Stellar collisions in globular clusters: the origin of multiple stellar populations

Valery Kravtsov¹, Sami Dib², Francisco A. Calderón³, José Antonio Belinchón⁴

¹Sternberg Astronomical Institute, Lomonosov Moscow State University, University Avenue 13, 119899 Moscow, Russia
²Max Planck Institute for Astronomy, Königstuhl 17, D-69117, Heidelberg, Germany
³Departamento de Física, Universidad Católica del Norte, Av. Angamos 0610, Antofagasta, Chile
⁴Departamento de Matemáticas, Universidad de Atacama, Av. Copayapu 485, Copiapó, Chile

18 March 2022

ABSTRACT

Two generations of stars, G1 and G2, typically populate Galactic globular clusters (GCs). The origin of G2 stars is unclear. We uncover two empirical dependencies between GC characteristics, which can be explained by the formation of G2 Main-Sequence (MS) stars due to collision/merging of their primordial counterparts (G1). A similar genesis of both G2 stars and peculiar objects like LMXBs and millisecond pulsars is also implied. Indeed, we find a significant (at a confidence level > 99.9%) anti-correlation between the fraction of G1 red giants \( (N_{G1}/N_{tot}) \) and stellar encounter rates among 51 GCs. Moreover, a Milky Way-like initial mass function (IMF) requires at least \( \sim 50\% \) of MS stars located in the mass range \([0.1 - 0.5]M_\odot\). Unlike cluster mass loss, stellar collisions/merging retain these G1 stars by converting them into more massive G2 ones, with mainly \( M_{MS} > 0.5M_\odot \). This process coupled with a decreasing relative mass loss with increasing GC masses implies a smaller \( (N_{G1}/N_{tot}) \) in more massive GCs with a shallower present day MF. From data for 35 GCs, we find that such an anti-correlation is significant at 98.3% confidence level (Spearman’s correlation) for the 12 most massive GCs \( (M_{GC} > 10^{5.3}M_\odot) \) and it is at a confidence level of 89% for the 12 least massive GCs \( (M_{GC} < 10^{5.1}M_\odot) \). Other fractions of G1 and G2 stars observed at the bottom of the MS as compared with the red giant branch in a few GCs are consistent with the scenario proposed.

Key words: globular clusters: general - (stars:) binaries (including multiple): close

1 INTRODUCTION

Globular clusters (GCs) are amongst the most ancient stellar aggregates formed both in the Milky Way and in other galaxies. They form very early during the galaxies’ assembly history and despite a continuous evaporation of stars, have remained bound in different galactic environments. Thanks to their unique characteristics, GCs are of special interest for studying a variety of key problems in astrophysics.

One well known problem arises from the observational fact that stellar populations (SPs) in GCs are not quite simple. A lot of evidence about a more complex nature of GC SPs have been obtained in numerous works up to now. This complexity is widely referred to as the multiplicity of the cluster SPs. Here, we are focusing on the most generic case of the multiplicity observed in virtually all but several monometallic Galactic GCs. Namely, two populations or generations of stars, G1 (primordial) and G2 (secondary), with no (or no significant) difference in mean iron abundance are typically distinguished both spectroscopically (e.g., Carretta et al. 2009) and photometrically (Milone et al. 2017) in GCs. Hereafter, we refer to this multiplicity as a complexity of SPs. Moreover, the complexity of SPs has been studied in a wider context, i.e not only in Galactic GCs but also in young and intermediate-age massive star clusters in the Large and Small Magellanic Clouds (LMC and SMC). These clusters normally exhibit a complexity of their SPs at the age around 2 Gyr in the form of an extended MS turnoff (Mackey & Broby Nielsen 2007; Mackey et al. 2008; Milone et al. 2009). The manifestations typical for multiple SPs in GCs are observed in LMC/SMC massive star clusters older than 2 Gyr (e.g., Hollyhead et al. 2017; Niederhofer et al. 2017; Hollyhead et al. 2018; Milone et al. 2020; Li et al. 2021). This has challenged our understanding of the origin of the second generation in both massive star clusters in general and GCs in particular.

A number of scenarios have been proposed already a decade ago and even earlier to explain both the origin of and key differences between the SPs in GCs (e.g., Decressin et al. 2007; Ventura & D’Antona 2009; Decressin et al. 2010; Valcarce & Catelan 2011; Krause et al. 2013). For the past decade, many aspects of this important subject were further

* E-mail, VK: vkravtsov1958@gmail.com

© 2016 The Authors
studied in more detail. They have been summarized and reviewed, in particular, by Bastian & Lardo (2018). Interested readers are referred to this paper for key particulars concerning the subject, including the scenarios proposed so far and problems met within the frame of each one. It appears that, although a much deeper insight into the matter has been gained, the origin of G2 stars remains an unclear and even more controversial issue than before. One of the most serious problems associate with each of the proposed scenarios is the so-called ”mass-budget problem”. Here, we report supporting evidence of essential contribution of stellar collision/merging to the formation of G2 stars, a mechanism that is probably able not only to overcome the ”mass-budget problem” but to simultaneously explain other features of massive clusters such as the origin of their initial mass function (IMF) (e.g., Dib et al. 2007).

We have recently paid attention to modified MS stars which should have formed from their primordial counterparts due to (presumably mainly binary mediated) stellar collisions in the central regions of GCs (Kravtsov 2019, 2020; Kravtsov & Calderón 2021). In particular, we argued for a potential role of these stars and their descendants on both the sub-giant branch (SGB) and red giant branch (RGB) in contributing toward creating (1) a range of masses and varying surface abundances of particular elements among stars naturally believed to be congenerical (i.e., which evolve(d) from primordial MS stars of the same mass with its negligible dispersion) and (2) radial effects among MS turnoff and SGB/RGB stars in GCs. Indeed, a small portion of such modified stars are well known as collisional blue stragglers (CBSs) actually observed in GCs (Ferraro et al. 2009; Dalessandro et al. 2013). The existence of CBSs is direct evidence of the collision process at work. However, there are reasons to believe that the amount of the presently observed CBSs is only ”the tip of the iceberg” of the total number of modified MS stars formed and accumulated in GCs during their lifetime, in particular within the mass range [0.5 – 0.9]M⊙. By their characteristics, these collision-generated MS stars are different from primordial MS stars of the same mass and could hardly therefore belong to the population of G1 stars in GCs. For instance, according to the models of collision products appropriate to an old open cluster like M67 (Glebbeek et al. 2008), their descendants on the RGB are expected to have different surface carbon and nitrogen abundances in comparison with normal red giants in the sense that [C/Fe] is depleted and [N/Fe] is enhanced (see also Wei et al. 2020 who calculated surface chemical anomalies of low-mass MS stars due to mass transfer in a binary system). The so-called luminous red novae (see Section 3.4) are known to be reliable evidence of merging binary systems in over a large mass range closely approaching the most typical mass of binary systems in GCs and are promising to shed more light on the role and outcome of the merging process. Moreover, collisionally formed MS stars normally complete their MS evolution and achieve (and go on) the RGB simultaneously with primordial stars of lower mass. This can be not only suspected based on basic laws of stellar astrophysics but also more reliably deduced from relevant models of MS stars formed due to collision of two primordial MS stars with different mass ratios in GCs, at different cluster age (Sills et al. 2009). However, one of the main problems is that the fraction of collisionally formed G2 stars in GCs is poorly known. It is estimated with a large uncertainty.

In this paper, we further study the role and effect of the collision process, binary-mediated and permanently acting in GCs, in the transformation of their stellar populations, in general. We explore the dependencies, specifically the fraction of G1 stars versus both the stellar encounter rates and mass function (MF) slope in GCs, giving insight into a non-negligible mechanism for G2 stars to form.

In the next section (Section 2), we describe the data used for our analysis. In Section 3, we present: (1) the results dealing with both new dependencies obtained from the data used; (2) a model testing the availability of G1 MS stars in initial GCs to form populations of G2 stars observed in the present-day GCs on the RGB; (3) a verifiable effect implied by our scenario and observational results on merging stars producing luminous red novae. A summary of our results and concluding remarks are in Section 4.

2 DATA

For the present study devoted to the contribution of stellar collisions to the formation of the populations of G2 stars in GCs, we used three key sets of recently published and publicly available data on Galactic GCs and their SPs. First, we use the atlas on multiple SPs in Galactic GCs, obtained by Milone et al. (2017). Using uniform multiband HST photometry (Sarajedini et al. 2007; Piotto et al. 2015) in the central parts of a large sample of 57 GCs, the authors isolated RGB stars, separated them into subpopulations belonging to different generations, and deduced a number of parameters characterizing each generation. Also, Milone et al. (2017) calculated the total number of RGB stars, Ntot, falling in the observed area of each GC and quantified the relative number of G1 red giants (NG1/Ntot) as the ratio between the number of G1 (NG1) and the total number of red giants in a GC. The number of GCs where Milone et al. (2017) have been able to estimate the relative number of G1 red giants is 54.

Second, we also make use of the recently published data of Ebrahimi et al. (2020) on the clusters’ stellar MFs, masses, and a number of other characteristics listed for a sample of 32 GCs. However, NGC 6584 was excluded from our analysis, since the data on this GC are incomplete. Instead, we added data, taken from Sollima & Baumgardt (2017) on four other GCs: NGC4833, NGC6205, NGC6397, and NGC6656. Therefore, we finally obtained a sample of 35 GCs with available data on both the fraction of G1 and MFs and cluster masses. The original papers list logarithms of both the luminous and dynamical masses of GCs, logMlum and logMdyn, respectively. We notice that the GC masses used for our present study are expressed in the form MGC = 10^4 M⊙, where the exponent A is the mean between A_lum=logMlum and A_dyn=logMdyn.

Third, the data on stellar encounter rates (and their errors) in GCs, Γ, are taken from Bahramian et al. (2013) where the data are normalized assuming Γ = 1000 for the GC NGC104. Since the values of logΓ were used instead of Γ, we calculated these quantities and expressed the corresponding errors in the logarithmic scale. Overall, we obtained a sample of 51 GCs with available data on both Γ and (NG1/Ntot). Finally, we refer the interested readers to the aforementioned original publications, for additional detail on both the data them-
selves and related information about the observations data reduction procedures.

3 RESULTS

We examine the potential key role of binary-mediated stellar collisions in the formation of multiple SPs in GCs through formation of modified MS stars (a small portion of which presently observed as collisional BSs) from primordial stars, at different epochs. Many observational studies show that the typical effects of multiplicity of SPs in GCs, such as anti-correlations between abundances of some chemical elements, are not manifested by SPs even in old Galactic clusters nor in the field, except perhaps for its dense regions in the Galactic bulge (Schiavon et al. 2017). This fact implies that one of the most probable factors responsible for this dissimilarity is binary-mediated stellar collisions which do normally occur in the densest parts of GCs but are less probable to occur in environments where the stellar density is low. In relation to this, it is relevant to mention a new type of somewhat exotic objects (at least due to their special location in the color-magnitude diagram of globular and open clusters), so-called sub-subgiant stars, recently introduced and studied by Geller et al. (2017a,b) ... Leiner et al. (2017). These authors argue that a portion of these stars are the product of merged MS stars, which are now "observed while settling back down onto the normal main sequence". Given the aforementioned arguments, we verify two expected outcomes of the process of collision-induced merging of stars in relationship with the multiplicity of SPs in GCs.

3.1 The fraction of G1 stars versus the stellar encounter rates

Extremely dense central regions of GCs are recognized as unique sites where exotic objects, like collisional BSs, millisecond pulsars, and low mass X-ray binaries (LMXBs), among others, are preferably observed. It is primarily due to the conditions favorable for the formation of close binaries, subsequent evolution of which leads to their transformation into such exotic objects responsible for the respective astrophysical phenomena. Binary stars play a crucial role in relation to all processes that occur in the central regions of GCs, in particular, and in the dynamical evolution of GCs, in general. This important topic and many closely related aspects like binary mediated collisions and their outcomes were the subject of numerous astrophysical investigations. For many details and respective key references, we refer here, for example, to the review by Meylan & Heggie (1997). As has been already noted, we assumed that a significant fraction of G2 stars, like their prototypes in the form of collisional BSs, are of binary mediated collisional origin.

It is now established that there is statistically significant correlation between the populations of both LMXBs and millisecond pulsars and the stellar encounter rate in Galactic GCs (Pooley et al. 2003; Bahramian et al. 2013). Given these facts, we explored the dependence of \( \frac{N_{G1}}{N_{tot}} \) on \( \log \Gamma \), the logarithm of the stellar encounter rate in GCs, by relying on the respective data available for a sample of 51 Galactic GCs. This is displayed in Figure 1 which indicates an anti-correlation between the two quantities. Spearman’s correlation coefficient is \( \rho = -0.729 \) and this implies that this anti-correlation is statistically significant at a high confidence level (>99.9%). We note that the GC that deviates most from the sequence shown in Figure 1 and has the smallest \( \frac{N_{G1}}{N_{tot}} \) is Omega Cen. It has not been excluded from the statistical test made. However, this GC is obviously a special case among Galactic GCs, since it is well known to show a fairly large star-to-star variation in the [Fe/H] ratio and the presence of genuine multiple stellar populations appropriate for galaxies. This is opposed to the complexity of stellar populations in the majority of Galactic GCs, which we deal with in the present study. So, the astrophysical concept of the parameter \( \frac{N_{G1}}{N_{tot}} \) in Omega Cen is not, strictly speaking, equivalent to that in other GCs. There is also another important point concerning the \( \frac{N_{G1}}{N_{tot}} \) ratio that should be commented on. It has been measured in the central parts of GCs. Hence, it characterizes the fraction of G1 stars (red giants) exactly in the central part of a GC but not in the GC as a whole. The measured values of this parameter should not be considered as absolute ones, i.e. appropriate for GCs as entire objects. Indeed, given that G2 stars are often more concentrated toward the GC centers, the expected absolute values of the fraction of G1 stars in GCs are expected to be, on average, systematically higher than the presently available data on the \( \frac{N_{G1}}{N_{tot}} \) ratio used in the present study.

Finally, it should be noted that the obtained dependence is not quite trivial, for different reasons. It is seen that the dispersion of the points is obviously larger than would be expected being due to the typical formal errors of both quantities and the deviations of two GCs (apart from Omega Cen) are notably large. We will analyze this dependence in more detail in a forthcoming paper.

3.2 The fraction of G1 stars vs the slope of the present-day mass function

There is ongoing, intense debates on whether the IMF of stars in individual stellar clusters is universal and similar to the mass function of stars in the nearby Galactic field (e.g., Bastian et al. 2010; Dib 2014; Dib et al. 2017). If the IMF of...
GCs is assumed to be canonical, similar to the MF of stars in the Galactic field irrespective of GC characteristics, then the bulk of MS stars (\(\sim 85\%\)) should have masses \(M_{MS} < 0.5M_\odot\) or more specifically, \(\sim 50\%\) with \([0.1 - 0.5]M_\odot\) (Kroupa 2001, 2002). The fraction of these low-mass stars in GCs decreases with GC age. One of the important processes responsible for that is dynamical interactions. At first glance, stellar collisions leading to merging of stars produce a similar effect, since they primarily reduce this reservoir of the most numerous stars. However, in contrast to cluster evaporation, stellar collisions transform these low-mass G1 stars and thereby retain them in GCs in other form, by converting them into G2 stars of a higher mass, primarily within the mass range \([0.5 - 0.9]M_\odot\). Therefore, the collisional formation of G2 stars implies that the shallower the actual stellar MF of GCs at a given cluster mass, the smaller the fraction of G1 stars. Recent simulations of dynamical cluster disruption (Reina-Campos et al. 2018) predict decreasing relative mass loss with increasing GC masses. Given this, one can expect that the effect of stellar collisions on the transformation of the IMF should be more competitive among the most massive GCs.

In the collision-based scenario for the formation of G2 stars, an increasingly shallower present-day mass function would imply a more efficient collision process under the assumption of the same IMF for G1 stars. This would result in an increasingly smaller fraction of G1 stars and larger fraction of G2 stars. We will discuss this from a theoretical point of view in Section 3.3. This process can also be impacted by variations in the IMF of G1 stars, particularly variations that could affect the initial fraction of stars present in the mass range \([0.1 - 0.5]M_\odot\). We will explore the effect of IMF variations in a forthcoming paper.

We derived the dependence between the present-day stellar MF slope and the fraction of G1 stars in GCs of a large range of mass. This is shown in the upper panel of Figure 2. It appears quite dispersed, with correlation between the two quantities at low confidence level of 70% according to Spearman’s rank correlation test. We have drawn a linear fit to the data, obtained the deviations of individual points from the line fitted to the data, and plotted them as a function of the GC mass in the middle panel of Figure 2. From this panel, one can see that the deviations depend on the mass of GCs. For this reason, the analyzed sample of GCs has been divided into subsamples by meeting the following requirements: as large as possible separation by mass between higher- and lower-mass GCs along with having a reasonable subsample size of objects in each category. We decided upon a minimum subsample size of 12 GCs in each category. High-mass GCs have masses \(M_{GC} > 10^{5.3}M_\odot\) and lower-mass GCs have masses \(M_{GC} < 10^{5.1}M_\odot\). The remaining 11 GCs fall in the intermediate mass range. GCs belonging to the three subsamples are shown by different colors in both the middle and lower panels of Figure 2. Spearman’s rank correlation test shows that a correlation in the expected sense is statistically significant at 98.3% confidence level for the subsample of the most massive GCs. For the subsample of the least massive GCs, however, there is more scatter, and the correlation is less significant with a confidence level of 89% (and at least formally it is statistically insignificant). See test results in Table 1. From the appearance of the dependence for GCs of different mass in the lower panel of Figure 2, one cannot exclude potentially interesting details. For example, the range of the MF slope of the least massive GCs is obviously larger than among the most massive ones. Perhaps this is a signature of a larger mass loss from the least massive GCs and its respective stronger impact on the observed variation of the MF slope.

We recall that understanding the origin of the complexity of stellar populations in GCs is mainly challenged by two key problems to be explained together, namely the "mass-budget problem" and the anti-correlations of the surface abundances of a number of light elements. Below, we deal with both issues in more detail.

### 3.3 Mass budget: a toy model

Here, we present a model that can help us interpret our finding for the correlation between the ratio \((N_G1/N_{tot})\) and the slope of the present-day MF \(\alpha_{pd}\) in the framework of our proposed collision-based scenario. Strictly speaking, understanding the exact effect of collisions is obviously a very complicated task. It requires a reliable knowledge of a number of key parameters such as the collision rates over the entire stellar mass range, the coalescence efficiency as a function of stellar mass (e.g., Dib et al. 2007) and the mass loss as a function of cluster mass, etc. However, in order to estimate whether, and under which conditions, the capacity of the main reservoir of primordial lower-mass MS stars would be sufficient to produce the amount of collision products comparable with the observed fraction of G1 stars in GCs, we apply a much simpler approach. By means of our model which assumes a Kroupa-like IMF for the G1 stars, we compare the fractions of the MS stars that collide to form the respective G2 stars of a higher mass. Moreover, a model MS analogue of the ratio \((N_G1/N_{tot})\) (empirically derived for the RGB in Galactic GCs by Milone et al. (2017)) and the slope of the present day MF \(\alpha_{pd}\) are deduced to obtain a model MS analog of the empirical dependence obtained for the RGB and presented in Figure 2.

For simplicity, (1) we are focusing on the mass range \([0.1 - 0.9]M_\odot\) in which the bulk of MS stars reside in both young and present-day GCs, (2) we divide it by two ranges, assuming that stars that merge are those exclusively found in the mass range \([0.1 - 0.5]M_\odot\), and (3) we consider the case of insignificant mass loss, i.e. more compatible with the most massive GCs. The former mass range is conditionally referred to as "source" mass range, since it is very probably the main reservoir of stars. The majority of the respective collision products will fall in the next mass range \([0.5 - 0.9]M_\odot\) referred hereafter to as "target" mass range. Recall that the actually observed cluster RGB stars, which have been used

| Table 1. Spearman’s rank correlation test. |
|------------------------------------------|
| Mass range  | \(N_{GC}\) | \(\rho^a\) | \(P(\%)^b\) |
|-------------|-----------|---------|-------------|
| Total       | 35        | -0.172  | 70.0        |
| \(M_{GC} > 10^{5.3}M_\odot\) | 12    | -0.664  | 98.3        |
| \(M_{GC} < 10^{5.1}M_\odot\) | 12    | -0.483  | 89.0        |

\(a\) Spearman’s \(\rho\) coefficient of anti-correlation between the MF slope \(\alpha_{pd}\) and \((N_G1/N_{tot})\).

\(b\) The probability of non-random deviation of the \(\rho\) coefficient from zero.
(in the centers of GCs) by Milone et al. (2017) to estimate the parameter \((N_1(G1)/N_{tot})\), are the descendants of former MS stars fallen in this mass range, with their mass \(\sim 0.8 M\odot\) (primordial MS stars) or somewhat higher (collision products). Note that the requirement about collisions between G1 stars exclusively belonging to the source mass range is a too strong constraint, since these G1 stars might also collide with G1 stars from the target mass range by forming G2 stars with mass that falls within the target mass range. In the framework of our scenario, this latter collision channel would be more favorable for the formation of G2 stars from the point of view of resolving the mass-budget problem. Indeed, this channel implies that only one G1 star from the source mass range would be consumed to contribute one G2 star to the target mass range in contrast to our model requiring two G1 stars from the former mass range to form one G2 star for the latter range.

As an IMF for the G1 stars, we employ a Galactic field-like mass function (Kroupa 2001, 2002). This is just a first approximation. It is evident that the real picture is more complicated and significantly nuanced. For convenience, we accept all the collisions in any initial GC as if they occurred at a time, sometime after the formation of primordial generation of stars G1.

If \(N_1(G1)\) and \(N_2(G1)\) are the total numbers of stars of the first generation that fall in the mass range \([0.1 - 0.5] M\odot\) and \([0.5 - 0.9] M\odot\), respectively, then the fractions of these two populations of stars to the initial total number of stars \(N_{tot}\) can be written as:

\[
f_{1,in} = \frac{N_1(G1)}{N_{tot}},
\]

and

\[
f_{2,in} = \frac{N_2(G1)}{N_{tot}}.
\]

If \(N_{1,ext}(G1) = f_{ext} N_1(G1)\) is the total number of stars that have experienced a collision then the new total number of stars becomes:

\[
N_{tot} = N_{tot}' - \frac{f_{ext} N_1(G1)}{2}
\]

After the collision process, a second generation of stars (i.e., G2 stars) is formed in the mass range \([0.5 - 0.9] M\odot\). The new fractions of stars in the above-mentioned mass ranges will be given by:

\[
f_{1,fl} = \frac{N_1(G1) - f_{ext} N_1(G1)}{N_{tot}} = \frac{N_1(G1)(1-f_{ext})}{N_{tot} - f_{ext} N_1(G1)/2} = \frac{2f_{1,in}(1-f_{ext})}{2-f_{ext} f_{1,in}}
\]

and

\[
f_{2,fl} = \frac{N_2(G1)+f_{ext} N_1(G1)}{N_{tot}} = \frac{N_1(G1)(f_{2,in} / f_{3,in} + f_{ext}/2)}{N_{tot} - f_{ext} N_1(G1)/2} = \frac{2f_{2,in} + f_{ext} f_{1,in}}{2-f_{ext} f_{1,in}}
\]

The value of \(g_1\) is defined for MS stars in the mass range

\(N_{1,ext}(G1)\) is the total number of stars of the first generation that fall in the mass range \([0.1 - 0.5] M\odot\) and \([0.5 - 0.9] M\odot\), respectively, while 11 GCs with intermediate mass are shown by the black dots.

**Figure 2.** Dependence of \(N_{G1}/N_{tot}\), the fraction of G1 (first generation) stars, on power-law MF slope \(\alpha_{pd}\) characterizing the present-day MF of Galactic globular clusters. Upper panel: general dependence for all the 35 Galactic globular clusters with available data; error bars are shown with cyan lines and the continuous line is a linear fit to the data. Middle panel: the deviations of individual points from the line fitted to the data as a function of the GC mass; Lower panel: the same as in the upper panel, but the dependence is shown for three subsamples of GCs separated by mass. The red and blue dots are 12 GCs of higher mass (\(M_{GC} > 10^{5.3} M\odot\)) and 12 GCs of lower mass (\(M_{GC} < 10^{5.1} M\odot\)), respectively, while 11 GCs with intermediate mass are shown by the black dots.
[0.5−0.9]M⊙ like its analog for RGB stars (i.e., N1(G1)/Ntot) as:

\[
\alpha = \frac{N_1(G1)}{N_2(G1) + N_3(G2)} = \frac{N_1(G1)}{(N_2(G1) + f_{ext}N_1(G1)/2) = 0.4 + 0.4 \frac{f_{2, in}}{f_{ext}f_1, in}}
\]

(6)

Neglecting the effects of dynamical evolution during the collision process, the model values of the slope in the post collision MF (αpc) would be the result of the stellar collision process. For any functional form of the IMF of G1 stars, and for a given value of fext, we can estimate the values of f1, in, f2, in, which are used to derive the values of f1, f1 and f2, f1. With f1, f1 and f2, f1, we can then derive the values of gi and αpc. We can calculate αpc as being

\[
\alpha pc = \frac{\log(f_{2, f1}) - \log(f_{1, f1})}{(\log(0.7) - \log(0.3))},
\]

where 0.3 and 0.7 (in units of M⊙) are the central values in the mass bins [0.1, 0.5]M⊙ and [0.5, 0.9]M⊙, respectively. We recall that the values of αpc characterize the present-day MF in the observational sample of GCs vary between ≈ 0.05 and ≈ −1.25 and these values have been derived by Ebrahimi et al. (2020) assuming that the present-day MF of GCs is a single power law over the mass range [0.2 − 0.8]M⊙. The total range considered for the two mass bins in the model is [0.1 − 0.9]M⊙ and it encompasses the stellar mass range over which the value of αpd is estimated in the observations. When calculating the fractions f1, in and f2, in, and subsequently f1, f1 and f2, f1, we assume that the minimum stellar mass is Mmin = 0.01 M⊙ and the maximum stellar mass is Mmax = 150 M⊙. We build a series of models relying on a Kroupa-like IMF (i.e., the pre-collision IMF) and with various values of fext. For each model, we derive the corresponding values of gi and αpc. We show in Figure 3 a set of models that tie the value of gi to αpc for various values of fext and which all employ the Kroupa IMF for the pre-collision G1 population. We can now compare it with the relation between the parameter (N1(G1)/Ntot) and αpd in Figure 2.

Figure 3 shows that for the canonical Kroupa IMF, the capacity of the source mass range (i.e., the initial fraction of primordial MS stars formed in this mass range) would be sufficient to form large fractions of G2 MS stars, depending on the fext value, comparable to those empirically derived on the RGB in many GCs. In other words, the model fraction gi of G1 MS stars left in the target mass range [0.5 − 0.9]M⊙ after collisions have occurred, follows the same trend in terms of its dependence on αpd as the one seen in the observations, in general agreement with the empirically deduced parameter (N1(G1)/Ntot). The parameter fext can formally run from 0 to 1 and result in lower and lower values of the fraction gi that would be more and more approaching to the minimum values of the parameter (N1(G1)/Ntot) in GCs. However, the real maximum value of fext defining the maximum fraction of G1 stars from the source mass range to be converted into the maximum fraction of G2 stars (minimum gi) in the target mass range, can hardly approach to 1. Our numerical experiments show that values of fext > 0.6 yield too large values of αpc. So, we accept fext ≈ 0.6 as being around the limiting value in this respect. Coming back to Figure 3, the minimum value of the parameter gi (in the target mass range) for the limiting case of fext = 0.6 is about gi ≈ 0.40. fext = 0.5 would perhaps be more realistic value for the limiting case.

It should be taken into account that an additional fraction of G1 stars from the source mass range will be converted into G2 stars within the same mass range. The spirit of this simple model is that the fraction of G2 stars is expected to be (much) smaller than G1 stars in the source mass range, in contrast to the target mass range. This additionally implies that gi in this mass range would normally be higher (larger fraction of G1 stars) than gi in the next mass range, [0.5 − 0.9]M⊙, and therefore on the RGB, too. Let us consider, for simplicity, a limiting case where all the formed G2 stars are kept in the source mass range, [0.1 − 0.5]M⊙. Then, given that two merging G1 stars result in a single G2 star, we immediately conclude that the fraction of G2 stars should remain less than the fraction of G1 stars as long as fext < 2/3. In reality, this is impossible, since at least half of the G2 stars fall in the target mass range. So, the fraction of G2 stars in the source mass range (especially at MMS < 0.2 − 0.3 M⊙, see next Subsection) should be smaller or even much smaller (at a given value of fext), than in this limiting case. Given the probable partition between maximum fractions of G1 stars from the source mass range, consumed for the formation of G2 stars in both mass ranges, we finally come to the total limiting fraction of G1 stars allowed to collide in the source mass range around 0.70. This corresponds to 35% of the total number of stars in an initial GC at these most favorable conditions.

As we noted at the beginning of the given Subsection, the above estimate of gi is an approximation. In principle, we should apply a correction reflecting the obvious fact that a portion of G1 stars from the target mass range, [0.5−0.9]M⊙, will participate in collisions and form G2 stars. Therefore the fraction f2, in will effectively be smaller, resulting in a lower gi. Instead of f2, in in Equation 6, we then use f2, in(1 − fextgi) where the factor (1 − fextgi) softly takes the above-mentioned fact into accounts. It would be natural to apply
a stronger factor \((1 - f_{\text{ext}})\) if the reducing population of G1 stars in this mass range was not simultaneously blended with increasing population of G2 stars mostly originated from the source mass range. So, in order to be more conservative and realistic, we multiply \(f_{\text{ext}}\) by the first approximation value of \(g_1\). The corrected value of \(g_1\) will finally be around \(g_1 \approx 0.32\) for \(f_{\text{ext}} = 0.6\) and \(g_1 \approx 0.38\) for \(f_{\text{ext}} = 0.5\).

Figure 3 shows that the model dependence of \(g_1\) on \(\alpha_{\text{pc}}\) reproduces, on the whole, the empirical dependence we obtained for a sample of Galactic GCs (Figure 2), both in terms of the slope and in term of the absolute values. It is evident that the nature of the empirical dependence is more complex as compared with the simplistic nature of the toy model presented here for a particular case of the IMF. In addition to the subsequent dynamical evolution of the GCs, the \((N_1(G1)/N_{\text{tot}}) - \alpha_{\text{pd}}\) relation (or equivalently the \(g_1 - \alpha_{\text{pc}}\) relation) could potentially be affected by primordial clusterto-cluster variations of some of the clusters properties. These variations could be either directly measurable ones (e.g., mass, metallicity) or be more subtle such as variations in the IMF of the G1 population, initial mass segregation ratios, and binary fractions. We will explore the effects of some of these variations in future work.

Finally, we would like to stress, in other terms, the following important detail we have already noted in the paper text. MS G1 stars from the mass range \([0.1, 0.5]M_\odot\) are the largest reservoir of these stars, which can merge/collapse and form G2 stars in the target mass range (and in the source mass range, too). However, the resulting \(g_1\) parameter we estimate in the target mass range, \([0.5, 0.9]M_\odot\), can be smaller (i.e., a smaller fraction of G1 stars in this mass range) at a smaller number of G1 stars consumed in the source mass range. It is due to an additional important factor favoring a smaller proportion between the fraction of G1 and G2 stars in the target mass range. It is the subtraction of G1 stars from the target mass range, \([0.5, 0.9]M_\odot\), for the formation of G2 stars (especially if these G2 stars will be left within this mass range). For example, \(0.6M_\odot + 0.2M_\odot = 0.8M_\odot\), i.e., a \(0.6M_\odot\) G1 star will be subtracted from the target mass range and consumed for the formation of a \(0.8M_\odot\) G2 star in the same mass range.

3.4 The scenario’s implication: a verifiable effect

The collision-based scenario for the origin of G2 stars described above and which is able to explain the trend in the \((N_1(G1)/N_{\text{tot}}) - \alpha_{\text{pd}}\) relation has a number of other implications which can be verified (or tested). Perhaps the most direct would be evidence of the capability of the merging of low-mass stars (at least with \(M_{\text{MS}} \sim 0.4M_\odot\)) to produce, in the collision products, the elemental abundances compatible to those observed in G2 stars. At present, there is unfortunately no definite information on this subject and there is an uncertainty in the understanding of the processes occurring in merging low-mass stars. The characteristics of the finally formed merger, including its surface elemental abundances, are poorly known. Below, we illustrate this with recently obtained relevant observational results (effects), unexpected earlier.

We draw attention to the so-called luminous red novae (LRNe), a class of optical transients. Below, we refer to a particular case of LRNe, which can be the most relevant concerning low-mass mergers in GCs. On the whole, these red transients of different maximum luminosity were interpreted as due either to the merging of the companions of close binary systems (e.g., Soker & Tylenda 2003; Kulkarni et al. 2007) or to common envelope evolution of binary systems (e.g., Blagorodnova et al. 2017). Observational study of LRNe and of their remnants have already brought more understanding of the nature and characteristics of their most probable progenitors and manifested elemental abundances. It has been established that the progenitors are binary systems not only of high and intermediate mass. Among the known LRNe detected to date, there is a low-mass counterpart, V1309 Sco, the most relevant case for our scenario. Its outburst, with maximum luminosity exceeding \(10^6 L_\odot\), has been caused by the merged stars of a contact binary with the mass of the main companion \(\sim 1M_\odot\) (Tylenda et al. 2011) or, more probably, slightly higher, \(\sim 1.5M_\odot\). This is now the least massive binary system among those presently known to give rise respectively to the least luminous LRN (Howitt et al. 2020). Recently, Ferreira et al. (2019) studied this merger 10 years after the outburst. They found a dramatic change in its near-IR colour, among other details, and concluded that the post-outburst state of the merger is consistent with a blue straggler star. Kamińska et al. (2018) studied cool molecular outflows produced by three best-known Galactic LRNe, including V1309 Sco, using ALMA observations. Interestingly enough, they found that the molecular composition of these outflows “appear similar to those of oxygen-rich envelopes of classical evolved stars”, such as red super-giants or (post-)asymptotic giant branch stars. Although these results are not directly linked to multiple populations, they are linked to our scenario as they clearly indicate that our knowledge of the common-envelope phase of close binary evolution and of its outcome is essentially incomplete. The phenomenon of LRN was not predicted theoretically (e.g., Goranskij et al. 2016). In relation to this, we draw attention to the particularly challenging step of the formation of a merger/(collision product) of two MS stars. It is the final formation of its one whole hydrogen burning nucleus from two merging ones. There is significant uncertainty about this process and in particular about the increase in the central temperature as a result of the merger. If the central temperature could temporally achieve at least \(\sim (18 - 20) \times 10^6\)K or exceed this value, then the CNO-cycle would be activated for some period in a low-mass merger/collision product. This is the most important requirement that should be met regarding low-mass merger/collision products in order to explain the difference of (at least) their surface CNO abundance from that of G1 stars in the framework of our scenario. Another requirement which has to be met is the availability of a mechanism capable of transporting the products of the CNO-cycle to the external parts of MS G2 stars. The likelihood for such a mechanism to operate is high for low-mass merger/collision products. Indeed, they are expected to be fast rotators at least in the early stages after their formation. So, rotational mixing is the likely mechanism that is able to facilitate the transport of synthesized light elements to the stellar surface.

We focused, however, on the scenario’s key implication that could be compared with the available observations. We came to the following outcome. Our scenario implies that the parameter \(g_1\) can vary along the MS. More specifically, as we explained above, it is expected to be, on average, larger (i.e., a higher fraction of G1 stars) in the source mass range than
at the upper MS and up to the upper RGB. In particular, it should be true at $M_{MS} < 0.3 M_\odot$. In the framework of our scenario of the collisional/merging origin of the bulk of G2 stars, we expect a significantly decreased or even a lack of G2 stars (statistically), as compared with their G1 counterparts at $M_{MS} < 0.3 M_\odot$. Indeed, if the low mass limit for hydrogen burning (primordial) MS stars is close to 0.1 $M_\odot$, therefore the low mass limit of collisionally formed G2 stars should strictly speaking be around $0.1 M_\odot + 0.1 M_\odot = 0.2 M_\odot$. It is the lower part of the source mass range that is expected to be least enriched with G2 stars, since the primordial stars capable of forming collision products with such a mass (i.e., $M_{MS} < 0.3 M_\odot$) fall within a narrow mass range above the low limit ($M_{MS} \sim 0.1 M_\odot$) of the source mass range. It turns out that Milone et al. (2012, 2014, 2019) estimated the fraction of MS G1 stars below the so-called MS knee (i.e., with $M_{MS} < 0.3 M_\odot$) in the GCs NGC 2808, M4, and NGC 6752, respectively. While the fraction of G1 stars on the RGB and the lower MS in NGC 6752 is formally (the authors excluded from consideration the lowest part of the cluster MS where the G1 branch is more extended to a lower luminosity) the same within the error, the situation in two other GCs is strikingly different, especially in NGC 2808. Indeed, in the latter GC, the fraction of G1 stars at the bottom of the MS has been estimated to be around $0.65$ (and characterized from photometry by "primordial helium and enhanced carbon and oxygen abundances"), as compared to the fraction of G1 stars on the RGB around 0.232 (Milone et al. 2017). In M4 the fractions of G1 stars at the same evolutionary sequences are $0.38$ and 0.285 (Milone et al. 2017), respectively. These are only three GCs and it is therefore too early to draw any definitive conclusions. However, these results in their actual appearance are rather in agreement with our scenario. A varying proportion between the fractions of G1 and G2 stars with varying stellar mass seems to be less compatible with the canonical way of the formation of G2 stars from a reservoir of gas with a modified elemental abundances appropriate to G2 stars, unless the IMF of the second generation stars is essentially different from that of the primordial generation.

Finally, we note that the most massive GCs with their very small fraction of G1 stars on the RGB, i.e. a low value of the parameter $(N_1(G1)/N_{tot})$, around 0.2, seem to be a somewhat special case. One can assume, that the significant contrast between the fraction of G1 stars on the the RGB and on the bottom MS in NGC 2808 is the feature appropriate to the most massive GCs. Partially, it is probably due to the radial effects in GCs we have already mentioned above, namely since the estimates of the parameter $(N_1(G1)/N_{tot})$ was done by relying on samples of the centrally located RGB stars in GCs. We cannot exclude that this contrast may additionally depend on metallicity in the sense that it might, on average, be higher at higher metallicity among the most massive GCs. A possible effect of the merging of GCs on the formation of very massive GCs and therefore of their multiple stellar populations should also be kept in mind.

4 SUMMARY AND CONCLUDING REMARKS

The present paper is devoted to the origin of multiple stellar populations in Galactic GCs. Two generations of stars, G1 and G2, are presently known to typically populate monometallic Galactic GCs. The origin of G2 stars remains an unclear and controversial issue in spite of various plausible astrophysical mechanisms that have been proposed to be responsible for the multiplicity. The current situation is somewhat paradoxical in the sense that thanks to these investigations the origin of generation G2 has become even more mysterious. We report, for the first time, evidence for a potentially crucial role of binary-mediated collisions between G1 low-mass MS stars in the formation of G2 stars in environments of high stellar densities (i.e., the centers of GCs). Such collision products do differ from the primordial counterparts of the same mass. If GCs form with the canonical Kroupa IMF similar to the Galactic field, the characteristic mass of the IMF is around 0.2 $M_\odot$ and $\sim 50\%$ of G1 stars have masses in the range $[0.1 - 0.5] M_\odot$ (or $\sim 85\%$ in the range $[0.01 - 0.5] M_\odot$). The IMF evolves with time due to both collisions and dynamical interaction between these low-mass stars. In contrast to the latter, the former process retains these G1 stars in GCs in the form of new stars, i.e. by converting G1 stars into G2 ones of a higher mass. Thereby, it is expected to gradually modify the IMF in the sense: the shallower actual MF of GCs, the smaller is the fraction of G1 stars in the GCs, for a fixed IMF. We explored available data on both the present-day MF and relative number of G1 red giants ($N_{G1}/N_{tot}$) for a sample of 35 GCs. We find that an anti-correlation exists for GCs of high mass ($M_{GC} > 10^{6-7} M_\odot$) and is statistically significant at the 98.3% confidence level. The scatter is larger in this relation for the least massive GCs ($M_{GC} < 10^{5-6} M_\odot$) and the anti-correlation is less significant (i.e., 89% confidence level). The importance of the binary-mediated stellar collisions and their prevalence against mass loss by dynamical interactions in high-mass GCs are supported by (a) very significant anti-correlation we found between $(N_{G1}/N_{tot})$ and the stellar encounter rate in GCs and (b) a smaller dynamical disruption of more massive GCs, as predicted by recent simulations.

The populations of both LMXBs and millisecond pulsars in GCs are closely related to the formation of close binaries in the dense central regions of the parent GCs. It seems highly likely that these objects and G2 stars share the same formation mechanism. All these populations in GCs show significant correlation with the stellar encounter rate. We found strong anti-correlation between G1 population and stellar encounter rate.

We note that Wang et al. (2020) consider the role of stellar mergers in the formation of multiple populations in GCs with masses lower than $M_{GC} = 3.2 \times 10^{5-6} M_\odot$. However, the authors primarily focus on super- and massive stars and rely on the canonical mode of the formation of G2 stars from the gas ejected by such massive and super-massive stars of population G1. Earlier Sills & Glebbeek (2010) studied the possible contributions of stellar collisions between massive stars to the formation of multiple populations in GCs. They found that such collisions can produce material with elemental abundances comparable to those observed in G2 stars, but the total mass of the material produced is an order of magnitude lower than the required amount. In contrast, our approach advocates for an important role of collision-induced merging of stars belonging to the lower-mass end of the IMF. The collisional formation of G2 MS stars helps to avoid the severe problem of the availability of a sufficient reservoir of...
gas, since this channel, thanks to a big reservoir of available MS G1 stars, is potentially capable of being the main contributor to the population of G2 stars in GCs. Also, our results point out to a higher efficiency of this mechanism among higher-mass GCs. Of course, a more reliable conclusion about the essential role of collisions between low-mass stars in the formation of G2 stars implies that this mechanism should meet another important requirement. Namely, the ability of the merging of low-mass stars, if only they had the mass of $M_{MS} \sim 0.4M_\odot$, to result in the elemental anti-correlations, such as Na-O, Al-Mg, observed in GC stars from the RGB tip down to the MS turnoff region, i.e. in stars with mass around 0.8$M_\odot$. In this regard, we stress the importance of the promising observational results on stellar mergers giving rise to the outbursts observed as luminous red novae, a class of optical transients. Especially interesting and relevant case is a merger of a low-mass contact binary system, with the mass of its main companion $\sim 1M_\odot$, which erupted as a LRN (V1309 Sco). Further work is needed to quantify the relative importance of each of these different mechanisms for the formation of G2 MS stars.

Our results would imply that being negligible or none in low to intermediate density environments, the formation of stars from the merger of other stars can be efficient in regions of high stellar density. Such conditions are appropriate for the central parts of GCs. We estimate that $\sim 1/3$ of the initial number of low-mass stars in a GC should have merged to produce the present-day fraction of G1 stars $\sim 0.35$ on the RGB. Hence, we suggest that the collisional nature of G2 stars in GCs can help shed more light on other surprising results obtained for astrophysical objects where stellar encounter rates are high.

From APOGEE data gathered in the Galactic bulge, Schiavon et al. (2017) have discovered a population of field stars with high [N/Fe] ratio correlated with [Al/Fe] and anti-correlated with [C/Fe] ratios, that is typical for G2 stars in GCs. However, in contrast with the bimodal metallicity distribution function (MDF) of Galactic GCs, the MDF of these stars is unimodal with its maximum around $[\text{Fe/H}] \approx -1.0$ that corresponds to the minimum of the GC MDF in the Milky Way and other galaxies. So, in situ formation of these stars seems to be more probable than their origin from disrupted GCs.

5 ACKNOWLEDGMENTS

The authors thank the anonymous referee for useful comments that improved the manuscript. F.A. Calderón acknowledges partial support from CONICYT PAI through Grant No. 79170075. J.A. Belinchón is supported by CONICYT through FONDECYT Grant No 11170083.

DATA AVAILABILITY

The data underlying this article are available in tables from the papers referred to here (at https://doi.org/10.1093/mnras/stw2531; https://doi.org/10.1093/mnras/staa069; https://doi.org/10.1093/mnras/stx1856; https://doi.org/10.1088/0004-637X/766/2/136).

REFERENCES

Bahramian A., Heinke C. O., Sivakoff G. R., Gladstone J. C., 2013, ApJ, 766, 136
Bastian N., Lardo C., 2018, ARA&A, 56, 83
Bastian N., Covey K. R., Meyer M. R., 2010, ARA&A, 48, 339
Balogorodnova N., et al., 2017, ApJ, 834, 107
Carretta E., et al., 2009, Astronomy & Astrophysics, 505, 117
Dalessandro E., et al., 2013, The Astrophysical Journal, 778, 135
Decressin T., Meynet G., Charbonnel C., Prantzos N., Ekström S., 2007, A&A, 464, 1029
Decressin T., Baumgardt H., Charbonnel C., Kroupa P., 2010, A&A, 516, A73
Dib S., 2014, MNRAS, 444, 1957
Dib S., Kim J., Shadmehri M., 2007, MNRAS, 381, L40
Dib S., Schmeja S., Hony S., 2017, MNRAS, 464, 1738
Ebrahimi H., Sollima A., Hagni H., Baumgardt H., Hilker M., 2020, MNRAS, 494, 4226
Ferraro F. R., et al., 2009, Nature, 462, 1028
Ferreira T., Saito R. K., Minniti D., Navarro M. G., Ramos R. C., Smith L., Lucas P. W., 2019, MNRAS, 486, 1220
Geller A. M., et al., 2017a, ApJ, 840, 66
Geller A. M., Leiner E. M., Chatterjee S., Leigh N. W. C., Mathieu R. D., Sills A., 2017b, ApJ, 842, 1
Glebbeek E., Pols O. R., Hurley J. R., 2008, Astronomy & Astrophysics, 488, 1007
Goranski V. P., et al., 2016, The Astrophysical Bulletin, 71, 82
Hollyhead K. D., et al., 2017, MNRAS, 465, L39
Hollyhead K. D., et al., 2018, MNRAS, 476, 114
Howitt G., Stevenson S., Vigna-Gómez A., Justham S., Ivanova N., Woods T. E., Neijssel C. J., Mandel I., 2020, MNRAS, 492, 3229
Kamiński T., Steffen W., Tylden R., Young K. H., Patel N. A., Menten K. M., 2018, A&A, 617, A129
Krause M., Charbonnel C., Decressin T., Meynet G., Prantzos N., 2013, A&A, 552, A121
Kravtsov V., 2019, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 61, 122
Kravtsov V. S., 2020, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 61C, 57
Kravtsov V., Calderón F. A., 2021, AJ, 161, 7
Kroupa P., 2001, MNRAS, 322, 231
Kroupa P., 2002, Science, 295, 82
Kulkarni S. R., et al., 2007, Nature, 447, 458
Leiner E., Mathieu R. D., Geller A. M., 2017, ApJ, 840, 67
Li C., Tang B., Milone A. P., de Grijs R., Hong J., Yang Y., Wang Y., 2021, ApJ, 906, 133
Mackey A. D., Broby Nielsen P., 2007, MNRAS, 379, 151
Mackey A. D., Broby Nielsen P., Ferguson A. M. N., Richardson J. C., 2008, ApJ, 681, L17
Meylan G., Heggie D. C., 1997, A&ARv, 8, 1
Milone A. P., Bedin L. R., Piotto G., Anderson J., 2009, A&A, 497, 755
Milone A. P., et al., 2012, ApJ, 754, L34
Milone A. P., et al., 2014, MNRAS, 439, 1588
Milone A. P., et al., 2017, MNRAS, 464, 3636
Milone A. P., et al., 2019, MNRAS, 484, 4046
Milone A. P., et al., 2020, MNRAS, 491, 515
Niederhofer F., et al., 2017, MNRAS, 465, 4159
Piotto G., et al., 2015, AJ, 149, 91
Pooley D., et al., 2003, ApJ, 591, L131
Reina-Campos M., Kroupa J. M. D., PfefJer J., Bastian N., Crain R. A., 2018, MNRAS, 481, 2851
Sarajedini A., et al., 2007, AJ, 133, 1658
Schiavon R. P., et al., 2017, MNRAS, 465, 501
Sills A., Glebbeek E., 2010, MNRAS, 407, 277
Sills A., Karakas A., Lattanzio J., 2009, ApJ, 692, 1411
Soker N., Tylden R., 2003, ApJ, 582, L105
Sollima A., Baumgardt H., 2017, MNRAS, 471, 3668
Tylenda R., et al., 2011, A&A, 528, A114
Valcarce A. A. R., Catelan M., 2011, A&A, 533, A120
Ventura P., D’Antona F., 2009, A&A, 499, 835
Wang L., Kroupa P., Takahashi K., Jerabkova T., 2020, MNRAS, 491, 440
Wei D., Wang B., Chen X., Chen H., Li L., Jiang D., 2020, MNRAS, 493, 5479