The paper "The Formation of Very Massive Stars in Early Galaxies and Implications for Intermediate Mass Black Holes" investigates star formation in two halos resimulated from a previous simulation including a larger halo sample. These halos were selected as representative of halos where supermassive stars could form.

The paper touches on an important topic, the methods are sound are generally clearly described and the results of great interest.

I detail below a number of comments/questions. First some major points that require care and significant changes to the text (especially the first point, which includes several questions), then minor points.

We thank the referee for a very detailed and informative report which we have used to improve the manuscript. Our answers to the referee's comments can be found inline below.

Major:

- Infall/accretion rate, multiplicity, consequences: Showing the infall rate towards the halo center is meaningless in the situation in which (i) stars are not in the halo center but scattered throughout the protogalaxy and (ii) there is no spherical symmetry (see Fig. 4, left). This is a misleading way of presenting the accretion rate, since it has no bearing with the evolution of the stars in the simulation and I argue it should be removed from the paper. Does is really make sense to talk about "infall rate" in a halo as a check for SMS formation when star formation is widespread and not central?

  The infall rate as presented in Figure 6 is the infall rate immediately prior to the onset of star formation in the haloes. We note that we did not explicitly say this in the caption and this has now been rectified. Our goal here was to allow the reader to see what the central infall rates were prior to the onset of star formation. We have decided to leave Figure 6 as is therefore and address the referee's concerns through the addition of 4 new figures (Figures 7, 8, 9, 10)

  Showing analogs of Fig. 7 for all stars in the simulation, plus information on where they are located within the galaxy and how the accretion rate changes as conditions change would instead provide the correct information.

  We have added Figures 7 and 8 to now display the accretion rates and mass histories of all stars in both Halo A and Halo B. Previously, in the old Figure 7 we showed just the mass accretion rates for the most massive star and the median accretion rates.

  Is the accretion rate on *individual* stars higher when they're located close to the center? Can you show how the accretion rate depends on the local environment? HaloA, as noted in the text, "contain[s] a web of dense gaseous filaments": is accretion enhanced when the sink crosses this filamentary structure? Or is it higher when a star is still and close to the center of mass of the halo?

We have now added Figure 10 which explicitly shows how the environment (distance from a high density clump) affects the accretion rate onto a star. We also show the number density of gas surrounding each star which shows a 1:1 correspondence with accretion rate. Finally, the relative velocity of stars is also plotted. Stars are never “still” and have relative velocities of between 5 and 20 km/s.

In view of having a critical threshold for SMS formation, how this critical threshold is reached and maintained in a realistically evolving halo is very important and this should be thoroughly discussed in the paper. A more thorough discussion on multiplicity is also needed. The standard picture in the synchronized halo case, for instance, is of accretion on a single sink located at the center of the halo. Is the large multiplicity found here specific of the dynamical heating+low-ish LW background and the formation of a bona fide SMS from monolithic collapse is still expected in the synchronized halo picture with a high incident LW radiation? Please address all these points directly in the manuscript.

We have now added a new subsection titled “Stellar Multiplicity” where we analyse the enclosed stellar mass, gas mass and number of stars. This is complemented by Figure 9. The issue of multiple massive stars forming in a halo is addressed in the conclusions where we discuss similar recent results from Chon et al. (2017) and those from Regan et al (2018). We believe that the findings from various research groups have led to a present consensus that monolithic collapse is very difficult to achieve in practice. We compare the similarities of our results to those from the PopIII simulations groups to underline this.

- Choice of the halos and consequences on the results: please significantly expand in the text the discussion on the choice of the halos and the consequences on the results. Which of the halos in Regan+2020b were chosen? Why HaloA with LW radiation and HaloB without and not viceversa? Why not both with and without LW radiation? Why not only one, but with and without LW radiation? This last point is important: HaloB does not form massive stars, but there are two factors (i) no LW radiation, but also (ii) a lower mass (“HaloB, on the other hand, contains essentially only a single site of star formation due to the significantly smaller halo mass.”). HaloB forms less massive stars but also fewer stars overall, so it seems that there is less fragmentation although the absence of LW radiation should favor fragmentation… the text says “”HaloB is therefore the control simulation and allows us to examine the impact of what happens when no LW field is present” but the control would have been sounder if it were the same halo to be simulated with and without radiation. Although I am not asking to run a new simulation at least a stronger justification in the manuscript is needed.

We completely agree that having two different haloes, one without LW radiation and the other without is not ideal. This was not the original plan but these simulations were arguably beyond anything attempted with Enzo before and so compromises had to be made along the way. In essence at the end of our compute time we ended up with HaloA and HaloB which we could report on. Of course all the other simulations, some as we note below proved intractable, contributed to our knowledge of what was achievable.

Running these simulations is extremely expensive in terms of CPU time. We initially selected three haloes from Regan et al. (2020) which demonstrated isothermal collapse at approximately 8000K. We ran the simulations at a maximum refinement of 24 (this is 16 times higher than the simulations reported here.). There was considerable work in understanding the limits of the simulations and of what could and what could not be achieved. We did manage to run the simulations for 10,000 years but this consumed approximately 500,000 core hours and the radiation solver was unable to cope with the steep ionisation gradients found close to the stars.

We have added some additional text to the paper to clarify our reasons for using HaloA with a LW field and HaloB with one (footnote 3).

- Section 3.2: Centering the measure on the most massive star, which can become second massive at some point, makes the interpretation of the results confusing. the most massive star can also be in a non-central region (see Fig. 4) then some of the volume included is out of the baryon-dominated region of the halo (?) A more robust choice should be used.

  Centering on the most massive star is not ideal as the referee points out. However, while the star which is most massive does change it does remain very close to the centre of mass of the halo at all times. We have tested other centre definitions and we found that because the most massive star follows closely the high densities environments it is a good choice. We did not find that changing the centering of the analysis had any effect on the conclusion that ionising feedback does not have an appreciable negative impact on the halo. These results are also backed up by a recent 1D analysis by Sakurai et al. (2020) which appeared on the arXiv after we submitted this manuscript.

Minor: Introduction: “SMSs have been invoked to explain the existence of SMBHs”, this misses logical steps, i.e., the collapse of SMSs can produce BHs that after growing can become SMBHs.

Amended

Introduction: the right-hand column in page 1 discusses that a mass accretion threshold must be “maintained” without ever giving some timescales over which it has to be maintaned. Please add this to the discussion.

Done

Introduction: top of page 2, left column. " furthermore the metallicity of the gas being accreted should be below 1e-3 Zsun”, please add a few words of explanation for why this is so (Do not assume that readers are only experts in this particular field)

Done

Define J21 (normalization and units) the first time it’s used

Done

Section 2.2: define what you mean by “found” in the sentence “the target halo at the redshift at which it was found” and provide halo masses for those redshifts.

This has now been clarified and the (original) halo masses given.

Section 2.3: I don’t understand if the LW “background” inserted in the simulation is spatially homogeneous or it conserves the spatial distribution from the original simulation. If the spatial information is not included, please say it explicitly.

The original spatial distribution is preserved. This is now clarified in the text.

Caption of Figure 1: “LW background rate” what is meant by “rate” and please change the units of measure if this is truly a rate

We have now updated the caption to say “field” instead of “rate” since the units are erg cm^-2 s^-1 Hz^-1 sr^-1

Section 2.4: “The mass of the less massive star is added to the accretion rate of the more massive star for that timestep”: is this mass from accreted stars excluded when calculating the accretion rate to be compared with the critical rate?

The accretion rate onto a star includes mergers that occur during that timestep.

Section 2.4: “All stars in this simulation emit radiative feedback below the ionisation threshold of hydrogen.” However, PopIII stars should emit radiative feedback also above the ionisation threshold of hydrogen.

Yes. This is dealt with in the final paragraph of this section.

The reference to Footnote 3 appears several times (in the text and in two captions)

Yes that’s correct.

Figure 4: please highlight in the figure the positions of all the stars listed in the inset (e.g., with different colors)

We initially coloured stars according to their masses but this quickly becomes unwieldy. Instead we now change the radius of each “star” according to their masses. We believe this to be sufficient. The legend is also, by choice, not complete and is only designed to give a flavour of the stellar masses in the simulations.

By the end of each simulation the total stellar mass in HaloA is 90,000 $M_\odot$ and 1300 $M_\odot$ in HaloB.” Please give also the fraction in terms of DM and baryon mass.

As a fraction of the baryon mass this is already given since we quote the SF efficiency (i.e. what fraction of the baryon mass is converted into stars). The stellar to DM fraction is higher since the mass is baryon dominated at this radius. We now note this in Section 3.5.

IMBH mergers: note that three-body interactions and GW recoils can cut short the growth of an IMBH by mergers unless the mass ratios are significantly far from unity (Gultekin+04,06; Holley-Bockelmann+08). This is the reason why specific conditions are required to form SMBH seeds from mergers of BHs and IMBHs (https://ui.adsabs.harvard.edu/abs/2011ApJ...740L..42D/abstract). Please revise the text accordingly.

We thank the referee for pointing out these conditions. We have now updated the text accordingly.

Inclusion of radiative feedback in dynamical friction would make merging IMBHs much more difficult, since it generally results in an acceleration and not a deceleration (Park & Bogdanovic 17 — note also the Erratum in 2019). In the presence of radiative feedback a dense stellar background is needed to decelerate the IMBH against the negative dynamical friction caused by radiative feedback

The work of Toyouchi et al. (2020) expands on the work of Park & Bogdanovic (2017, 2019) by investigating not only the more rarefied environments but also high density environments similar to those investigated by Inayoshi et al. 2016. Toyouchi et al. (2020) find that in the denser environments with \rho_\inf > 10^6 cm^-3 that the dense clumps provide a sufficiently negative force to cause a declaration of the black hole (for black holes with masses of M_{BH} > 10^4 \msolar). We have updated the text to clarify this.

"The merger of a 10,000 M⊙ binary black hole system will enter the LISA band several 1 month before merger and complete thousands of orbits before the final plunge", this seems optimistic for a binary at z=15. Please state the redshift at which the binary would have to be for staying in band for >1 month with SNR above detection threshold, or remove this sentence: for low SNR systems like these, the mass estimate, in any case, will be obtained from the information at merger (i.e., or a few hours before) and not from the previous (low SNR) cycles.

We have now updated the text here to clarify that the model we describe here is the merger of two 5000 Msolar mass black holes at z = 15 which gives a SNR of ~ 32. This is consistent with what others find and in fact is derived using the equations (and python code) available from Robson et al. (2019). The referee is indeed correct that the higher SNR values are only found in approximately the last day prior to the merger but even in this last dive such a binary will complete a few hundred orbits. We have modified the text accordingly.