassa et al. 2009; Stothers 2010). The observational studies (Josselin & Plez 2007; Gray 2008; Stothers 2010; Kravchenko et al. 2019) show that RSGs can exhibit intrinsic variations as large as $10\,km\,s^{-1}$. This value can be considered as an upper limit, because larger variations are expected for larger CSGs (Stothers 2010) and the amplitude varies irregularly between cycles (e.g. see Kravchenko et al. 2019). Thus, typical RV variations for most CSGs are significantly below that limit. In addition, our sample consists largely of G- and K-type stars (smaller in average than the M supergiants from the Galaxy for which the upper limit of $10\,km\,s^{-1}$ was measured). In consequence, we expect the average $\Delta V_{\rm lim}$ to be much smaller than this.

Despite the fact that we have no more than eight epochs for any target, we have the advantage of the size of our sample to study the multiplicity of CSGs statistically. We examine the distribution of the RV variability of our sample (Fig. 3) and find three distinct regions:
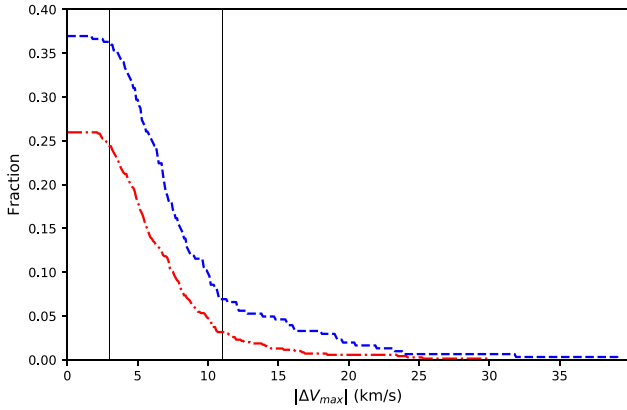
Figure 3. Cumulative distribution of the largest significant (equation 2) $|\Delta V_{max}|$ of each CSG. Red-dash–dotted line is for the LMC sample, whereas blue-dashed line is for the SMC sample. These distributions include only those CSGs with at least one significant (according to equation 2) value of $|\Delta V|$, but the fraction is calculated over the total number of CSG with multi-epoch data from the corresponding galaxy. The black vertical lines mark the limits at 11 km s^{-1}, cutting the lines at 0.07 (SMC, 21 CSGs) and 0.03 (LMC, 22 CSGs), and at ~3 km s^{-1} (see text for the details).

(i) Below ~3 km s^{-1}, the distribution is truncated because most RV variations below this value are not significant according to equation (2).

(ii) Between 3 and 11 km s^{-1}, the slope is approximately constant, with only a small number of CSGs showing variations larger than that. This region is dominated by three factors: intrinsic variations, binary motions, and uncorrected errors. While binary systems are only a fraction of the population, whose identification also depends strongly on the system inclination, intrinsic variations are expected for all CSGs. Also, any individual error not corrected by our statistical method is expected to be small (given the systematic errors found and corrected in Table 1), and thus their effects will have much more weight here than among those few CSGs with RV variations above 11 km s^{-1}.

(iii) Above 11 km s^{-1}, the flattening of the curve indicates that it is no longer affected by the ubiquitous intrinsic variations. Therefore, the only explanation for these variations is orbital motion.

By adopting $\Delta V_{lim} = 11$ km s^{-1} in equation (1), we get a reliable sample of binaries. With this limit, we find that 21 CSGs from the SMC (6.9 ± 1.4 per cent[3] of the total sample), and 22 (3.2 ± 0.7 per cent) of the CSGs from the LMC can be considered trustworthy binaries.

3.2 Detection through RV curves

For most of our sample, we have only two or three epochs available (see Table 3). However, a small fraction of the sample present RVCs whose behaviour can be analysed. This analysis is based on two factors: the RV variation (i.e. if the variability resembles orbital motion) and the time-scale of such variation. Our objective is to search for RV measurements that cannot be explained by intrinsic variations, but by orbital motions only. We note that no attempt

is made to estimate orbital parameters, owing to limitation in the observational data.

All RVCs with more than 4 epochs are analysed and for those that delineate a regular curve, we estimate a rough limit for the lowest period possible. In this sense we can, in principle, distinguish between intrinsic variations and orbital motion. The intrinsic variability of CSGs has typical periodicities on two different scales: (i) main periods are in the range between few hundred to one thousand days and (ii) longer secondary periods up to 5000 d (13.7 a; Kiss, Szabó & Bedding 2006; Chatys et al. 2019). Although these periods were derived from photometric information, the extensive following of the RSG μ Cep shows that its light and RV curves are linked: they have the same periodicity with a phase shift (Kiss et al. 2006; Jorissen, Van Eck & Kravchenko 2016; Kravchenko et al. 2019). Thus, these periods can be compared with those derived from our RVCs.

The minimum orbital period (P_{min}) for a CSG is limited by its stellar radius (see Appendix A1), as any orbit smaller than that would lead to a merger or to a mass transfer that would interrupt the normal CSG evolution in most cases (Eldridge, Izzard & Tout 2008). We calculate P_{min} for CSGs of different initial masses under the most optimistic circumstances (zero eccentricity), for different mass ratios (q). We detail the calculations in Appendix A1. We find that for any mass in the range between 9 and 25 M$_\odot$, P_{min} is below 14 a. Although not all CSGs present a long secondary period, and only 14 per cent (33 out of the 227 LMC CSGs studied by Chatys et al. 2019) have periods longer than 3000 d, we use 5000 d as a limit, to increase the reliability given our sparse sampling. Thus, we discard RVCs whose shapes suggest a period below 14 a.

Eight targets were initially selected with potential periods above 14 a. Most of the RVCs selected exhibit a linear trend over a decade or more, suggesting periods of several decades. However, given the sparse sampling of the RVCs, such trends can be caused by the combination of intrinsic variations of shorter periods over larger baselines. To evaluate this effect, we performed a Monte Carlo simulation where for each star we simulated 10^6 random sinusoidal RVCs, with phases and periods determined randomly as follows. For those CSGs in Chatys et al. (2019) with a known period for their intrinsic variability, we used random periods within ± 10 per cent of the nominal period. For the rest, we selected randomly intrinsic-variability periods in the range 160–2000 d. These period limits are chosen based on the distribution of measured periods in Chatys et al. (2019). For each simulated RVC, we extract RV measurements at the same epochs as the target stars taking into account its corresponding uncertainty,[4] and for the obtained data we calculated the absolute value of the Pearson's correlation coefficient. We compared these coefficients with those obtained from the observational data. We calculate the probability of having a random alignment with a correlation coefficient equal or higher than the observed one. To consider that the correlation from the RVC observed is not product of a random alignment of data points caused by intrinsic variations, we require that the probability calculated is lower than a certain threshold. We chose a significance level of 0.27 per cent (3σ for an equivalent Gaussian distribution) for these comparisons, as it is equivalent to the significance level used in the method explained in the previous subsection. In other words, we accepted as non-random those

---

[3]These fractions can be considered as a binomial distribution. Thus, their uncertainties can be calculated as $\sqrt{F(1 - F)/N_s}$, where N_s is the size of the sample and F is the fraction of positives in the sample (N_p/N_s).

[4]We randomly calculated the RV value by drawing it from a Gaussian distribution centred in the obtained RV value, and with a sigma equal to the corresponding uncertainty.

Table 2. CSGs in our sample that were reported to have BSSFs in previous works. The fractions are calculated respect the number of stars observed at least in two epochs, but without taking into account any CSG located in a cluster, as their BSSFs may be caused by contamination.

Origin	CSGs with BSSFs #	Observed at least in 2 epochs #	Signficiant by equation (2) #	In clusters #	Identified as binary in this work # (per cent)
From GDN15	16	14	10	2	4 (33)
From Neugent et al. (2019)	24	18	11	0	3 (17)
Total (unique CSGs^a)	38	30	19	2	7 (25)

^{a}There are two CSGs in common between these two catalogues.

RVCs that have a probability of change alignment of lower than 0.27 per cent.

Using this methodology, we find two CSGs (both from the SMC) whose RVCs cannot be explained via a combination of random fluctuations and periodic intrinsic variability with typical periods in the range 160–2000 d. We therefore assume these trends to result from genuine orbital motion. We provide the discussion for each individual case and we show the corresponding RVC in Appendix B.

3.3 CSGs with blue-star spectral features

A number of CSGs in the MCs were reported in previous works as having spectral features of early-type components (blue-star spectral features; BSSFs). Considering these targets as confirmed binaries is risky. The probability of random alignment with a blue star is not negligible, particularly in the case of those located in crowded star-forming regions. However, they can be considered as strong candidates. In this section, we analyse the cross-match of such CSGs with our own catalogue, which is also summarized in Table 2.

GDN15 reported 12 CSGs in whose spectra they found also the spectrum of a OB-type star at their bluest wavelengths. Also, these authors reported 12 CSGs with Balmer lines in emission. However, these lines may be caused by nebular emission. Thus, we checked in detail these spectra, as they are available in Dorda et al. (2018). We have tentatively confirmed the composite nature of 4 of these stars (identified by its catalogue name), obtaining a total of 16 CSGs with BSSFs:

[M2002] 23463: Emission is observed in H β, H γ, and H δ in the 2010 and 2012 data, with evidence for a slight blue asymmetry to H β. Furthermore there is a hint of an underlying absorption component in H δ which may suggest the presence of a B or Be star as a companion. This star is also the list presented of RSG+B-type stars for the SMC presented in Neugent et al. (2019). Unfortunately, we cannot confirm for sure if it is the same Be star as they identified, because they did not provide the ID of that star.

YSG010: This star has strong narrow Balmer lines in emission that may be nebular in origin. However, there is clear evidence for a blue continuum in the spectrum, and we have positively identified He I emission lines at 4009 and 4026 Å for example, implying the secondary is a B-type star. This system is discussed further in Section 4.3.

[M2002] 55355: Both H γ and H δ are visible as absorption lines. There are no clear signs of emission.

RM2-093: The spectrum is peculiar having many emission lines in addition to the Balmer series. A number of these are coincident with the positions of allowed Fe II lines and may indicate that the companion is an emission line B-type star. Further data are necessary to confirm the nature of this system.

Neugent et al. (2019) reported 24 CSGs from the MCs with spectral features from an OB-type star (Balmer's lines too strong for a CSG). Two of these stars are in common with the list from GDN15 (both of them with significant RV variations below 11 km s^{-1}).

In total, from the 38 CSGs with BSSFs reported in literature, we have at least two epochs for 30 of them, among which 11 have no significant RV variations (according to equation 2). This is not unexpected, as 9 of them have only two or three epochs available. The other two may be explained by high inclinations or long orbital periods, as none of them seems to be located close to any potential contaminant.

Among the 19 CSGs with significant RV variations, 10 have RV variations between 4.6 and 10.2 km s^{-1}, and their RVCs do not exhibit clearly a binary behaviour. There are also two CSGs which are located in clusters, and thus, their BSSFs are possibly caused by contamination from a unrelated blue star. The rest (seven CSGs) are considered binaries by the dynamical criterion, and we checked that none of them are located close to any potential contaminant. Therefore, the BSSFs observed in these CSGs belong to their companions, which must to be relatively luminous OB-type stars.

4 DISCUSSION

We have found 45 binary systems among the sample of CSGs studied (see Table C1). Moreover, 41 of them, were previously unknown. In this section, we explore the implications of this finding.

4.1 A lower limit for the binary fraction

The binary fraction (BF) is the fraction of stars from a given population (CSGs in this case) in a binary (or multiple) system. In this work, we find that 7.6 per cent of the CSGs studied in the SMC and 3.2 per cent in the LMC are reliable binaries (see Table C1 for the list of binaries found). These values are significantly lower than those obtained in previous works, but it is not unexpected. Our method is affected by a number of biases, such as inclination, sampling, orbital parameters, etc. Previous works (Patrick et al. 2019, 2020) corrected their samples of observational biases using Monte Carlo simulations, estimating the intrinsic BF from their observed fraction. However, to perform such corrections an orbital-parameter distribution, must be assumed. Observationally, the orbital parameter distribution for CSGs is not well constrained. Because of this (Patrick et al. 2019, 2020) extrapolated the orbital parameter distributions from earlier evolutionary phases (OB stars) (Moe & Di Stefano 2017). In this study, we have a sample large enough to estimate the minimum binary fraction (BF$_{min}$) expected for CSGs with an acceptable statistical significance without assuming any orbital parameter distribution. Therefore, our results can be used as independent constraints for

Table 3. CSGs split by number of RV epochs, and binaries detected among them (number and percentage).

	LMC sample			SMC sample		
Number of epochs	CSGs # (per cent[a])	Binaries # (per cent)	Accumulated (≥# of epochs) # (per cent)	CSGs # (per cent[a])	Binaries # (per cent)	Accumulated (≥# of epochs) # (per cent)
2	404 (58)	5 (1.2)	22 (3)	122 (40)	2 (1.6)	23 (8)
3	173 (25)	7 (4)	17 (6)	63 (21)	6 (10)	21 (12)
4	65 (9)	3 (5)	10 (9)	23 (8)	2 (9)	15 (13)
5	32 (5)	4 (12)	7 (14)	23 (8)	2[b] (9)	13 (14)
6	16 (2.3)	2 (13)	3 (16)	54 (18)	8[b] (15)	11 (15)
7	3 (0.4)	1 (33)	1 (33)	16 (5)	3 (19)	3 (17)
8	–	–	–	2 (0.7)	0 (0)	0 (0)
All	693	22 (3.2)	–	303	23 (7.6)	–

[a]Calculated over the total number of CSGs.

[b]We have taken into account here one extra binary found through the RVC method (see Section 3.2).

evolutionary models for binary populations. Also, in this way our results can be used to check, in an independent manner, the results obtained by Patrick et al. (2019, 2020). In fact, it is encouraging that our result for the SMC (7.6 ± 1.5 per cent) is compatible at 2σ with the observed BF (10 per cent) from Patrick et al. (2020) for the CSG in NGC 330, a cluster located in that galaxy, when the limit $\Delta V_{\mathrm{lim}} = 11\,\mathrm{km\,s^{-1}}$ is used in both cases.

Our main bias is the very limited number of epochs. As we show in Table 3, the maximum number of epochs is 7 and 8 for the samples from the LMC and the SMC, respectively, and only a few stars reach that level of sampling. In addition, a significant fraction of both samples (58 per cent in the LMC and 40 per cent in the SMC) have only two epochs. As can be deduced from Table 3, to find more binaries (by any of the methods used here) is more likely when more epochs are available, but our sample is clearly dominated by those with a very low number of epochs. Fortunately, we can minimize the main bias in our sample, without need of theoretical assumptions or simulations, by restricting the CSG sample that we consider to those objects with a given minimum number of epochs. However, the drawback in this scenario is that as we increase the minimum number of epochs, fewer stars are included in the minimum BF calculation, and thus, the uncertainties (see Section 3.1) will be larger. Given the fractions found previously, we decided to assume a maximum uncertainty for the percentage of 5. This condition is fulfilled in the LMC sample using those CSGs observed in five or more epochs (51 stars), while in the SMC is fulfilled using those with six or more epochs (72 stars). Under these limits, we found 11 binary CSGs in the SMC and 7 in the LMC, which represent 15 ± 4 and 14 ± 5 per cent, respectively, of the CSGs observed at least the number of epochs selected in each case.

Finally, we considered if there is any reason to not to combine the results from both galaxies. The differences in typical T_{eff} ($\sim300\,\mathrm{K}$; Tabernero et al. 2018) between the SMC and the LMC, affect to the average sizes of the stars from both galaxies. However, these differences are small and, therefore, the P_{min} for the CSGs from both galaxies is very similar (see Fig. A1). As we can expect systems in a very wide range of periods, this small difference seems unlikely to affect in a noticeable way the abundance of binaries. In the absence of a strong theoretical argument against, together with the fact that the BF_{min} calculated for both galaxies from the subsamples selected are statistically indistinguishable, makes us to think that we can use the BF of both MCs combined (15 ± 3 per cent), as global BF_{min} for CSGs. This value agrees well with the bias-corrected results obtained by Patrick et al. (2019, 2020) and Neugent et al. (2020) for samples from both galaxies.

4.2 Companions

4.2.1 Expected orbital biases

A detailed characterization of each binary system will be performed in a future paper. Here, we analyse the sample statistically, in order to understand its biases. The dynamical binary detection method only detects systems with a RV variation larger than $11\,\mathrm{km\,s^{-1}}$, which limits the detectable semi-amplitude velocity (K_1) of the systems, and in turn limits the periods and mass ratios detectable by this method. We note that this limit on K_1 does not affect binaries found through the RVC method alone. However, such stars represent only ~4 per cent of binaries found and as a consequence they do not significantly affect the following analysis.

In the most optimistic scenario, the measured RV variation is equal to the peak-to-peak velocity, which is twice K_1 by definition. However, given our limited sampling (less than 10 epochs per target), the $|\Delta V_{\mathrm{max}}|$ measured in most cases is not a reliable estimate of the true binary peak-to-peak velocity (Simón-Díaz et al., submitted). In most cases the peak-to-peak velocity will be greater than the measured $|\Delta V_{\mathrm{max}}|$.

To quantify this effect, we performed a Monte Carlo simulation where 10^6 synthetic RVCs were generated (with an unit amplitude for simplicity, and thus a peak-to-peak value of 2). We then took a number of RV measurements (n) from each curve, where n is randomly determined for each curve using as a probability distribution the sum by rows of the columns named 'Binaries' in Table 3. By calculating the differences between the values obtained from each curve, we obtain $|\Delta V_{\mathrm{max}}|$. Putting together the $|\Delta V_{\mathrm{max}}|$ obtained for every synthetic RVC, we obtain a distribution of values between 0 and 2 (the minimum and maximum values possible, respectively, for the $|\Delta V_{\mathrm{max}}|$ measured in the synthetic RVCs). Despite the most probable value in the distribution is 2, the median value is only 1.62. This simple experiment confirms that the measured $|\Delta V_{\mathrm{max}}|$ is, in general, not a reliable estimate of the peak-to-peak velocity of our stars. In consequence, if we consider that the peak-to-peak amplitudes in our sample can be typically as low as the limit $|\Delta V_{\mathrm{max}}| = 11\,\mathrm{km\,s^{-1}}$, we would be too optimistic. Statistically, using as lower limit $|\Delta V_{\mathrm{max}} = 11\,\mathrm{km\,s^{-1}}$ for binary identification implies that we are imposing a limit to the typical peak-to-peak amplitude significantly higher than $11\,\mathrm{km\,s^{-1}}$. We can estimate a more realistic peak-to-peak value using the ratio between the peak-to-peak amplitude in the synthetic RVCs and the median value obtained: $1.62/2 = 0.81$. Thus, we should expect that in median the $11\,\mathrm{km\,s^{-1}}$ limit is equivalent to a limit in the peak-to-peak velocity corresponding of $11/0.81 = 13.6\,\mathrm{km\,s^{-1}}$. Due to that, the expected lower limit value for K_1 in the of observed

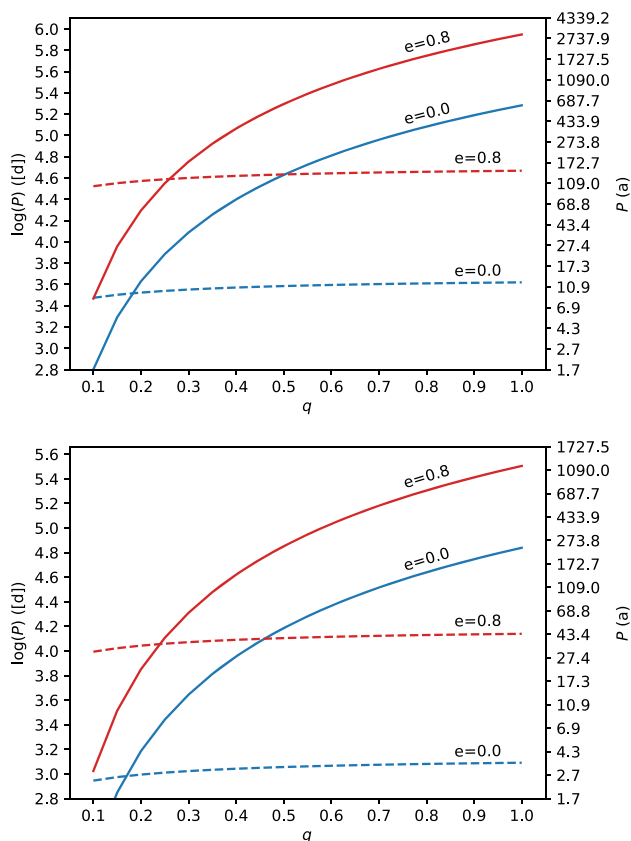**Figure 4.** $P_{orb}-q$ plane for CSGs from the LMC. The colour of the lines indicates the eccentricity used for their calculation. We show the results for two extreme eccentricities, 0.0 (in blue) and 0.8 (in red). Solid lines indicate the limit imposed by $K_1^{min} = 6.8 \, km \, s^{-1}$ in the $P-q$ plane: any binary CSG detected by our method should be at the right of these lines. Dashed lines indicate the minimum periods imposed by the size and the eccentricity, for given a mass (the lower panel is for $9 \, M_\odot$, whereas upper panel is for $25 \, M_\odot$). For the SMC, the figure is equivalent, but with minimum periods 0.1 dex lower.

binary systems is $6.8 \, km \, s^{-1}$. Take into account that this limit is under the assumption of $i = 90°$. However, an inclination $i < 90°$ implies that the real peak-to-peak velocity is even greater than the $|\Delta V_{max}|$ observed, and thus that K_1 is also even higher. In consequence, the inclination does not affect to the lower limit on K_1 calculated.

K_1 is related to the orbital period (P_{orb}), the orbital eccentricity (e), and the mass-ratio ($q = M_{companion}/M_{CSG}$). A lower limit to K_1 defines for each eccentricity an area in the $P - q$ diagram where systems cannot be detected by our method. In Fig. 4, we show the limits imposed by the minimum K_1 on orbital configurations for CSGs in the LMC as solid lines. The detectable parameter space (for a given eccentricity) is located below these solid lines. The minimum period possible, imposed by the size of the CSGs is also shown, which, as explained in Section 3.2, depends on stellar radius, mass, and eccentricity (see Appendix A1 for details). The dependence on eccentricity is important as, for a given q, a higher eccentricity implies a shorter periapsis separation and the minimum separation is limited by the physical size of the CSG.

The combination of P_{min} and minimum semi-amplitude velocity implies a number of biases for the binary systems detected in our sample. We can detect q values down to 0.20, over the entire CSG

mass range (see Fig. 4), in the most optimistic circumstances ($e \sim 0$) and down to $q \sim 0.25$ at $e = 0.8$. However, low-q systems can only be detected for short orbital periods (as a result of the minimum semi-amplitude velocity cut-off). High-q systems can be detected in a wider range of periods. This results in a favourable detection bias towards higher q systems. In addition, systems with short periods can be detected at a wider range of q's. This biases our detections towards shorter periods. We must consider also that high-eccentricity systems are less likely to be detected by our method than those with a lower eccentricity. A system with a K_1 large enough to be detected but also a high eccentricity will display a detectable RV difference only during a small fraction of its orbit. Given our limited sampling, the detection probability for high-eccentricity systems is significantly lower than for low-eccentricity systems. Therefore, we expect to be biased towards low and intermediate eccentricities.

Moe & Di Stefano (2017) studied the frequency distribution for main-sequence OB-type stars (the progenitors of CSGs). For those with periods >20 d the companion frequency peaks at $\log (P_{orb}) \sim 3.5$ dex (d) (see their fig. 37), with a strong preference for lower q's ($q \sim 0.2$–0.3), and a eccentricity distribution weighted towards larger values. This implies that our RV limit of $11 \, km \, s^{-1}$ misses a significant part of the binary population: those systems with a combination of a low q and a long orbital period. Combining the observational biases together with the expected distribution of parameters, we can conclude that our sample is likely dominated by companions with intermediate q's (higher than 0.20), and periods around $10^{3.5}$ d. As those systems with mid to high eccentricity have P_{min} significantly above 3.5 dex (d), probably most of them interacted or merged. Only those few with periods high enough would have survived. Thus, we expect that systems with low eccentricities are significantly more frequent than those with high eccentricities.

The minimum q expected is ~ 0.20 regardless of the CSG mass. As most of our sample are mid-to-high luminosity CSGs (see e.g. González-Fernández et al. 2015; Tabernero et al. 2018), the lowest mass expected is roughly $10 \, M_\odot$. Therefore, the companions of these stars can have masses as low as $2 \, M_\odot$. We must take into account that companions with masses lower than $\sim 3 \, M_\odot$ (later than B9 V) have not entered in the main sequence by the time lowest mass CSGs explode as supernova (Neugent et al. 2019). For a typical CSG of $15 \, M_\odot$, the lowest mass companions should have masses around $3 \, M_\odot$. Therefore, we conclude that our method is able to detect any O or B companion, although late-B types cannot be detected in all cases, only in low-mass CSGs. In addition, we expect CSGs plus compact object systems to be detectable. Such objects are remnants of the original primary star (Kruckow et al. 2016; Klencki et al. 2021) of the system, while the CSG was the original secondary. Neutron stars have a highest theoretical mass of $2.16 \, M_\odot$, which is just slightly higher than the lowest mass detectable by our method. Thus, it is very unlikely to have detected them in this study. On the other hand, black holes can have higher masses, and thus they can be present among the companions detected. However, to identify the nature of the companions in the systems found, the orbital properties of the systems must be constrained in future works.

4.2.2 Comparison with single-epoch surveys.

The method proposed by Neugent et al. (2018) uses the identification of BSSFs in single-epoch spectra to detect binary CSGs. They obtained that 4 per cent of their MC sample present BSSFs. This value is significantly below our estimated minimum BF of 15 ± 3 per cent, but we must take into account that we are comparing different
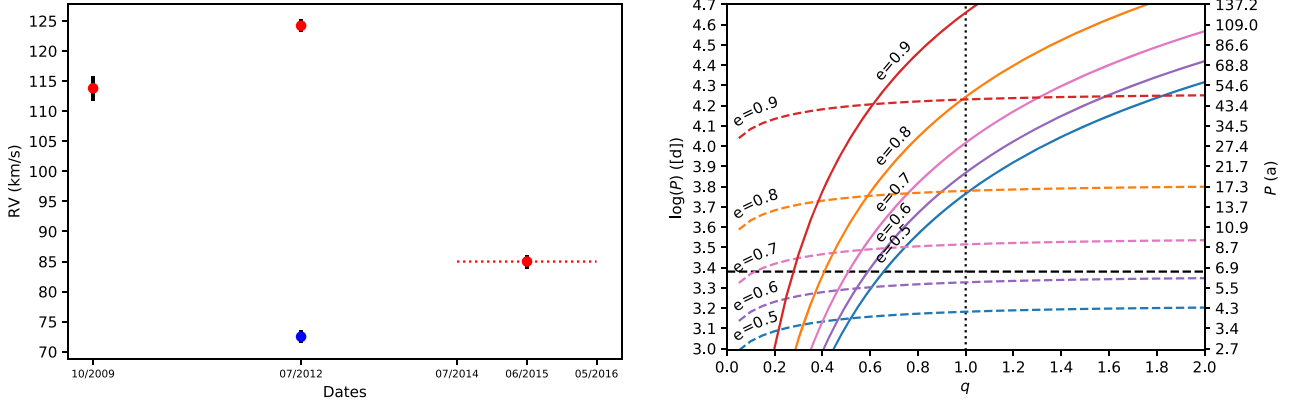
Figure 5. *Left* (5a): RVC of YSG010. Red circles are the epochs available, while the blue circle is the RV obtained for the blue companion. The red dotted line shows the time-span of *Gaia* observations. Whenever the error bars are not visible is because they are smaller than the circles. *Right* (5b): Study of the P_{orb} possible according to the estimated size of YSG010. Each coloured line indicates a given eccentricity. Dashed lines mark the P_{min} given its estimated size. Solid lines indicate the limit imposed by assuming $K_1 = 19.6$ km s^{-1} (see text). The black-dashed horizontal line indicates $P = 6.6$ a, the minimum P_{orb} derived from our observations. The dotted vertical line indicates $q = 1$.

methods. To compare these results we must consider the fraction of CSGs with BSSFs in our sample. We cannot use our whole sample for this calculation, as it has not been searched entirely looking for BSSFs. Instead, we use the sample from González-Fernández et al. (2015), composed by 544 CSGs for which they reported 16 CSGs with BSSFs (including the two located in clusters). We also take into account 10 CSGs in that list with BSSFs reported by Neugent et al. (2019). These 26 CSGs imply a BF of 4.6 ± 0.9 per cent, which is fully compatible with the results obtained by Neugent et al. (2019). In addition, our value is compatible with the bias-corrected BF ($19.5^{+7.6}_{-6.7}$ per cent) obtained by Neugent et al. (2020) for the LMC. All this confirms that the difference between the results is not intrinsic to the samples considered, but to the different methodologies used.

According to Neugent et al. (2018), the coolest OB companion they considered (15 000 K, ~B5 V), can be concealed easily by a relatively warm CSGs (see the yellow-supergiant spectrum in their Fig. 7). Moreover, they found that their warmest CSGs can conceal the contribution of even the earliest-type main-sequence B-type stars. On the contrary, as explained in Section 4.2.1, our sample seems to be dominated by companions with q's slightly above 0.25. Such companions, if present in the main sequence, are expected to have intermediate- and late-B types. In fact, our method is theoretically able to detect companions as late as B9 V ($\gtrsim 3\ M_\odot$). However, their BSSFs cannot be detected in the CSG spectra according to Neugent et al. (2018). Thus, the presence of such companions can explain the difference between the results of each method.

We also compare our results with those obtained for M31 and M33 by Neugent et al. (2019). Among the 149 candidates these authors observed, they found 63 RSGs with BSSFs (42 per cent). This value is well above the lower limit for the BF found in this work. However, it is much higher (one order of magnitude) than the results obtained by Neugent et al. (2019) for the MCs (or by us in this work, as explained above), and about twice the bias-corrected value obtained by Neugent et al. (2020) for the LMC. We think that this striking difference is due to three factors. First, their samples from M31 and M33 were selected using photometric criteria (in the $U-B$ and $R-I$ plane) optimized to find RSGs with OB companions. On the contrary their MC sample belongs to another unpublished

project focused on finding RSGs, without regard of any potential OB companion. That is also the case of the catalogues used in this work. Secondly, according to Neugent et al. (2018) warmer RSGs can conceal hotter B companions. As CSGs in the MCs are in average warmer than those from higher metallicity environments (as M31 and M33), they should be more effective concealing their blue companions. Thirdly, because of the larger distances to M31 and M33 the probability of having RSGs aligned by chance with OB-type stars is significantly greater. M31 and M33 candidates were observed with the fiber-fed spectrograph Hetospec, whose fibers have a projected diameter of 1.5 arcsec. At the distance of these galaxies (778 and 840 kpc, respectively), the projected diameter of a fiber covers 5.7 pc in M31 and 6.1 pc in M33. On the other hand, their MC RSGs (as well those from González-Fernández et al. 2015) were observed with AAOmega, whose fibers are slightly lager (2 arcsec of diameter). However, the LMC and the SMC are much closer (50 and 62 kpc, respectively), and thus, the projected diameter of a fiber is significantly smaller (0.5 pc in the LMC and 0.6 pc in the SMC). The probability of an alignment by chance in the MC observations is non-negligible, however, in the case of M31 and M33, given that open clusters with RSGs have typical sizes about few pcs[5] the probability is significant.

Despite the fact that the multi-epoch methodology is able to find faint OB-type stars or compact-object companions that cannot be detected through BSSFs, it has a number of important limitations (biased periods and eccentricities; and the effect of the orbital inclination) that do not affect to the detection of BSSFs. In fact, as shown in Table 2, among the 28 CSGs with BSSFs for which we have at least 2 epochs, we only detect as binaries seven of them (25 per cent). This percentage is significantly higher than the percentage of binaries found in the whole sample (25 ± 8 per cent against 4.5 ± 0.6 per cent), and than the minimum BF calculated in Section 4.1 (15 ± 3 per cent). This suggests that detecting BSSFs, although does not prove physical connection, is an acceptable proxy for binarity in the MCs. However, the level of contamination increases with distance, therefore we expect this technique to become more

---

[5]E.g. NGC 2345 has a core diameter of of 5.2 pc and a halo of 28.6 pc (Alonso-Santiago et al. 2019), and given the 4 kpc to NGC 7419 (Marco & Negueruela 2013), its diameter can be estimated in ~4.9 pc.

ineffective for galaxies at larger distances. From all these results, we conclude that both techniques should be used complimentary, as no single technique is able to provide a complete picture of the binary CSG population.

4.3 The first CSG in a triple system?

The CSG with the largest RV variation in our sample is GDR2-ID 4685850173680692480 (YSG010 in the GDN15 catalogue, where it was classified as G6.5 Ia–Iab). Despite only three epochs are available for this star within a time span of 6 yr (see Fig. 5a), it has a change in RV of 39.2 km s^{-1}. As it is relatively isolated, we discard an incorrect cross-match as an explanation for this.

YSG010 has BSSFs identified in GDN15. We checked its spectra (available in the atlas of Dorda et al. 2018), and we clearly identified the atomic lines He I 4 009 and 4 026 Å from its blue companion, shifted -52 ± 1 km s^{-1} with respect to the CSG component of YSG010. This difference adds confidence to the -39.2 km s^{-1} change observed. Thus we should expect a high RV error in *Gaia* (see Section 2.2.2), as YSG010 was observed five different epochs between 2014 and 2016. Surprisingly, it is only 1.1 km s^{-1}. This indicates that its RV did not vary much during these 2 yr, in contrast to the change of 39 km s^{-1} observed between 2012 and 2014, and the difference of 52 km s^{-1} respect its blue companion. The simplest explanation that reconcile these two facts is a very eccentric orbit, whose semimajor axis points roughly in our direction. Thus, the RV change would correspond to the periapsis or to the apoapsis. The lack of large RV variations between 2014 and 2016, together with the RV observed in 2012, indicates that the period cannot be shorter than 4 yr. Thus implies that the RV from 2009, 3 yr apart from the maximum in 2012, cannot belong to a previous period. In consequence, we should expect an orbital period significantly longer than the time span covered by our data (6.6 a; from 2009 October 2009 to 2016 May).

We studied the minimum orbital period possible for this CSG, for a range of eccentricities (eccentricity between 0.5 and 0.9) and mass-ratios (q between 0.1 and 2.0), using the methods explained in Appendix A2. From its $M_{bol} = -6.5$ and $T_{eff} = 4454$ K, we estimate a mass of ~ 12 M$_\odot$ and a radius of ~ 300 R$_\odot$. These results are shown in Fig. 5b. In addition, we took into account the limits in the P–q plane imposed by the observed change in RV : $K_1 = 39.2|/2 \geq 19.6$ km s^{-1} for the different eccentricities considered. As can be seen, the area of the P–q plane where we can have $q < 1$ and also satisfy our observational constraints ($P_{orb} > 6.6$ and at the right of the K_1 limit) is small. From a statistical point of view, it is likely that this system has $q > 1$. In fact, if we take into account an inclination $i < 90°$ (which would shift the K_1 limit to the right), or a period a few years longer, then this system would not be compatible with $q < 1$.

Although any companion whose mass was originally larger than the mass of YSG010 should have evolved (and probably died) before the former reached its current evolutionary stage, there are two scenarios able to explain a CSG in a $q > 1$ system. The first scenario is that a CSG transferred (or is transferring) a significant mass amount to the companion, but not so much that its evolution towards CSG would have been aborted. This case can happen due to pulsation instabilities which cause a mass transfer during the periapsis in mid- to high-eccentricity systems (as is the case of well-known high-mass Galactic binary systems such as VV Cep, AZ Cas, and KQ Pup). Such very specific conditions are not frequent. According to Podsiadlowski et al. (1992) only 1–2 per cent of the high-mass stars evolve through this scenario. The alternative scenario is a triple system. According to Moe & Di Stefano (2017), hierarchical triple system, with a pair in

a short-period orbit, and a third component orbiting the first pair, are frequent (10–20 per cent) among high-mass stars. Thus, the triple-system scenario is one order of magnitude more likely than the scenario with a binary and mass-transferring CSG.

If YSG010 was born in a triple system, which is the most likely case, there are two possible current configurations. In one hand, the companion is a pair of stars bonded in a short-period system, which orbits YSG010 with a very long period. On the other hand, the original OB-type companions merged (a frequent fate among short-period OB-type stars) forming a single B-type star more massive than YSG010. The lack of double He I lines suggest that the companion is a single star, but it can be also explained if the companions have an asymmetric mass, with one of them dominating the blue flux over the other. Thus, the available information do not allow to decide between these scenarios. Further observations are required to confirm the q of this system, and if it is $q > 1$, and the nature of the companion.

5 CONCLUSIONS

In this work, we study the RV variations of a large sample (~ 1000) of CSGs in the MCs. We collected data from different surveys, covering a time-span of almost 40 yr and performed a statistical homogenization to avoid systematic errors (see Section 2.3). Our objective was to find CSGs in binary systems by detecting RV variations attributable to orbital motions. For this, we studied the RV variability in our sample by two methods designed to distinguish genuine binary motions from intrinsic variations. We found a total of 45 binary CSGs, 41 of them are previously unknown. 23 are from the SMC and 22 from the LMC (see Table C1). As a result of observational biases we do not provide an estimate of the intrinsic BF but instead we calculate a lower limit of 15 ± 4 per cent for the SMC and 14 ± 5 per cent for the LMC. By combining the data from both galaxies, we obtained a minimum BF of 15 ± 3 per cent. These results are in in good agreement with previous studies of CSGs (Patrick et al. 2019, 2020) done using their RV variations, and with those obtained by Neugent et al. (2020) through single-epoch spectroscopy and photometric identification.

We compare our sample with the distributions obtained from OB-type stars (Moe & Di Stefano 2017), which indicate that our binary population is likely dominated by low to intermediate eccentricities, intermediate mass ratios, and intermediate periods. Moreover, the comparison with the theoretical study performed by Neugent et al. (2018) suggest that a significant fraction of the companions in our sample are intermediate and late B-type stars. When we compared our results with those obtained by detecting BSSFs (Neugent et al. 2018, 2019), we conclude that is vital to combine the results of both methods, BSSFs and RV variations, in order to obtain a global estimate of the BF in the future.

Finally, we report the first CSG candidate to be part of a triple system (YSG010). The data available suggest that is much more likely that it is (or was) a triple system, than a binary. However, the data available are not enough to confirm this finding, and further investigation is necessary.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

REFERENCES

Alonso-Santiago J., Negueruela I., Marco A., Tabernero H. M., González-Fernández C., Castro N., 2019, A&A, 631, A124
Beasor E. R., Davies B., Smith N., Bastian N., 2019, MNRAS, 486, 266
Belczynski K., Kalogera V., Bulik T., 2002, ApJ, 572, 407
Bessell M. S., Wood P. R., 1984, PASP, 96, 247
Britavskiy N. et al., 2019, A&A, 624, A128
Brott I. et al., 2011, A&A, 530, A115
Burki G., Mayor M., 1983, A&A, 124, 256
Chatys F. W., Bedding T. R., Murphy S. J., Kiss L. L., Dobie D., Grindlay J. E., 2019, MNRAS, 487, 4832
Chiavassa A., Plez B., Josselin E., Freytag B., 2009, A&A, 506, 1351
De Marco O., Izzard R. G., 2017, PASA, 34, e001
De Mink S. E., Langer N., Izzard R. G., Sana H., de Koter A., 2013, ApJ, 764, 166
Doherty C. L., Gil-Pons P., Siess L., Lattanzio J. C., Lau H. H. B., 2015, MNRAS, 446, 2599
Dorda R., Negueruela I., González-Fernández C., Tabernero H. M., 2016, A&A, 592, A16
Dorda R., Negueruela I., González-Fernández C., Marco A., 2018, A&A, 618, A137
Dunstall P. R. et al., 2015, A&A, 580, A93
Eggleton P. P., 1983, ApJ, 268, 368
Ekström S., Georgy C., Meynet G., Groh J., Granada A., 2013, in Kervella P., Le Bertre T., Perrin G., eds, EAS Publications Series, Vol. 60. p. 31
Eldridge J. J., Izzard R. G., Tout C. A., 2008, MNRAS, 384, 1109
Gaia Collaboration, 2018, A&A, 616, A12
González-Fernández C., Dorda R., Negueruela I., Marco A., 2015, A&A, 578, A3
Graczyk D. et al., 2014, ApJ, 780, 59
Gray D. F., 2008, AJ, 135, 1450
Hirai R., Sato T., Podsiadlowski P., Vigna-Gomez A., Mandel I., 2020, MNRAS, 499, 1154
Jorissen A., Van Eck S., Kravchenko K., 2016, Velocity Fields in Stellar Atmospheres Probed by Tomography, Astronomy at High Angular Resolution, Astrophysics and Space Science Library, Vol. 439, Springer International Publishing, Switzerland. p. 137
Josselin E., Plez B., 2007, A&A, 469, 671
Katz D. et al., 2019, A&A, 622, A205
Kiss L. L., Szabó G. M., Bedding T. R., 2006, MNRAS, 372, 1721
Klencki J., Nelemans G., Istrate A. G., Chruslinska M., 2021, A&A, 645, 54
Kobulnicky H. A. et al., 2014, ApJS, 213, 34
Kravchenko K., Chiavassa A., Van Eck S., Jorissen A., Merle T., Freytag B., Plez B., 2019, A&A, 632, A28
Kruckow M. U., Tauris T. M., Langer N., Szécsi D., Marchant P., Podsiadlowski P., 2016, A&A, 596, A58
Kunder A. et al., 2017, AJ, 153, 75
Langer N. et al., 2020, A&A, 638, A39

Marchant P., Langer N., Podsiadlowski P., Tauris T. M., de Mink S., Mandel I., Moriya T. J., 2017, A&A, 604, A55
Marco A., Negueruela I., 2013, A&A, 552, A92
Massey P., Olsen K. A. G., 2003, AJ, 126, 2867
Maurice E. et al., 1987, A&AS, 67, 423
Moe M., Di Stefano R., 2017, ApJS, 230, 15
Neugent K. F., Massey P., Skiff B., Drout M. R., Meynet G., Olsen K. A. G., 2010, ApJ, 719, 1784
Neugent K. F., Massey P., Skiff B., Meynet G., 2012, ApJ, 749, 177
Neugent K. F., Levesque E. M., Massey P., 2018, AJ, 156, 225
Neugent K. F., Levesque E. M., Massey P., Morrell N. I., 2019, ApJ, 875, 124
Neugent K. F., Levesque E. M., Massey P., Morrell N. I., Drout M. R., 2020, ApJ, 900, 118
Patrick L. R. et al., 2019, A&A, 624, A129
Patrick L. R. et al., 2020, A&A, 635, 49
Podsiadlowski P., Joss P. C., Hsu J. J. L., 1992, ApJ, 391, 246
Prevot L. et al., 1985, A&AS, 62, 23
Robitaille T. P. et al., 2008, AJ, 136, 2413
Sana H. et al., 2012, Science, 337, 444
Sana H. et al., 2013, A&A, 550, A107
Sana H. et al., 2014, ApJS, 215, 15
Schwarzschild M., 1975, ApJ, 195, 137
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Soszyñski I. et al., 2009, Acta Astron., 59, 239
Soszyñski I. et al., 2011, Acta Astron., 61, 217
Stothers R. B., 2010, ApJ, 725, 1170
Tabernero H. M., Dorda R., Negueruela I., González-Fernández C., 2018, MNRAS, 476, 3106
Walker A. R., 2012, Ap&SS, 341, 43
Wang C., Langer N., Schootemeijer A., Castro N., Adscheid S., Marchant P., Hastings B., 2020, ApJ, 888, L12

APPENDIX A: VARIATION PERIODS AMONG CSGS

A1 Comparison between secondary periods and minimum orbital periods

We have calculated the minimum periods (P_{min}) for which CSGs do not interact with their companions. The interaction would happen if the radii of the CSG (R_{CSG}) grow large enough to overflow its Roche Lobe at the periapsis. By using the ratio r_L between the separation of both stars at periapsis (r_p) and the size of the Roche Lobe (r_{RL}) proposed by Eggleton (1983) it can be demonstrated that to avoid the Roche lobe overflow this condition must be satisfied:

$$R_{CSG} \geq r_L(q)\,a\,(1 - e),$$

where a is sum of the orbital semimajor axes, and e the eccentricity. From here we can calculate the minimum period (P_{min}) possible for a given CSG mass (M_{CSG}) and mass ratio q

$$P_{min} = \left(\frac{R_{CSG}}{r_L(q)\,(1 - e)} \right)^{3/2} 2\,\pi \left(\frac{q}{G\,M_{CSG}\,(1 + q)} \right)^{1/2}.$$

We used the stellar parameters predicted by Brott et al. (2011). In particular, we used as R_{CSG} the maximum size reached in any moment by the CSG of each initial mass (M_{CSG}) considered. The calculations were done for a number of q, but always with zero eccentricity, as higher eccentricity would imply necessarily longer P_{orb}. The results are plotted in Fig. A1, for the SMC and the LMC, respectively, and they are compared with the longest P_2 found among CSGs (5000 d Chatys et al. 2019).
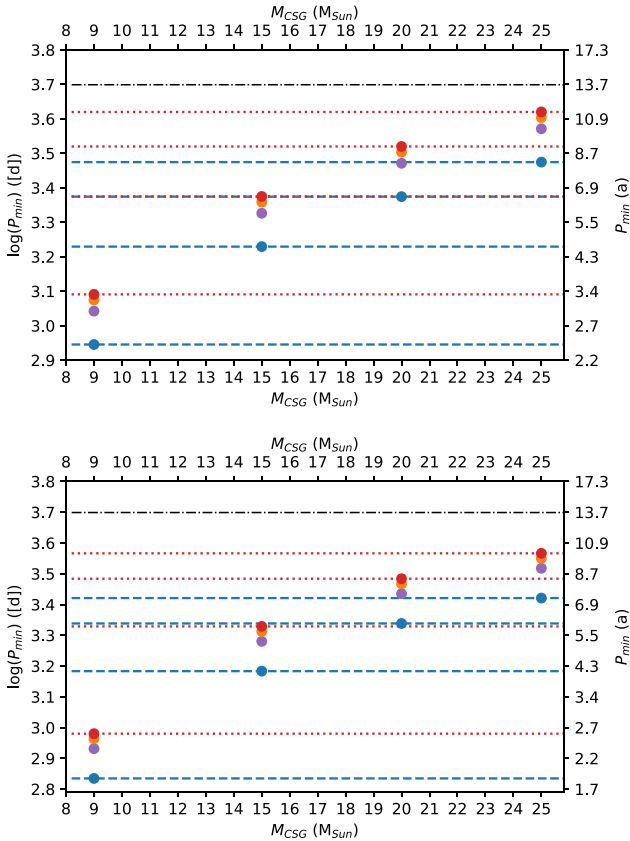
Figure A1. Minimum period depending on the CSG mass. The period is expressed as the log of days (left axis), and also in years (right axis). The colour of the circles indicate different mass ratios: blue for $q = 0.1$, magenta for $q = 0.4$, orange for $q = 0.7$, and red for $q = 1.0$. The black lines indicate the longest P_2 considered (5000 d) for comparison. The coloured horizontal lines are a visual aid to better see which is the range of P_{min} for a given M_{CSG} and a different q. The upper panel shows the results for CSGs from the LMC and the lower plane for CSGs from the LMC.

A2 Mass and radius estimation

In this paper, we estimated the mass and radius for a number of CSGs. We calculated first their luminosity using their 2MASS photometry (Skrutskie et al. 2006), transformed to the AAO system, to calculate the bolometric correction according to Bessell & Wood (1984). By using infrared photometry, we avoid any significant contamination from any hypothetical blue companions and we reduce the effects of interstellar reddening to a negligible level. In addition, as the photometric variations of CSGs decrease in amplitude with the wavelength (Robitaille et al. 2008), the infrared provides a more reliable estimation. To calculate the absolute magnitudes, we used a distance modulus of 18.5 mag for the LMC (Walker 2012) and of 19.0 mag for the SMC (Graczyk et al. 2014). Then, we estimated the mass by comparing the position of the CSG in the Hertzsprung–Russel diagram with the evolutionary tracks from Brott et al. (2011). Finally, we calculated the radius using the luminosity and the T_{eff} published in Tabernero et al. (2018) when available. As they provide one T_{eff} per epoch, we use the average (weighted by their signal-to-noise ratio) as reference value. When a CSGs is not available in that work, we used the relation between $(J - K_s)$ and T_{eff}. For the LMC, this relation was calculated by Britavskiy et al. (2019). For the SMC,
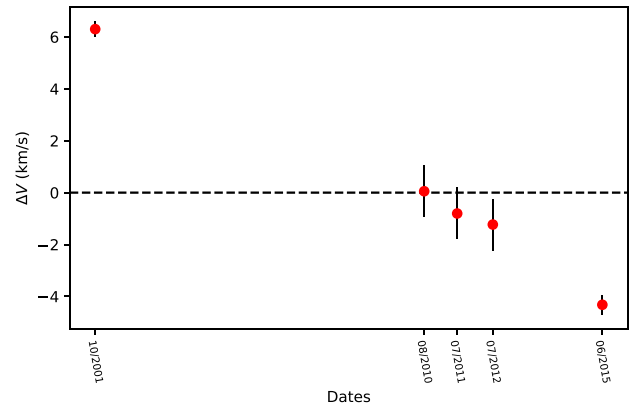

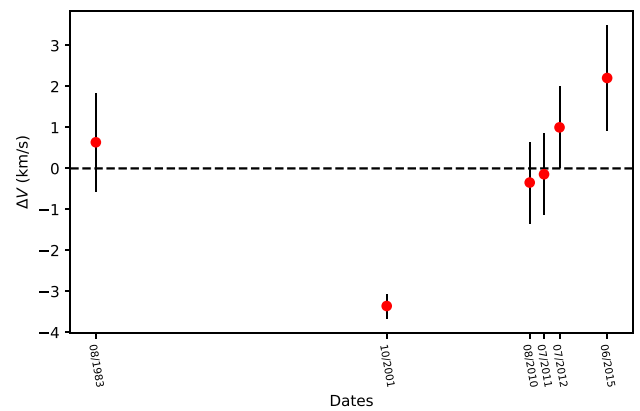**Figure B1.** RVC of GDR2-ID 4685947794049430656.

Figure B2. RVC of GDR2-ID 4690517158168014848.

we calculated this relation, and we obtained:

$$T_{eff} = -1571(J - K_s)_0 + 5660\,(K).$$

APPENDIX B: RADIAL VELOCITY CURVES WITH ORBITAL-TYPE BEHAVIOURS

In this Appendix, we show the RVCs that cannot be explained by intrinsic variations and thus we have considered them caused by orbital motions (see Section 3.2). We discuss each CSG in detail, and we show their corresponding RVCs.

Take into account that many of the CSGs displayed here present a high *Gaia* RV error. As explained in Section 2.2.2, this is due these errors are related to the dispersion of the epochs measured by *Gaia* between 2014 and 2016. Thus, those with large values should not be considered as a unreliable measurement, but a sign that during these 2 yr the RV of that CSG varied significantly.

GDR2-ID 4685947794049430656 (Fig. B1): All the five data of this CSG have a Pearson's correlation coefficient $r = -0.995\,97$. The probability of finding a random alignment of data points in the synthetic RVCs able to explain this coefficient is 0.01 per cent, which is lower than our threshold level of 0.27 per cent. In addition, its trend covers a variation time-span of 14 a, which suggest a P_{orb} of at least ~3 decades, although it can be much longer. Thus, we think that intrinsic variations cannot explain this curve.

GDR2-ID 4690517158168014848 (Fig. B2): The five data between 2001 and 2015 have a Pearson's correlation coefficient $r = -0.983\,87$. The probability of finding a random alignment of

data points in the synthetic RVCs (with a period randomly selected from the range between 3928 and 4802 d) able to explain this coefficient is 0.09 per cent, which is lower than our threshold level of 0.27 per cent. If we assume the data points observed in 2001 and in 2015 as the minimum and maximum, respectively, of the RVC,

the P_{orb} would be almost 30 a. Thus, we think that intrinsic variations cannot explain this curve.

APPENDIX C: BINARIES FOUND

Table C1. List of binaries found, identified by their *Gaia* DR2 ID, together with their fundamental parameters for this work.

GDR2-ID	Catalogue name	RAJ2000	Dec.J2000	$|\Delta V_{max}|$ (km s^{-1})	$\frac{|v_i - v_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}}$	#Epochs	Galaxy	Selected by
4690519185392371456	[M2002] 50360	01:01:03.58	−72:02:58.50	23.2	22.3	6	SMC	$|\Delta V_{max}| > 11$
4690518910514527616	[M2002] 50348	01:01:03.26	−72:04:39.40	20.4	21.5	6	SMC	$|\Delta V_{max}| > 11$
4690517158168014848	[M2002] 48122	01:00:09.42	−72:08:44.50	5.6	4.2	6	SMC	RVC
4690512858880950656	[M2002] 57386	01:03:47.35	−72:01:16.00	31.9	30.5	4	SMC	$|\Delta V_{max}| > 11$
4690509152347949184	[M2002] 56732	01:03:34.30	−72:06:05.80	12.1	11.54	5	SMC	$|\Delta V_{max}| > 11$
4690507193841115008	[M2002] 51000	01:01:19.92	−72:05:13.10	14.8	10.5	6	SMC	$|\Delta V_{max}| > 11$
4690505200976492160	[M2002] 49478	01:00:41.56	−72:10:37.00	18.1	7.8	7	SMC	$|\Delta V_{max}| > 11$
4689267700627387776	[GDN15] SMC206	00:51:50.50	−71:59:24.23	19.1	12.6	3	SMC	$|\Delta V_{max}| > 11$
4688992032495995520	PMMR 061	00:53:44.50	−72:33:19.31	19.6	9.7	2	SMC	$|\Delta V_{max}| > 11^a$
4688979555567809792	[M2002] 19743	00:51:23.28	−72:38:43.80	11.4	5.4	7	SMC	$|\Delta V_{max}| > 11$
4688965850370232704	[M2002] 18592	00:51:03.90	−72:43:17.40	23.8	16.8	6	SMC	$|\Delta V_{max}| > 11$
4688861014472093568	SkKM13	00:45:04.57	−73:05:27.51	15.5	6.6	6	SMC	$|\Delta V_{max}| > 11$
4687501747539145216	[GDN15] SMC375	01:05:27.44	−72:17:04.35	15.7	10.0	3	SMC	$|\Delta V_{max}| > 11$
4687489584188959872	[M2002] 64663	01:06:47.62	−72:16:11.90	12.7	8.4	7	SMC	$|\Delta V_{max}| > 11$
4687481024294949376	[M2002] 55355	01:03:06.43	−72:28:35.10	22.0	23.2	6	SMC	$|\Delta V_{max}| > 11^b$
4687412064305984768	[M2002] 63188	01:06:03.21	−72:52:16.00	19.2	18.4	4	SMC	$|\Delta V_{max}| > 11$
4686455764056822016	[M2002] 83593	01:30:33.92	−73:18:41.90	12.2	7.4	3	SMC	$|\Delta V_{max}| > 11$
4686417315506422016	PMMR 198	01:18:17.70	−73:09:27.48	13.9	5.6	3	SMC	$|\Delta V_{max}| > 11$
4685989781659094144	[GDN15] YSG070	01:00:09.60	−72:33:59.30	16.2	7.1	3	SMC	$|\Delta V_{max}| > 11$
4685947794049430656	[M2002] 11101	00:48:31.92	−73:07:44.40	10.6	22.4	5	SMC	RVC
4685926761539086848	[M2002] 11939	00:48:51.82	−73:22:39.83	12.1	6.1	6	SMC	$|\Delta V_{max}| > 11$
4685850173680692480	[GDN15] YSG010	00:47:08.69	−73:14:11.90	39.2	25.9	3	SMC	$|\Delta V_{max}| > 11^b$
4685836605893771520	[GDN15] SMC099	00:46:41.68	−73:22:54.20	16.0	6.9	2	SMC	$|\Delta V_{max}| > 11^b$
4662153228519937664	RM1-080	04:57:26.36	−66:23:25.78	13.06	4.7	3	LMC	$|\Delta V_{max}| > 11$
4661877491594670208	RM1-151	05:04:14.12	−67:16:14.40	13.84	4.8	3	LMC	$|\Delta V_{max}| > 11$
4661772144640288384	RM1-024	04:52:53.64	−66:55:52.14	11.93	9.6	3	LMC	$|\Delta V_{max}| > 11$
4661446757944600448		05:04:53.93	−68:01:58.64	24.14	4.5	2	LMC	$|\Delta V_{max}| > 11$
	2MASS J05045392−6801585							
4661306016119437056		05:03:02.35	−68:47:20.23	29.94	6.0	2	LMC	$|\Delta V_{max}| > 11^a$
	2MASS J05030232−6847203							
4660179639518872576	[M2002] 147928	05:30:33.47	−67:17:15.07	23.47	14.4	3	LMC	$|\Delta V_{max}| > 11$
4658466355538857728	RM2-093	05:31:24.27	−68:41:33.64	14.25	6.2	4	LMC	$|\Delta V_{max}| > 11^b$
4658433443240057216	[M2002] 145716	05:29:54.75	−69:04:15.69	12.52	7.1	7	LMC	$|\Delta V_{max}| > 11$
4658429143945343872	[M2002] 150577	05:31:18.44	−69:09:28.16	25.29	4.6	6	LMC	$|\Delta V_{max}| > 11$
4658287100783922304		05:19:15.79	−68:56:03.89	18.47	5.1	2	LMC	$|\Delta V_{max}| > 11$
	2MASS J05191579−6856039							
4658097198899548928	[M2002] 139413	05:27:47.51	−69:13:20.59	11.73	8.3	6	LMC	$|\Delta V_{max}| > 11^a$
4658057685185383424	[M2002] 144217	05:29:27.58	−69:08:50.33	11.69	4.5	5	LMC	$|\Delta V_{max}| > 11$
4658052943542064896	[M2002] 148381	05:30:41.43	−69:15:33.84	14.06	13.5	4	LMC	$|\Delta V_{max}| > 11^b$
4657674642790217600	[M2002] 174742	05:40:25.32	−69:15:30.20	11.46	5.1	5	LMC	$|\Delta V_{max}| > 11$
4657598364163823872	SP77 54-38	05:41:10.66	−69:38:04.04	11.94	7.3	4	LMC	$|\Delta V_{max}| > 11$
4657099387712019072	RM1-603	05:33:04.42	−70:48:30.13	16.14	14.0	3	LMC	$|\Delta V_{max}| > 11$
4655456756015488000	[M2002] 21369	04:54:36.85	−69:20:22.18	12.34	7.1	5	LMC	$|\Delta V_{max}| > 11$
4655369100006728320	RM1-015	04:51:20.58	−69:29:13.97	16.82	4.7	3	LMC	$|\Delta V_{max}| > 11$
4655321000684432384		04:51:09.48	−69:56:14.43	14.4	4.7	2	LMC	$|\Delta V_{max}| > 11$
	2MASS J04510944−6956146							
4651906394290251520	RM2-069	05:23:35.76	−70:44:55.34	15.38	3.1	3	LMC	$|\Delta V_{max}| > 11$
4651885194331239424		05:21:01.36	−70:58:45.90	16.61	4.8	2	LMC	$|\Delta V_{max}| > 11$
	2MASS J05210134−7058459							
4657696972286770432	[M2002] 164506	05:36:06.37	−68:56:40.70	11.30	8.4	5	LMC	$|\Delta V_{max}| > 11$

[a] This CSG has been reported to have BSSFs in Neugent et al. (2019).
[b] This CSG has been reported to have BSSFs in González-Fernández et al. (2015).

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