

Peer Review File

Manuscript Title: Turbulent cold flows gave birth to the first quasars

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

How the first supermassive black holes, which are observed to power quasars at $z=7$, grew so massive in such a limited period of time has remained a scientific puzzle since their discovery. The problem can be split into two separate questions: a) how and when did the seed black holes, which later grew into supermassive black holes, first form? b) Under which circumstances can these seed black holes grow into supermassive black holes by their observed redshift? The latter part of the question was recently answered by DiMatteo and her research group (Tenneti2018,Huang2019), who showed that low-shear environments, which host massive halos that are fed predominantly by accretion flows, are able to efficiently grow black holes from 10^5 Msun to more than 10^8 Msun. However, the first part of the question (where the 10^5 Msun black hole came from in the first place) remains open and is the one addressed by the authors of this paper.

In this paper, the authors present high resolution simulation of one halo that fulfils the criteria highlighted by Tenneti2018 the formation of seed black holes in such halos. They report that high levels of turbulence suppress early star formation and that gas instead collapses directly into dense clumps which could form black holes in the mass range 10^4 - 10^5 Msun. If confirmed, this would be a key scientific insight that would address one of the main open questions in our understanding of the origin of the first supermassive black holes.

While I do not doubt the importance of the scientific question addressed in this paper, I do have some concerns regarding the methodology, and its consequences for the scientific conclusions drawn from this work.

One of the key results in the paper presented here is that the host galaxy does not form Pop III stars early on, despite efficient cooling. Usually, the formation of massive seed black holes via massive gas clouds, as advocated in this paper, relies crucially on maintain a sufficiently high temperature to dissociate H_2 , either via Lyman-Werner radiation from nearby young stars (Omukai2001,Oh&Haiman2002,Begelman2006) or from repeated injection of energy via repeated mergers (Wise2019,Regan2020). Neither scenario is appropriate here, as the authors explicitly mention that there is no radiating neighbour and the halo is "accretion dominated" i.e. experiences few mergers. H_2 fractions are high and gas cools efficiently. These are usually conditions in which Pop III stars form efficiently, yet the authors report explicitly that no such stars form in their simulations. The authors argue that Pop III star formation is suppressed by the supersonic turbulence in the halo. However, research concerned with the formation of stars in turbulent clouds have shown that stars do continue to form in highly turbulent environments, including the "more

than an order of magnitude above the sound speed" (i.e. $\text{Mach} > 10$) quoted in this paper (see e.g. Federath & Klessen, and Hennebelle & Chabrier)

To substantiate the claims, I would like authors to address the following questions:

1. What criteria were used to check whether Pop III stars form?
2. What impact does the heating on the smallest scale (added to ensure the Jeans length remains refined at all times) have on the fragmentation of gas within the halo, and as a consequence on the formation of stars and on preventing the massive gas clumps reported here from fragmenting?
3. How do conditions in the halo compare to those in turbulent star formation simulations (e.g. the ones discussed in Hennebelle & Chabrier), and what does this mean for the expected star formation rate?

In my opinion, significantly further work is needed to show that the lack of Pop III stars is physically expected in this environment, despite the efficient cooling and the inherently turbulent nature of star forming clouds, and not merely an artefact of the simulation. If Pop III stars are able to form, their feedback would disrupt gas flows further, and the metals released in the process would significantly change the cooling of the gas. Both would most likely prevent the formation of the massive clumps from which the authors claim seed black holes could form. The lack of Pop III stars is therefore at the heart of the result presented here, and needs to be justified much more robustly for the conclusion of the paper to uphold.

Referee #2 (Remarks to the Author):

The manuscript "The turbulent origins of the first quasars" by Latif M. A. and collaborators presents an extremely high-resolution cosmological simulation aimed at assessing the conditions under which massive seed black holes can form in the high-redshift Universe. The work suggests that the large inflow rates onto the progenitor of a massive galaxy host at $z=6$ are responsible for the driving of strong turbulence, which inhibits star formation and favours the catastrophic collapse of massive gas clumps into a supermassive star. By means of the MESA code, the authors show that the accretion of gas onto the protostar cannot be halted, and the star grows up to about $10^4 M_{\text{sun}}$ in a short time.

The study is original and timely, given the ever stronger constraints on the mass of massive black holes at high redshift obtained from observations, and the great interest of the community in this topic.

It also suggests that the formation of black holes via the direct collapse mechanism might simply be the result of cosmic structure formation, and does not require any specific conditions like strong LW background or streaming velocities, contrary to what all previous work claimed.

The methodology of the study is surely valid, and the results are striking. Nonetheless, there are some potential caveats in the study that do not allow to properly assess the robustness and reliability of the results. In particular, I have a few concerns about the simulations, that I report in the following:

- 1) The chemical network used in the simulations is not state of the art, for instance compared to

Glover 2015, where it was shown that there are several processes that should be considered in order to accurately follow gas cooling and fragmentation in primordial conditions, processes not typically included in numerical simulations (to my knowledge, the Enzo chemical network does not include such processes). For instance, apart from the missing chemical species and reactions in the network relative to Glover 2015, the densities in the current simulation are so high that the transition to the LTE cooling rates should have been accounted for, but I did not find any mention of it in the draft. I understand that the simulations are already very expensive, and I am not asking to improve the network and rerun them, but the authors should at least discuss potential caveats of their work in the light of these missing processes.

2) The authors employ a pressure floor in their simulations for cells at the maximum refinement level, in order to force the Jeans length to be resolved. I am concerned that this choice, which the authors justify with the need to stabilise the collapse, might significantly affect their results. In particular, while it is true that the pressure floor helps avoiding numerical fragmentation, this is obtained at the expense of energy conservation, actually injecting energy into the system. This energy/pressure increase might artificially support gas against collapse and inhibit *physical* fragmentation and subsequent star formation that should instead occur, until the mass in the centre becomes extremely large.

In particular, several recent works on star formation in molecular clouds, with similar resolutions to the current study, have opted for not including a pressure floor, but instead introducing sink particles in maximally refined cells where the Jeans length was not properly resolved. Why did the authors not employ this approach, especially given the extremely high resolution achieved?

3) Moreover, when the authors restarted the simulation at a lower level of refinement (0.01 pc), I guess the pressure floor was adjusted accordingly. I wonder whether this instantaneous change can potentially affect the entire halo core, introducing a bias in the subsequent evolution, especially in the light of the lower resolution (which corresponds to not being able to resolve the Jeans length on larger scales).

In general, the article is properly written, and it also gives appropriate credit to most previous works, although not completely (see my comments her below).

Apart from these major concerns, the article is properly written, and also gives appropriate credit to previous works, although not completely. Here below, I detail some additional minor comments to the article that would improve the clarity and quality of the paper:

i) Abstract (references 5/6): Given the achievable resolution and the limitations of the modelling (especially in the paper by Q. Zhu), I find this claim a bit too strong, and a lot of effort is still needed in my opinion.

ii) Abstract (references 9/10/11): I disagree with this sentence. Recently, Lupi et al. 2021 showed that the conditions for the formation of a massive seed are actually favoured in quasar progenitors, and that the only antagonist is metal pollution resulting from the high sfr in the region.

iii) Abstract: “without the need ... , or even atomic cooling”. What the authors mean here? In order to collapse, gas needs to cool down, and in primordial environments cooling is provided by atomic

cooling down to 10^4 K, and molecular cooling below 10^4 . Are the authors claiming that cooling is provided by H₂ only, without resulting in fragmentation?

iv) Main text: Why does the temperature never exceeds 2000 K? Is it because there is molecular gas cooling gas efficiently? Or is the atomic phase simply not resolved in time because when the gas heats up above 10^4 K atomic+molecular cooling bring the temperature down very quickly?

With respect to the turbulence, although I agree with the authors that turbulence can provide support against collapse, it is also true that supersonic turbulence produces shocks that, in case of efficient cooling, can enhance significantly the gas density and favour the formation of small scale fragments that might form stars before enough mass has reached the centre of their massive halo.

v) Main text: About the massive self-gravitating disc, how is this related to the pressure floor? I am wondering whether the same disc would have formed without pressure floor. Later during collapse, the authors say that H₂ formation is boosted via three-body reactions, and additional cooling brings the temperature to drop. I am wondering what is the impact of the pressure floor on the fragmentation that might appear because of this strong cooling. This is important to understand why a massive gas distribution of 4×10^6 can be present in the halo without any star having formed so far.

vi) Main text: The authors claim that C1 and C2 collapse before any star has formed in the halo. I do not find anywhere in the draft how star formation is modelled in the simulation. PopIII stars are also expected to live for a few Myr (assuming a top-heavy IMF), hence metal enrichment via SN feedback might alter the evolution. No details on these processes are given in the manuscript, and are crucial to assess how reliable the presented simulations are. On top of this, even if star formation has been included, it is not clear how the presence of a pressure floor alters the formation of stars in the halo.

vii) Supplementary material (Halo identification): What do the authors mean with “Lagrangian volume of the halo”? How many virial radii does their volume enclose?

Author Rebuttals to Initial Comments:

We thank both referees for their careful inspection of our paper and are happy to address their points below.

General Comments

We see from the comments of both referees that we left out details about pressure floors and star formation in our letter, this was an oversight on our part. We answer questions that both reviewers had here and address their other points below.

Our simulation was partitioned into two stages. In the first, we allowed up to 25 levels of refinement (a maximum resolution of ~ 20 AU) from $z = 150$ down to the onset of collapse in the halo at $z = 25.7$. When the simulation reached 25 levels of refinement during collapse, we halted the run and turned on a pressure floor on the 25th level so gas could not collapse to even smaller scales. We then evolved the halo with this floor for 20 kyr to study its collapse down to ~ 20 AU.

Although we did not have star or sink formation turned on, no normal Pop III stars formed in the halo in this first stage because turbulent velocities prevented the collapse of the gas, even though it was cooled by H_2 . H_2 cooling was also not efficient enough to allow density fluctuations due to supersonic turbulent shocks to collapse to stars, in contrast to star-forming clouds today in which cooling by dust, metals, and molecules can cause the collapse of such fluctuations. Had normal Pop III stars begun to form anywhere in the halo they would have triggered AMR down to the maximum level of refinement allowed, and this never happened before the halo itself began to collapse.

In the second stage of the run we restarted the simulation at a slightly earlier time, when it was only at a maximum of 18 levels of refinement (a resolution of ~ 0.01 pc), and restricted refinement to 18 levels thereafter to be able to evolve the halo for the times required for the two supermassive stars to form and collapse. We set a new pressure floor on the 18th level to prevent gas from collapsing to scales below 0.01 pc. In spite of its large H_2 fractions, the halo does not form normal Pop III stars prior to the two supermassive stars because supersonic turbulence driven by the cold flows modified the Jeans mass of the gas:

$$M_{\text{Jeans}} = \pi/6 (V_{\text{eff}})^3 / (G^{3/2} \rho^{1/2}),$$

where

$$V_{\text{eff}} = \sqrt{c_s^2 + v_t^2}.$$

Here, v_t is the turbulent velocity (see Stacy et al 2011, ApJL, 730, 1), and c_s is the sound speed. After the onset of collapse, densities in the core soon rose above $1e-16$ g/cm³ and $v_t \sim 20$ km/s and $c_s \sim 2$ km/s, which produced

an effective Jeans mass of about 10^4 Msun. Consequently, in spite of H₂ cooling the gas could only fracture on large mass scales and no normal Pop III stars formed, only the two massive clumps that created the 31,000 and 40,000 Msun stars. Although no star or sink formation was turned on during this second stage of the simulation, C1 and C2 were found by our clump finder to be the first self-gravitating clumps to form that were capable of collapsing to stars after the halo began to collapse. No other structures capable of forming stars formed at this time, a result that was corroborated with a test run with sink particles as discussed in a reply to referee 2 below. As we show in the attached plot, cold flows continue to drive highly supersonic turbulence throughout stage 2 (the magenta and red lines are at 1.0 Myr and 1.2 Myr, respectively).

Our results are consistent with Hirano et al. 2017, Sci, 357, 1375 in which supersonic gas flows relative to the DM potential of a halo modify the Jeans mass of the gas and prevent it from collapsing until the halo grows to above 10^7 solar masses and forms a supermassive star. In Hirano et al. 2017, turbulence in a rare halo with 3-sigma fluctuation streaming velocities suppressed star formation until the halo reached a mass of $2e7$ Msun, collapsed because of H₂ and atomic cooling, and formed a supermassive star. Like us, they found that no normal Pop III stars formed before the halo reached this mass. Although the origin of the supersonic turbulence in our study is quite different, its suppression of SF prior to collapse is similar.

The pressure floors we imposed on the highest level of refinement at the end of stage 1 and during stage 2 of our simulation only suppress fragmentation at or below their resolution limits of 20 AU and 0.01 pc. The scales on which collapse leads to normal Pop III SF are larger and would have been resolved in our runs. It is possible for fragmentation to occur in accretion disks on scales below a few tens of AU (e.g., Greif et al. 2011, ApJ, 737, 75; Becerra et al. 2015, MNRAS, 446, 2380; Becerra et al. 2018, MNRAS, 480, 5029) but these low-mass clumps are expected to migrate inwards and merge with the central object on timescales that are much shorter than the Kelvin-Helmholtz contraction time of the object onto the main sequence. This has been found to be true in 3D simulations by Hosokawa et al 2016, ApJ 824, 119 and analytical studies of Pop III disk fragmentation by Latif & Schleicher 2015, A&A, 585, 151 and Inayoshi & Haiman 2014, MNRAS, 445, 1549).

Finally, contrary to what is sometimes thought in the community, H₂ cooling has been shown to play a role in supermassive star formation when large reservoirs of gas and high halo temperatures are present. Only the most extreme LW backgrounds destroy all the H₂ in a halo because it can self-shield in the high densities at the center. Not even energy injection by mergers in Wise et al. 2019, Nature, 566, 85 destroyed it all, even in moderate LW backgrounds, and their halo was still predicted to form a supermassive star. Following up on Wise et al. 2019, Regan et al. 2020, OJA, 3, 14

showed that a $6e3$ solar-mass star formed in the presence of H_2 cooling and Chon et al. 2020, MNRAS, 494, 2851 even reported the formation of a 10^4 Msun star in the presence of metals ($Z/Z_{\text{sol}} < 0.001$). These studies show that any rapid dynamical collapse can produce supermassive stars, regardless of origin (see also Latif & Volonteri 2015, MNRAS, 452, 1026 and Regan & Downes 2018, MNRAS, 478, 5037).

We now clarify these points in the letter.

Referee 1:

>> One of the key results in the paper presented here is that the host galaxy does not form Pop III stars early on, despite efficient cooling. Usually, the formation of massive seed black holes via massive gas clouds, as advocated in this paper, relies crucially on maintain a sufficiently high temperature to dissociate H_2 , either via Lyman-Werner radiation from nearby young stars (Omukai 2001, Oh & Haiman 2002, Begelman 2006) or from repeated injection of energy via repeated mergers (Wise 2019, Regan 2020). Neither scenario is appropriate here, as the authors explicitly mention that there is no radiating neighbour and the halo is "accretion dominated" i.e. experiences few mergers. H_2 fractions are high and gas cools efficiently. These are usually conditions in which Pop III stars form efficiently, yet the authors report explicitly that no such stars form in their simulations. The authors argue that Pop III star formation is suppressed by the supersonic turbulence in the halo. However, research concerned with the formation of stars in turbulent clouds have shown that stars do continue to form in highly turbulent environments, including the "more than an order of magnitude above the sound speed" (i.e. $\text{mach} > 10$) quoted in this paper (see e.g. Federath & Klessen, and Hennebelle & Chabrier).

Excellent point -- as noted above, supersonic turbulent shocks in our halo create local density fluctuations but they do not collapse into stars because H_2 cooling is so less efficient than metal and dust cooling in the galaxy today. Another difference between our models and Federath 2018, Phys Today, 71, 6, 38 and Hennebelle & Chabrier 2008, ApJ, 684, 395 is that they proceed from pre-stellar cores drawn from a stochastic density distribution in the cloud. However, like us, they found that turbulence increased the Jeans mass and shifted the mass spectrum of the clumps to higher values (see Fig 1 of Hennebelle & Chabrier).

>> To substantiate the claims, I would like authors to address the following questions:

>> 1. What criteria were used to check whether Pop III stars form?

Explained in the general comments, but we ran an additional test with sink particles as described in our response to point (2) to referee 2 below. No normal Pop III stars formed

in either approach (now also explained in the Methods section of the paper).

>> 2. What impact does the heating on the smallest scale (added to
>> ensure the jeans length remains refined at all times) have on
>> the fragmentation of gas within the halo, and as a consequence
>> on the formation of stars and on preventing the massive gas
>> clumps reported here from fragmenting?

Explained in the general comments.

>> 3. How do conditions in the halo compare to those in turbulent
>> star formation simulations (e.g. the ones discussed in Henebelle
>> & Chabrier), and what does this mean for the expected star
>> formation rate?

They are dramatically different. Supersonically turbulent clouds today fragment on relatively small mass scales because of efficient metal and dust cooling. Primordial clouds cooled by much less efficient H₂ had thermal Jeans masses that were 100 times higher and are thought to produce stars with characteristic masses of at least a few tens of solar masses. Simple analytic estimates show that mass accretion rates ($\dot{m} \sim c_s^3/G$) in present day clouds are $\sim 10^4$ times lower than in primordial gas clouds ($\sim 100 \cdot (200/10)^{(1.5)}$) because dust and metals can cool cloud cores to ~ 10 K while H₂, Ly α or some combination of the two can only cool gas to 500 - 8000 K. The main drivers of turbulence in present day SF clouds are ionizing UV, jets, winds and SNe while gravitational infall primarily drives turbulence in primeval environments.

As explained in the general comments, H₂ cooling prior to the collapse of the halo is too inefficient to cause density fluctuations due to shocks in the supersonic turbulence in our halo to collapse to stars. After the onset of collapse, the turbulent Jeans mass becomes far too large for normal Pop III stars to form before the two supermassive stars do.

>> In my opinion, significantly further work is needed to show
>> that the lack of Pop III stars is physically expected in this
>> environment, despite the efficient cooling and the inherently
>> turbulent nature of star forming clouds, and not merely an
>> artefact of the simulation. If Pop III stars are able to form,
>> their feedback would disrupt gas flows further, and the metals
>> released in the process would significantly change the cooling
>> of the gas. Both would most likely prevent the formation of the
>> massive clumps from which the authors claim seed black holes
>> could form. The lack of Pop III stars is therefore at the heart
>> of the result presented here, and needs to be justified much
>> more robustly for the conclusion of the paper to uphold.

We're happy to discuss your questions in further detail as needed. We do expect Pop III stars to appear at later times in the halo because X-rays from the DCBH can catalyze H₂ formation and trigger their formation (e.g., Machacek et al 2003, MNRAS, 338, 273; Aykutalp et al. 2014, ApJ, 797, 139; Aykutalp et al. 2020, ApJL, 898, 53) and they play an important role in regulating its growth at later times (Smidt et al. 2018, ApJ, 865, 126). But even if normal Pop III stars had formed before the two supermassive ones they probably would not have been able to interfere with their growth. Recent radiation hydrodynamical simulations indicate that

such stars would have had to be at least 1000 Msun before their ionizing UV flux could disrupt massive inflows in their vicinity in similar environments (Latif et al. 2021, MNRAS, 508, 1756). We now discuss this point in the letter.

Referee 2

>> 1) The chemical network used in the simulations is not state of the art, for instance compared to Glover 2015, where it was shown that there are several processes that should be considered in order to accurately follow gas cooling and fragmentation in primordial conditions, processes not typically included in numerical simulations (to my knowledge, the Enzo chemical network does not include such processes). For instance, apart from the missing chemical species and reactions in the network relative to Glover 2015, the densities in the current simulation are so high that the transition to the LTE cooling rates should have been accounted for, but I did not find any mention of it in the draft. I understand that the simulations are already very expensive, and I am not asking to improve the network and rerun them, but the authors should at least discuss potential caveats of their work in the light of these missing processes.

Good point. Glover 2015, MNRAS, 451, 2082 (a) and Glover 2015, MNRAS, 453, 3 (b) used a reaction-based reduction technique to construct the minimal model of 26 reactions required to assess the effect of the choice of chemical model on J_{crit} , the minimum Lyman-Werner (LW) UV flux required to destroy H₂ in a halo. Glover 2015a found that Enzo's native reaction network yielded results that were within a factor of two of their own network (section 3.6.2 of Glover 2015a states that "the Enzo code consists of the same subset of the reactions from our minimal model"). Glover 2015b found that the chemical network in the most recent version of Enzo omits two reactions that can affect J_{crit} : collisional ionizations of H by H (reaction 10) and the dissociative tunnelling contribution to the H₂ collisional dissociation rate. However, there are no LW fluxes in our simulation and the temperatures in the halo are too low to produce many collisional ionizations of H by H so these omissions have little impact on our simulation. We now cite both papers and note the omitted processes in the paper.

Our H₂ cooling rates (Glover & Abel 2008, MNRAS, 388, 1627) are valid at low densities and high densities, at which LTE effects become important. They also include the transition from optically-thin to optically-thick gas cooling at densities at which H₂ becomes opaque to its own cooling lines (Ripamonte and Abel 2004, MNRAS, 348, 1019). We now clarify these points in the letter.

>> 2) The authors employ a pressure floor in their simulations for cells at the maximum refinement level, in order to force the Jeans length to be resolved. I am concerned that this choice, which the authors justify with the need to stabilise the collapse, might significantly affect their results. In particular, while it is true that the pressure floor helps avoiding numerical

>> fragmentation, this is obtained at the expense of energy
>> conservation, actually injecting energy into the system. This
>> energy/pressure increase might artificially support gas against
>> collapse and inhibit *physical* fragmentation and subsequent star
>> formation that should instead occur, until the mass in the centre
>> becomes extremely large. In particular, several recent works on
>> star formation in molecular clouds, with similar resolutions to
>> the current study, have opted for not including a pressure floor,
>> but instead introducing sink particles in maximally refined cells
>> where the Jeans length was not properly resolved. Why did the
>> authors not employ this approach, especially given the extremely
>> high resolution achieved?

Now addressed in the general comments. That said, to follow up on your concerns we ran an additional sanity check by turning on sink particles during all of stage 1 and for the first Myr of stage 2 instead of pressure floors. Sinks were created if five criteria were met:

- (1) the grid cell is at the maximum refinement level
- (2) the gas density in the cell is higher than the Jeans density
- (3) the flow is convergent
- (4) the cooling time is shorter than the free-fall time
- (5) the cell is at a local minimum of the gravitational potential

Besides accreting gas, these sinks, once created, also act as star particles so they can be a source of ionizing UV radiation, winds and SN explosions. This algorithm has been tested and used in a number of studies such as Regan & Downes 2018, MNRAS, 478, 5037, Regan et al. 2020, OJA, 3, 14, Latif et al. 2021, MNRAS, 508, 1756 and Latif et al. 2021, arXiv:2109.10655 and is similar to those used in Federrath et al 2010, ApJ, 713, 269, and Krumholz 2004, ApJ, 611, 399.

No sink particles formed during either stage of the run, confirming that no normal Pop III stars formed before the two supermassive stars did. This was consistent with Latif et al. 2013, MNRAS, 436, 2989, who found essentially no differences when using sink particles or pressure floors in simulations of massive BH seed formation. We now discuss these points in the text.

>> 3) Moreover, when the authors restarted the simulation at a lower
>> level of refinement (0.01 pc), I guess the pressure floor was
>> adjusted accordingly. I wonder whether this instantaneous change
>> can potentially affect the entire halo core, introducing a bias
>> in the subsequent evolution, especially in the light of the lower
>> resolution (which corresponds to not being able to resolve the
>> Jeans length on larger scales).

Good question. When the pressure floor was activated in the second half of the simulation at the 18th level of refinement it would have been a little different from the floor turned on at the 25th level at the end of the first stage of the run.

But any differences in internal gas energies deposited at those scales was immediately dwarfed by the subsequent collapse of the halo, when supersonic turbulence in the much higher densities drove the Jeans mass to over 10^4 Msun, effectively cutting off normal Pop III SF at small scales until later times. This is why the two massive clumps appeared and formed supermassive stars first. We ensured that the Jeans length was always resolved by at least 64 cells.

>> i) Abstract (references 5/6): Given the achievable resolution >> and the limitations of the modelling (especially in the paper >> by Q. Zhu), I find this claim a bit too strong, and a lot of >> effort is still needed in my opinion.

We now state that supermassive stars are "probably" the seeds of the first quasars.

>> ii) Abstract (references 9/10/11): I disagree with this >> sentence. Recently, Lupi et al. 2021 showed that the conditions >> for the formation of a massive seed are actually favoured in >> quasar progenitors, and that the only antagonist is metal >> pollution resulting from the high sfr in the region.

Our understanding of Lupi et al. 2021, MNRAS, 503, 5046 is that only one pair of synchronized halos in which DCBHs could likely form were among the progenitors of the quasar host halo, although there were also a large number of dynamically heated halos. But because these halos were clustered they in reality may have been polluted by metals due to SF in unresolved halos because the DM resolution in Lupi et al was limited to $\sim 10^7$ Msun. Metals would make DCBH formation in the candidate halos more difficult. We spoke with Lupi, who confirmed these points and noted that finding even a single halo that meets all the conditions for DCBH formation in their study (LW flux, tidal torques, photo-evaporation and metal enrichment) might be difficult, but this remains to be verified in future work. We now cite Lupi et al. in the paper. Because supermassive stars were not actually demonstrated to form in Lupi et al, we feel that our statement in the abstract is correct.

>> iii) Abstract: "without the need..., or even atomic cooling". >> What the authors mean here? In order to collapse, gas needs to >> cool down, and in primordial environments cooling is provided >> by atomic cooling down to 10^4 K, and molecular cooling below >> 10^4 . Are the authors claiming that cooling is provided by H2 >> only, without resulting in fragmentation?

During collapse, the halo is mainly cooled by H2 but a few cells with a small total mass do reach 8000 - 10000 K and cool by Ly α emission. However, these few cells have far too little mass to affect the collapse of the halo as a whole, which is why average temperatures in the halo are only ~ 2000 K. As discussed above, fragmentation in the halo on smaller mass scales is suppressed by supersonic turbulence and the relative inefficiency of H2 cooling prior to the collapse of the halo and by large effective Jeans masses after the onset of collapse that are also due to the turbulence. Now clarified in the text.

>> iv) Main text: Why does the temperature never exceeds 2000 K? >> Is it because there is molecular gas cooling gas efficiently?

>> Or is the atomic phase simply not resolved in time because
>> when the gas heats up above 10^4 K atomic + molecular cooling
>> bring the temperature down very quickly? With respect to the
>> turbulence, although I agree with the authors that turbulence
>> can provide support against collapse, it is also true that
>> supersonic turbulence produces shocks that, in case of
>> efficient cooling, can enhance significantly the gas density
>> and favour the formation of small scale fragments that might
>> form stars before enough mass has reached the centre of their
>> massive halo.

Efficient H₂ cooling prevents temperatures in the halo from exceeding ~ 2000 K except for small pockets of gas that are cooled by Ly α emission as noted above. Your question as to whether Ly α cooling could produce snapshots of 2000 K average temperatures is a good one. In our past experience with halos in a variety of LW backgrounds, atomic cooling is strong but cannot reduce gas temperatures below the Balmer thermostat at 6000 - 8000 K. H₂ cooling rates peak at 2000 - 3000 K, and at these temperatures can be two orders of magnitude higher than in normal Pop III star halos, which is why H₂ can keep the gas at lower temperatures even in a halo of this mass (also, there is no LW background here to reduce H₂ mass fractions). As a sanity check, we switched off cooling in a test simulation and temperatures rose to 10^4 K during collapse, ultimately halting it. As noted above, we'd expect turbulent shocks to fragment gas in the presence of metal and dust cooling, but H₂ is a much less efficient coolant, which is why density fluctuations due to supersonic turbulent shocks do not collapse to stars in our model.

>> v) Main text: About the massive self-gravitating disc, how is
>> this related to the pressure floor? I am wondering whether the
>> same disc would have formed without pressure floor. Later
>> during collapse, the authors say that H₂ formation is boosted
>> via three-body reactions, and additional cooling brings the
>> temperature to drop. I am wondering what is the impact of the
>> pressure floor on the fragmentation that might appear because
>> of this strong cooling. This is important to understand why a
>> massive gas distribution of 4×10^6 can be present in the halo
>> without any star having formed so far.

Good question. The small disk forms when the maximum resolution of the simulation is 20 AU. Any fragmentation suppressed by the pressure floor would occur at or below these scales, so the floor has no effect on the formation of the disk because it is much larger, several thousand AU. The combination of highly supersonic turbulence driven by the cold flows and the relative inefficiency of H₂ cooling is what prevents SF in the gas, as noted above. If cooling would lead to fragmentation on scales below our resolution limit we'd expect the clumps to migrate inward onto the central object on timescales that are short in comparison to the object's contraction onto main sequence burning, as found in other studies cited above. However, we note that the clump with this disk becomes part of C2 before it can form a star.

>> vi) Main text: The authors claim that C1 and C2 collapse before
>> any star has formed in the halo. I do not find anywhere in the
>> draft how star formation is modelled in the simulation. Pop III
>> stars are also expected to live for a few Myr (assuming a
>> top-heavy IMF), hence metal enrichment via SN feedback might

>> alter the evolution. No details on these processes are given in
>> the manuscript, and are crucial to assess how reliable the
>> presented simulations are. On top of this, even if star
>> formation has been included, it is not clear how the presence
>> of a pressure floor alters the formation of stars in the halo.

Now explained in the general comments, but let us know if we can
clarify anything in more detail.

>> vii) Supplementary material (Halo identification): What do the
>> authors mean with "Lagrangian volume of the halo"? How many
>> virial radii does their volume enclose?

The Lagrangian volume of the halo is the region that contains
all the particles that later become part of the halo by $z = 6$.
Our most refined region encloses twice the Lagrangian radius
of the halo.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I am happy that the authors have adequately addressed my concerns in the revised manuscript, and would recommend it for publication. Please find below my slightly adjusted summary of the key results:

How the first supermassive black holes, which are observed to power quasars at $z=7$, grew so massive in such a limited period of time has remained a scientific puzzle since their discovery. The problem can be split into two separate questions: a) how and when did the seed black holes, which later grew into supermassive black holes, first form? b) Under which circumstances can these seed black holes grow into supermassive black holes by their observed redshift? The latter part of the question was recently answered by DiMatteo and her research group (Tenneti2018,Huang2019), who showed that low-shear environments, which host massive halos that are fed predominantly by accretion flows, are able to efficiently grow black holes from 10^5 $M_{\odot}$ to more than 10^8 $M_{\odot}$. However, the first part of the question (where the 10^5 $M_{\odot}$ black hole came from in the first place) remains open and is the one addressed by the authors of this paper.

In this paper, the authors present high resolution simulation of one halo that fulfills the criteria highlighted by Tenneti2018, and studies in detail the formation of seed black holes in such halos. The work is therefore directly complementary to existing work in the field, and addresses one of the key open questions in our current understanding of the origin of the first observed quasars. The authors report that high levels of turbulence suppress early star formation. This instead allows gas to collapse directly into dense clumps which are projected to form black holes in the mass range 10^4 - 10^5 $M_{\odot}$.

The simulations presented here are not unique in studying the formation of such supermassive stars, and their potential to become seed black holes. As acknowledged by the authors in the paper, such objects have been studied in various environments before. The work presented here stands out in comparison to previous work in its choice of environment: it explicitly focuses on the same type of environment that has been shown to be able to grow such black holes into observed quasars by Tenneti2018. It shows that such environments can form massive seed black holes via a supermassive star, solving the puzzle where the black holes grown in Tenneti2018 first came from. This is a key scientific insight that addresses one of the main open questions in our understanding of the origin of the first supermassive black holes. With the information added in response to the first round of reviews, the methodology is clearly explained and appropriate for the problem. Information is clearly communicated.

Referee #2 (Remarks to the Author):

I thank the authors for their answers, that helped clarifying many aspects of their study.

The manuscript is now much clearer. However, I still have some concerns, that I would like the authors to address before being ready for publication.

1) I really appreciated the authors' effort in rerunning the simulation using sink particles. I was wondering whether the gas distribution is somewhat different in this case, despite not changing the final result. In this respect, I think it would be useful for the reader to see a surface density map similar to Fig. 2 in the extended material.

2) About sink particles, there are a few aspects that should be still clarified.

— First, I am wondering whether the authors have assessed the main condition that prevents PopIII star formation from occurring. Their sink formation scheme is based on 5 criteria, hence it would be important to understand which plays the most important role and how it is affected by the physical conditions of the halo, but also by potential numerical limitations of the study. For instance, I may naively guess that the first and the fifth conditions are more easily matched than the other three conditions. In particular, I am wondering whether condition 4 is always matched, given the almost constantly very low gas temperature (which might hint for the cooling time being very short compared to the simulation time-step), or if instead such low temperatures result in the cooling time becoming artificially longer (a condition that easily occurs when we approach the equilibrium temperature of the gas). In this respect, many other studies did not consider this condition for sink formation (see Federrath et al., and even Krumholz et al.)

— Secondly, I was wondering how the Jeans density is estimated? Is the approach the same as in Regan & Downes (2018), based on Krumholz et al. (2004), which uses the sound speed only? Or does it also include turbulence support? Moreover, is the number of Jeans factor (64) included in the calculation? These details should be mentioned, since they would have crucial implications for the results in the manuscript.

2) After re-reading the manuscript carefully, I also noticed that there is no discussion in the text of the mass growth history of the halo, in particular in relation to the previous results by Wise et al. 2019 and Regan et al. 2020, that the authors mention also in their first answer. I know that Wise et al. found that repeated mergers were responsible for dynamical heating, but dynamical heating by itself (as discussed by Yoshida et al. 2003) can also result from smooth accretion, as long as the mass accretion rate is large enough. This is also discussed in Lupi et al. (2021). I think it would be useful to assess whether the peculiar conditions of the simulated halo in this work match those of the dynamical heating scenario discussed in Yoshida et al. 2013 and Wise et al. 2019 or if they represent a completely different mechanism.

3) About my third comment from the previous report, I disagree with the authors about it being “a little different” when the resolution is changed by 7 levels of refinement. In particular, at fixed density and temperature, the jeans length would be resolved with $1/2^7$ of the previous number of cells, i.e. less than 1 cell. This means that the pressure to correct for this ($\sim c_s^2$) would become 2^{14} times larger. On a different side, but still related to the pressure floor, I disagree with the authors

about the pressure floor playing a role in the unresolved fragmentation in the disc at 20 AU or below. Since their pressure floor ensures the Jeans length is resolved with 64 cells, the fragmentation is more likely suppressed at or below $20 \times 64 \text{ AU} \sim 1280 \text{ AU} \sim 0.01 \text{ pc}$ in Stage 1, and at or below $0.01 \times 64 \text{ pc} \sim 0.64 \text{ pc}$ in Stage 2. In any case, considering that the same results have been found with sink formation, I think this comment can be considered minor.

4) About Lupi et al. (2021), after re-reading their paper, I can state that the mass resolution in the zoom-in region is 10^5 Msun and not 10^7 Msun . Nonetheless, I agree with the authors that this means that some minihaloes in their early formation stage might be unresolved, and this could potentially affect the number of candidates. However, I honestly think that, although Lupi et al. (2021) did not actually demonstrate that a SMS would form in the halo, it still represents a first attempt at addressing the conditions for SMS formation in quasar host haloes at high redshift, and proper credit should be given. In particular, the simulation here is a zoom-in as that in Lupi et al. (2021), although with a higher resolution and including hydrodynamics. Therefore, the region outside the Lagrangian volume simulated might be affected as well by unresolved star formation and feedback, that might potentially enrich the candidate halo in which the SMSs form, invalidating the current results.

5) About the chemistry network, it is true that the Enzo network performed reasonably well in Glover (2015). Nonetheless, the authors are not considering that Glover explored the differences between his network and Enzo in the value of J_{crit} to avoid H_2 cooling, and there is no mention in that work of the impact of the missing reactions in $J=0$ conditions. In particular, it is not obvious the extrapolation of conclusions in Glover (2015) is reasonable. I therefore suggest the authors to add a caveat about this.

Author Rebuttals to First Revision:

>> 1) I really appreciated the authors' effort in rerunning the
>> simulation using sink particles. I was wondering whether the
>> gas distribution is somewhat different in this case, despite
>> not changing the final result. In this respect, I think it
>> would be useful for the reader to see a surface density map
>> similar to Fig. 2 in the extended material.

We think you may have meant Fig. 4 because Fig. 2 was made at a much higher resolution of 20 AU. We put a new Fig. 4 for the sink particle run in the Methods section and, as expected, the structure of the core is very similar to that in the pressure floor run.

>> 2) About sink particles, there are a few aspects that should
>> be still clarified. First, I am wondering whether the authors
>> have assessed the main condition that prevents Pop III star
>> formation from occurring. Their sink formation scheme is
>> based on 5 criteria, hence it would be important to understand
>> which plays the most important role and how it is affected
>> by the physical conditions of the halo, but also by potential
>> numerical limitations of the study. For instance, I may
>> naively guess that the first and the fifth conditions are
>> more easily matched than the other three conditions. In
>> particular, I am wondering whether condition 4 is always
>> matched, given the almost constantly very low gas temperature
>> (which might hint for the cooling time being very short
>> compared to the simulation time-step), or if instead such low
>> temperatures result in the cooling time becoming artificially
>> longer (a condition that easily occurs when we approach the
>> equilibrium temperature of the gas). In this respect, many
>> other studies did not consider this condition for sink
>> formation (see Federrath et al., and even Krumholz et al.)

This is an interesting question, but the code does not record which criteria for SF are violated in the cells that are at the maximum level of refinement during the run (to our knowledge, no code does). In principle, one could hardwire Enzo to deactivate the criteria one by one, rerunning the simulation each time to determine the effects of each one, but this would require a computational study several times larger than the one we've already performed and is beyond the scope of the letter. The state-of-the-art sink particles we use are well tested and produce results similar to those in the studies mentioned by the referee (Krumholz 2004 and Federrath 2021) as well as in Bleuler & Teyssier 2014, Regan et al. 2018,2019,2020 and Latif 2021a, 2021b so we are confident of its results.

That said, we suspect that normal Pop III SF is suppressed because conditions II and III are generally not met. The highly supersonic turbulence disrupts convergent flows, in turn preventing cell densities from reaching the Jeans density. We examined the cooling and free fall times in our simulation and found that below a few pc the cooling time is always shorter than the free fall time, as shown in the plot attached below, so condition IV should not be inhibiting SF in our runs. Turbulence delays collapse in our runs so it probably extends free fall times.

>> Secondly, I was wondering how the Jeans density is estimated?
>> Is the approach the same as in Regan & Downes (2018), based
>> on Krumholz et al. (2004), which uses the sound speed only?
>> Or does it also include turbulence support? Moreover, is the
>> number of Jeans factor (64) included in the calculation?
>> These details should be mentioned, since they would have
>> crucial implications for the results in the manuscript.

We estimate the Jeans density as was done in Krumholz et al 2004 and Regan & Downes 2018: $\rho_{\text{jeans}} \sim 1/4 * \pi * c_s^2 / (G \Delta x^2)$. To be consistent with the state of the start in the field we did not include turbulent support, something we will consider in future work to formulate a recipe for SMS formation. We use the Truelove criterion, a factor of 4 instead of 64.

>> 2) After re-reading the manuscript carefully, I also noticed
>> that there is no discussion in the text of the mass growth
>> history of the halo, in particular in relation to the
>> previous results by Wise et al. 2019 and Regan et al. 2020,
>> that the authors mention also in their first answer. I know
>> that Wise et al. found that repeated mergers were responsible
>> for dynamical heating, but dynamical heating by itself (as
>> discussed by Yoshida et al. 2003) can also result from smooth
>> accretion, as long as the mass accretion rate is large enough.
>> This is also discussed in Lupi et al. (2021). I think it would
>> be useful to assess whether the peculiar conditions of the
>> simulated halo in this work match those of the dynamical
>> heating scenario discussed in Yoshida et al. 2013 and Wise et
>> al. 2019 or if they represent a completely different mechanism.

We discussed the mass growth of our halo in paragraph 2 of the Methods section and showed it in two plots in Extended Data Figure 1. There are a number of differences between the environment and assembly history of our halo and those in Wise et al. 2019 (W19; note that Regan et al. 2020 studied the same halo). For example, the halo MMH in W19 reaches $2e7 M_{\text{sun}}$ by $z \sim 16.5$, having grown by a factor of 30 in mass in just 30 Myr from $z = 21 - 19$ because of a major merger. The DM resolution

in W19 is $\sim 3e4$ Msun, an order of magnitude lower than ours, so they barely resolved the halo when it was below 10^6 Msun. It was exposed to a LW flux of 3 J₂₁, which results in a halo mass of $\sim 2e6$ Msun at collapse (see Fig 2 of W19).

Does dynamical heating delay collapse in our halo? As explained in the text, our halo forms about 100 Myr earlier and grows by about two orders of magnitude in 40 Myr primarily by accretion. No major mergers contribute to its growth. Although the average growth rate of our halo is larger than in W19, below $3e6$ Msun its instantaneous growth rate is a few 10^5 Msun per unit redshift, an order of magnitude smaller than in W19. The max growth rate of our halo is still a factor of 2 - 3 times lower than in W19 and peaks when the halo has already reached 10^7 Msun.

Typical H₂ mass fractions in W19 are 10^{-5} - 10^{-6} , at least an order of magnitude lower than those in our halo, $\sim 10^{-4}$. At these H₂ fractions the growth rate of our halo would have to exceed that in W19 by an order of magnitude for dynamical heating to be effective (see their Extended Data Figure 1). Since the growth rate of our halo is about 20 times lower than the critical rate for dynamical heating at the H₂ fractions in our halo, it has little or no impact on its collapse.

Note that turbulent and radial infall velocities in W19 are about a factor of ten lower than in our halo (~ 30 - 40 km/s) so turbulence has little effect on the collapse of their halo. The conditions leading to the formation of SMSs in our halo are thus quite distinct from those in W19. Now discussed in the Methods section.

>> 3) About my third comment from the previous report, I disagree with the authors about it being 'a little different' when the resolution is changed by 7 levels of refinement. In particular, at fixed density and temperature, the jeans length would be resolved with $1/2^7$ of the previous number of cells, i.e. less than 1 cell. This means that the pressure to correct for this ($\sim c_s^2$) would become 2^{14} times larger. On a different side, but still related to the pressure floor, I disagree with the authors about the pressure floor playing a role in the unresolved fragmentation in the disc at 20 AU or below. Since their pressure floor ensures the Jeans length is resolved with 64 cells, the fragmentation is more likely suppressed at or below 20×64 AU ~ 1280 AU ~ 0.01 pc in Stage 1, and at or below 0.01×64 pc ~ 0.64 pc in Stage 2. In any case, considering that the same results have been found with sink formation, I think this comment can be considered minor.

Machacek et al 2001 found that any effect of the pressure floor

implemented in Enzo should not extend beyond a distance equal to the length of 4 cells at the maximum refinement level. We therefore take our results to hold above 80 AU and 0.04 pc in the first and second stages of the simulation, respectively and now make this point in the draft. We agree with the referee, it's a minor issue that we can now consider to be closed.

>> 4) About Lupi et al. (2021), after re-reading their paper, I
>> can state that the mass resolution in the zoom-in region is
>> 10^5 Msun and not 10^7 Msun. Nonetheless, I agree with the
>> authors that this means that some minihaloes in their early
>> formation stage might be unresolved, and this could
>> potentially affect the number of candidates. However, I
>> honestly think that, although Lupi et al. (2021) did not
>> actually demonstrate that a SMS would form in the halo, it
>> still represents a first attempt at addressing the conditions
>> for SMS formation in quasar host haloes at high redshift, and
>> proper credit should be given. In particular, the simulation
>> here is a zoom-in as that in Lupi et al. (2021), although with
>> a higher resolution and including hydrodynamics. Therefore,
>> the region outside the Lagrangian volume simulated might be
>> affected as well by unresolved star formation and feedback,
>> that might potentially enrich the candidate halo in which the
>> SMSs form, invalidating the current results.

We now cite Lupi et al. 2021 in the summary paragraph (abstract). The halo in our run is the first to collapse not just in twice its Lagrangian volume (which is fully resolved) but in the entire $(37.5 \text{ Mpc})^3$ box. It is therefore highly unlikely that other stars will contaminate the halo with metals before the two SMSs appear, especially since the halo forms at $z = 35$, well before the vast majority of Pop III stars form.

>> 5) About the chemistry network, it is true that the Enzo
>> network performed reasonably well in Glover (2015).
>> Nonetheless, the authors are not considering that Glover
>> explored the differences between his network and Enzo in
>> the value of J_{crit} to avoid H₂ cooling, and there is no
>> mention in that work of the impact of the missing reactions
>> in $J=0$ conditions. In particular, it is not obvious the
>> extrapolation of conclusions in Glover (2015) is reasonable.
>> I therefore suggest the authors to add a caveat about this.

Recall that we noted in the text that neither reaction (the dissociative tunnelling contribution to the H₂ collisional dissociation rate nor the collisional ionization of H by H) are included in our simulations. Although Glover 2015a, 2015b did not consider the effect of these reactions in the case of no LW background, it is clear from the two reaction rates that their impact would be negligible. The H-H collisional

ionization rate is

$$4.5e-21*(T)^{(1.5)}*\exp(-157800.0/T) \text{ (ref: Glover 15)}$$

At the temperatures in our simulation this reaction rate is $\sim 10^{-50} \text{ cm}^3/\text{s}$. It would have been relevant had T been $\sim 8000 \text{ K}$ because the reaction rate increases by 20 orders of magnitude. Collisional dissociation of H_2 by H atoms is strongest at $2000 - 4500 \text{ K}$ (see Fig. 13 in Martin et al 1996) and was considered in Jcrit studies by Latif et al. 2014,2015 before Glover 2015a,2015b. Again, our runs never reach these temperatures so these reactions would have little impact on the collapse of the halo. We now discuss the temperature ranges in the text to make it clear why the two reactions would not change our results.

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #2 (Remarks to the Author):

I thank the authors for their response and the clarifications. However, I am not fully convinced by some of the authors' comments.

1) About the new map, it seems that the colormap used is slightly different from the one in the main text (the colours are less bright). It's interesting to notice that the clumps in the map are less and also less dense (and still being a bit more compact) than in the pressure floor case. These differences might hint to a lower mass. Is that the case or is it due to a different limit in the colorbar/choice of the colormap? Have the authors measured the clump properties in this case?

On top of this, I found very difficult to compare the results, since the maps correspond to different times. I think reporting the maps at the same times of Fig. 4 would be much more useful and would also support their claim about the pressure floor not having a significant effect.

2) About the second point, I am fine with the authors comment. Nonetheless, I think one could simply write in a log file the first criterion that is not met every time the sink formation routine is called. This would avoid running many simulations and, at the same time, would allow to record a relevant piece of information (this is a suggestion for future works).

2b) The jeans density in Krumholz+04 scales with J^2 , i.e. $1/N_J^2$. According to the Truelove criterion, the factor should then be $1/16$, not $1/4$. Is it a typo? Or did the author actually use $1/4$ (corresponding to $N_J=2$)? In the latter case, their threshold is a factor of 4 higher than the Truelove one (which is already different from the choice they made in the pressure floor case, hence possibly biasing the comparison).

3) I reread the paper by Machacek+01, but I did not find any mention of the 4 cells in the text. Actually, the factor $K=100$ entering the pressure floor in Machacek+01 corresponds to $N_J \sim 10$ (assuming $\mu=1.22$), much lower than the value in this work. This choice resulted in fragmentation being suppressed below 1-2 pc (as Machacek wrote in the paper), when the resolution was $dx \sim 0.05$ - 0.03 pc (at $z=19$ and 30 , respectively). By comparing these two numbers the authors can immediately notice that $1\text{-}2\text{ pc} > \sim 10 * dx$, i.e. not 4 cells but more than N_J . In this respect, I would conclude that the choice of $N_J=64$ is expected to suppress fragmentation on scales comparable to $N_J * dx$, contrary to what the authors claim.

4) About the chemical network, I think the authors misunderstood my point. I think the reasonable agreement between Glover+15 and the Enzo chemical network has been demonstrated only in the $J>0$ case by Glover, while there is no confirmation that the $J=0$ would lead to the same results (and not because of the two missing reactions, but because of the other differences in the network). The comment they added to the text is therefore useless (it is a repetition of the sentence just above the added one). My request of adding a caveat was referring to a general comment about the possible uncertainties due to the chemical network employed, nothing more than that.

Author Rebuttals to Second Revision:

>> 1) About the new map, it seems that the colormap used is slightly
>> different from the one in the main text (the colours are less
>> bright). It's interesting to notice that the clumps in the map are
>> less and also less dense (and still being a bit more compact) than
>> in the pressure floor case. These differences might hint to a lower
>> mass. Is that the case or is it due to a different limit in the
>> colorbar/choice of the colormap? Have the authors measured the
>> clump properties in this case? On top of this, I found very
>> difficult to compare the results, since the maps correspond to
>> different times. I think reporting the maps at the same times of
>> Fig. 4 would be much more useful and would also support their
>> claim about the pressure floor not having a significant effect.

The two figures had the same colorbar range and color maps but EDF 3 had to be converted from PS to JPEG format because of formatting rules with Nature. This is likely why the shades looked a little different, but the identical colorbar ranges nevertheless enable good comparisons of the densities. We changed EDF 3 to include two panels at 0.6 Myr and 1.0 Myr. The structures in the core at those times are very similar to those in Figure 4. Minor differences are expected between two runs because of the stochastic nature of turbulence and the non-linear dynamics of gravitational collapse. As we show in the new EDF 4, the masses of C1 and C2 in the two simulations are also quite similar and were the first clumps to collapse in both runs (and with very similar central infall rates). We updated the text accordingly.

>> 2b) The jeans density in Krumholz+04 scales with J^2 , i.e $1/N_J^2$.
>> According to the Truelove criterion, the factor should then be
>> $1/16$, not $1/4$. Is it a typo? Or did the author actually use $1/4$
>> (corresponding to $N_J=2$)? In the latter case, their threshold is
>> a factor of 4 higher than the Truelove one (which is already
>> different from the choice they made in the pressure floor case,
>> hence possibly biasing the comparison).

We thank the referee for pointing this out, there was a typo in our previous reply to the referee report ($\rho_{\text{jeans}} \sim 1/16 \cdot \pi \cdot c_s^2 / (G \cdot \Delta x^2)$). We resolve the Jeans density by 16 cells, as in Krumholz et al. (2004).

>> 3) I reread the paper by Machacek+01, but I did not find any
>> mention of the 4 cells in the text. Actually, the factor $K=100$
>> entering the pressure floor in Machacek+01 corresponds to $N_J \sim 10$
>> (assuming $\mu=1.22$), much lower than the value in this work. This
>> choice resulted in fragmentation being suppressed below 1-2 pc
>> (as Machacek wrote in the paper), when the resolution was $\Delta x \sim 0.05$
>> -0.03 pc (at $z=19$ and 30, respectively). By comparing these two
>> numbers the authors can immediately notice that 1-2 pc $\gg 10 \cdot \Delta x$,
>> i.e. not 4 cells but more than N_J . In this respect, I would
>> conclude that the choice of $N_J=64$ is expected to suppress
>> fragmentation on scales comparable to $N_J \cdot \Delta x$, contrary to what
>> the authors claim.

A few clarifications: first, the pressure floor is imposed on every cell at the deepest level of refinement. It is constructed by balancing the thermal and gravitational potential energy of the gas to ensure that when it is imposed on a cell that the new Jeans length of the cell is $\sqrt{K} \cdot \Delta x$, regardless of by how many cells with which the Jeans length is resolved in the run.

Consequently, matter in that cell *may* later be dispersed in a sphere with a radius of at most $\sim 0.5 * \sqrt{K} * dx$. We have confirmed this to be true in a private communication with an Enzo code developer, Britton Smith. We now therefore state at the bottom of p. 13 in the draft that

"The pressure floors we imposed on the highest level of refinement at the end of stage 1 and during stage 2 suppress fragmentation on scales of at most 10 times their resolution limits (200 AU and 0.1 pc, respectively..."

That said, while the pressure floor in principle can suppress fragmentation below these scales it did not in our simulation because no such fragmentation occurred in the sink particle run, in which there were no floors. C1 and C2 were the first fragments to form stars in both runs.

>> 4) About the chemical network, I think the authors misunderstood
>> my point. I think the reasonable agreement between Glover+15 and
>> the Enzo chemical network has been demonstrated only in the $J>0$
>> case by Glover, while there is no confirmation that the $J=0$ would
>> lead to the same results (and not because of the two missing
>> reactions, but because of the other differences in the network).
>> The comment they added to the text is therefore useless (it is a
>> repetition of the sentence just above the added one). My request
>> of adding a caveat was referring to a general comment about the
>> possible uncertainties due to the chemical network employed,
>> nothing more than that.

Thanks for clarifying your request, we now state

"However, since there are no LW fluxes in our simulation and the temperatures at which these two reactions become significant are 8000 K and 2000 - 4500 K, respectively (which are greater than those in our halo), their omission has little impact on our results. In principle, the choice of chemistry network, reaction coefficients and update scheme can lead to some variations in the thermal evolution of the gas."

Reviewer Reports on the Third Revision:

Referees' comments:

Referee #2 (Remarks to the Author):

I thank the authors for addressing all my comments. The results are now more convincing, hence I think the manuscript is worth publication.

Author Rebuttals to Third Revision:

>> Referees' comments:

>> Referee #2 (Remarks to the Author):

>> I thank the authors for addressing all my comments. The results are now more
>> convincing, hence I think the manuscript is worth publication.

Thanks for signing off on the final version of the letter,

DW