Introduction

We thank the referee for a thoughtful and elaborate report. We address the comments in sections with headings that match those of the referee report, starting with the following listing of our response to the first 2 paragraphs in the introduction.

1. We updated the abstract and conclusion to clarify that we meant to describe what would be needed for the environmental dependence of the formation and disruption of bound clusters for the Auriga simulations to be consistent with the observations, rather than to argue that the Auriga simulations reflect these trends.

2. We feel that our conclusion would be too strongly phrased if we would generalise our results obtained using the Auriga simulations to cosmological simulations in general. We find that age-selected stellar particles in the Auriga simulations do not reproduce key properties of observed GC systems in the Milky Way and M31. However, our results might reflect characteristics that are specific to Auriga, e.g. that the galaxy formation model overmixes metals at early times. Therefore we did add a final bulletpoint to our conclusion to clarify the above, but we limit the scope to the Auriga simulations. Thanks for pointing this out!

   Furthermore, we clarified in the manuscript that we investigate the metallicity and galactocentric radius because these properties are available for the observed GC systems and in the simulations.

General comments about style

- Numerical quantities are now described in arabic numbers instead of words.
- We now first use the name of the physical quantities in the given example, followed by the symbolic representation.
- The text has been updated to reduce the usage of parenthesis.
- Colloquial expressions have been rephrased.
- We replaced 'Sec.' by 'Section' to refer to sections in the present work. Moreover, we changed 'Sec.' to 'section' to refer to a section in one of the references.
- All acronyms are introduced the first time they appear in the text (MW, M31, GC, GCS, YMC, ISM, SN, AGB, DM, and Au). The subset thereof that appears in the abstract is also introduced the first time it appears.
- The colour palette has been updated to improve colourblind-friendliness.
- We no longer use blue and red to prevent confusion because they are typically used for metal-poor and metal-rich populations.

Major comments and concerns

- We updated the introduction to make it more clear why we consider the metallicity and galactocentric radius in light of *in situ* formation or accretion of GC candidates.
- We added section 5.3 to discuss how our results change of we adopt an age cut of 8 or 6 Gyr.

Additional major comments

- We added a section 5.3 to discuss how our results change if we adopt an age cut of 8 or 6 Gyr instead of 10 Gyr.
- We clarified that the number of observed GCs differs between Figure 1 and Figure 2 because the latter shows metallicities between -2.5 and 0, galactocentric radii between 1 and 250 kpc. Figure 1, on the other hand, shows the full range of radii, and the top panel of Figure 3 shows the full range of metallicities.
- We normalise the Auriga simulations by the virial radius of the dark matter halo to compensate for scatter between different simulation runs. We deliberately chose for r_{vir} because there are several stellar length scales that could be used (e.g. the effective radius of the bulge, the radial scalelength or vertical scaleheight of the disk, the optical radius, the extent of the stellar halo, the half-mass stellar radius, etc). The simulations show a wide variety between properties of the Auriga galaxies with little correlation between these different stellar length scales. Therefore we could introduce unknown biases by normalising to any one of these options which is why we use the virial radius instead.

Comments for each section

Introduction

- We clarified that the blue and red GCs are observed to have different colour distributions that may reflect a different metallicity distribution which is argued to arise from two distinct formation channels. We briefly summarise that differences between their properties include the metallicity and radial distributions, which is why we investigate these properties in the Auriga simulations.
- We removed the quoted sentence.

Sect. 2 -

- We added the halo mass range of the Auriga simulations.
- The last paragraph should now make more sense after updating the introduction.

Sect. 3 -

- We added references to section 3.3, 3.4, and 3.5 in the first paragraph of section 3.
- Section 3.2 has been restructured.
- We removed the standard error on the mean and replaced 'dispersion' by 'standard deviation' in section 3.3.
- We added a sentence to clarify the last statement in section 3.4.

Sect. 4 -

- We added the mean number of GC candidates compared to the mean number of star particles in the Auriga L4 galaxies. In addition, Figure 3 now shows the number of star particles included in each subset.
- We clarified at various places whether we mean the mean or median among the simulations, or if we refer to the absolute value of one of the galaxies.
- We added the suggestion in the 3rd paragraph of section 4.1.
- Bimodality in GC systems is now discussed in the introduction.
- Figure 4 - now has a vertical line to indicate the metallicity cut between metal-poor and metal-rich GC subpopulations in the Milky Way.
- We added a sentence about the cluster formation efficiency model of Kruijssen (2012) in the last paragraph of section 4.1.
- We now show the median instead of the mean for the galactocentric radius in section 4.2 and Figure 6.
- Figure 6: no changes were made for rescaling the galaxies (i.e. we still use the virial radius)
- Figure 5 and 7: shaded regions now show the 25-75th percentiles.
- We added a brief discussion of Figure 8.
- We clarified that we are referring to Figure 5 and Figure 7 for the claims about the *in situ* and accreted component in the third paragraph of section 4.3.

- Figure 9: The values are removed from the masked upper right corner bins.
- Figure 2,8,9: The values of the upper and right axes have been removed from Figure 2 and 9 (observations) for clarity (rather than the suggestion to add axis labels and units). However, we did not mask these bins for Figure 8 (Auriga) because it illustrates our finding that the simulations produce a considerable amount of mass in GC candidates with low metallicities at large radii.

Sect. 5 -

- We added references to Deason et al. (2015) and Conroy et al. (2019) in the first paragraph of section 5.1.
- We included a discussion of bimodality in metallicities in the introduction. In addition to that, the results of Usher et al. (2012) "imply that most early-type galaxies in the SLUGGS survey have bimodal metallicity as well as colour distributions", and "Combined with previous studies, nine of 12 galaxies show evidence for bimodality while in the remaining three galaxies bimodality cannot be ruled out.". Our interpretation of this study seems to differ from that of the referee.
- We still find that the accreted star particles are more metal-poor than those that have formed *in situ* if we lower the age cut to 8 or 6 Gyr. This has been added in section 5.3.
- We removed the quoted sentence.
- We clarified that our selection function is inconsistent with the definition
- We clarified which which galaxy Brodie & Strader provide the quoted number.
- We clarified footnote 8. of Kruijssen (2019) for *ex situ* GC formation.

Sect. 6 -

- We repeat the goal of the study in section 6.
- We clarified the 4th point.

Minor comments

- We clarified that 'blue' GC subpopulation means metal-poor with [Fe/H] $<$ 1
- The units of mass-to-light ratio have been corrected.
- The indicated typos have been corrected.

- ‘missing error bars and references for the virial radii’ → The reference was given in a footnote, but is now moved to the main text.
- 4.1: ‘which model do the authors refer to?’ → added ‘star formation’.
- 4.1: ‘The top half of the left figure’ → ‘The top panel of Figure 3’.