Response to referee report for manuscript ref. MN-14-0750-MJ

```
The authors perform sophisticated galaxy formation simulations in a
fully cosmological context. Using explicit radiative transfer
calculations, the authors can derive, rather than assume, photon
escape fractions of the simulated galaxies and hence consistent
reionization simulations are performed. The authors conclude
low-mass galaxies actually dominate the ionizing photon budget in
the early phase of reionization. The simulations are of high quality
and the conclusions are interesting. The authors have also checked
caveats and done reasonable comparisons with previous works. I
recommend the paper for publication in MN, subject to revisions as
suggested in the below. I request mostly clarifications of a number
of points in the text.
```

We thank the referee for his/her review of our work, and we have addressed the points individually below. In the manuscript, we have boldfaced the added text and crossed out the deleted text. In addition to these changes below, we have cited Kimm & Cen (2014) and Salvadori et al. (2014) in Section 5.3.2 and 5.3.3, respectively, during the revision process.

```
1 Introduction, 1st para.
"Lyman alpha forest lines imply that"
Please describe a little more precisely.
The authors probably mean the complete Gunn-Peterson
trough in the spectra of QSOs beyond z=6, and they do
not really mean the IGM temperature inferred from Lyman-a
lines, is this right ?
```

We were referring to the Gunn-Peterson trough, not the IGM temperature in Lyman-a lines. We have changed that sentence to read,

"First, the transmission fraction of $z \sim 6$ quasar light blueward of Lyα through the intergalactic medium (IGM) indicates that the universe was mostly ionised by this epoch."

```
2 Introduction 2nd para
"which both are too rare"
-> "both of which are too rare" ??
```

Fixed.

```
3. Intro, the end of 3rd para.
"but starting at which limiting magnitude?"
Sounds like a conversation.
```

This is in the first author's writing style, and we would prefer to keep this question as originally stated because it leads into the next paragraph, which answers this exact question.

```
4. Intro. in the middle of 4th para.
"... Lyman-Werner (LW) background".
I find the authors use this term a little carelessly
```

```
throughout the manuscript.
With "background", I imagine some uniform cosmic
(large-scale) background, but with
"Lyman-Werner radiation", I think of some external
(nearby) radiation source as well.
In some parts, the authors mean "LW radiation
(including from near-by sources) suppress H2 formation"
by writing "by LW background".
A related point is the treatment of the global LW background
later in section 2.2. Shouldn't the authors include the global
background in addition to the collective LW photons
from local sources in the simulation box ?
```

We treat the LW radiation field as the sum of a time-dependent uniform background and optically-thin fields from point sources in the simulation. This is discussed at the end of Section 2.2. We agree with the referee that we should have been more specific about the terminology concerning soft UV radiation. We have replaced all relevant references to a soft UV or LW background with either "radiation field", "external radiation", or "intensity".

```
5. Somewhere in Introduction, perhaps around
6th-8th para, please write that the escape fraction
is intrinsically a quantity of an object, rather than
of a population. Namely, galaxies with the same mass
can have very much different fesc because of complex
structures, dust content etc. It is also very likely
time-dependent even within the active time of a galaxy.
I agree with the authors that it is important to discuss
the mean value of fesc as a function of mass, formation
epoch etc. Nevertheless, fesc is often misused or confused
in the literature, as if it were some global quantity.
```

We agree with the referee that this is important to stress to the readers. We have added the following sentences to the 6th paragraph,

"Before moving forward, it should be stressed that the UV escape fraction is an intrinsic quantity for a given galaxy not an entire population. Galaxies with the same mass can have very different escape fractions, arising from, e.g., complex gaseous and stellar morphologies, dust content, and cosmological mass inflow. Furthermore, variable star formation rates (SFRs) and the associated radiative feedback in the ISM can result in an escape fraction that is highly time-dependent."

```
6. Section 2.1
This is not really a comment in a referee report,
but I just wonder why the authors still use GRAFIC
when more useful ones are available such as MUSIC.
```

This is a fair question, and the answer isn't obvious to someone outside our collaboration. When we started this project in mid-2010, MUSIC was not fully tested, thus we used GRAFIC for the initial conditions. The simulations in the *Birth of a Galaxy* series all use the same initial conditions, which were calculated four years ago. As a note, all of our simulations started in the past two years have used MUSIC exclusively.

```
7. Section 2.2, about Population III IMF.
"a Salpeter IMF that exponentially ...."
I believe the Salpeter IMF is rather specific, and so
I'd suggest to write "a power-law IMF that exponentially...".
```

Changed.

```
8. Figure 2 is interesting in that the most massive halos
at z=8-9 actually cool by H/He rather than by metal
cooling. Is this because metal mixing is inefficient in
the large halos ? When I see Sutherland-Dopita's cooling
function, metal (Fe, Si etc) cooling is effective also at
T>10000K.
Can one think of some numerical artifact causing
inefficient mixing in the simulated galaxies ?
```

Most of the ISM has a temperature between 8,000 and 30,000 K, where the Sutherland and Dopita cooling curve does not depend on metallicity greatly, where the increased metal-line cooling occurs at higher temperatures. See the Figure for a density-temperature histogram of the gas inside the most massive halo at $z = 7.3$. Because cooling is so efficient at $T > 10^5$ K, there is very little gas mass to contribute to the total cooling rate of the halo, thus most of the cooling comes from hydrogen transitions.

Inefficient mixing does not appear to be a significant problem in our simulation. See the right panel of the Figure to see the average metallicity for a given density and temperature inside the most massive halo at z = 7.3. Here the metals are well mixed between all of the three ISM phases. There is some metal-free gas at low densities in the outskirts of the halo ($r \sim r_{\rm vir}$), which could be worth investigating in a later paper.

```
9. Halo mass function in Section 3.3.
This is the most inaccurate part of the paper from my
view point. To obtain the LF, the authors devise a correction
factor n_analytic/n_sim of haloes. If the correction is
actually small overall, I would not do any correction, because
the comparison of LFs shown in Figure 6 is only qualitative
anyway. If the correction factor is significant, I would not
do such rough correction either. In any case, I request the
authors to clarify if the main conclusions are affected
if they simply use the measured LF without the correction.
```

Fortunately, there was a bug in the Press-Schetcher (Sheth-Tormen) halo mass function code that we were using. We have fixed this bug, and the analytical halo mass function is consistent with the simulated halo mass function now. There is a slight deficiency of $\sim 5 \times 10^7 M_\odot$ halos in the simulation. There is an underproduction of halos with masses $M < 4 \times 10^5 M_\odot$ that is caused by limited mass resolution effects, which does not affect the luminosity function because galaxies do not form in such small halos. We have updated Figure 5 to reflect this change.

Figs/halo0-z.pdf

We have removed the correction factor to the luminosity function and updated Figure 6. The mean galaxy luminosity function at $M_{\rm UV} > -12$ has changed by a factor of two at $z > 10$, and we have updated the abstract and Table 2 with these uncorrected values.

```
10 Figure 8
Please explain the large jump at log M ~ 8.0 of the
solid line at z=7.3, if it has some physical meaning.
Or, in other words, what happened between z=8.0
and z=7.3 in such a short physical time ?
Maybe a diffuse UV background rapidly increases
its intensity ?
```

This large jump occurs when a single halo with $M \sim 10^8 M_\odot$ undergoes a strong star formation event with an escape fraction around 60 per cent at $z = 7.3$. We have added the following text to the end of the next to last paragraph in Section 3.4.2.

"The large jump at $M \sim 10^8 M_\odot$ is caused a single halo undergoing a strong star formation event with a sSFR $\simeq 30$ and $f_{\rm esc} \simeq 0.6$ that produces 70 per cent of the ionising photon budget in the simulation with the remaining 20 and 10 per cent originating from more massive and less massive haloes."

```
11. Section 3.5
It is unclear to me if the authors include intra-galactic
gas to calculate the clumping factor. Shouldn't one
include only clumping in the diffuse IGM ?
In other words, I couldn't get clear consistency between
clumping, f_esc and the photon production rate for
individual galaxies.
```

Thank you for catching this point. We indeed restrict the clumping factor calculation to the IGM, where we use the definition of 20 times the critical density as the IGM, which also excludes filaments. We have clarified by this oversight in a few places in Section 3.5 and the caption of Figure 11.

```
12. Occupation fraction f_host, equation (16)
The functional form looks like nothing but a function.
Please explain if eq (16) has some physical meaning
except the limit to f0.
```

This equation fits the simulated occupation fraction shown in Figure 3. With the exception to the sound-crossing time, the functional form has no physical basis except to fit the simulated data. We have added the following phrase after Equation 16 to clarify the equation's motivation.

"which provides a functional fit to the simulated values of $f_{\rm host}$ that are shown in Figure 3."

```
13. Section 5.2
The observability by JWST is mentioned but remain
unclear. Figure 15 does not seem to help our
understanding. Please clarify which part of Figure 15
is indeed within reach of JWST ? (Or TMT, ELT etc)
```

We have added the detection limits of the HUDF12 and Frontier Fields campaigns and JWST for a 10^5 second exposure to the figure. We have also added the "robust" sample of $z \geq 6.5$ galaxies from McLure et al. (2013) to the figure and made the associated changes to the text to reflect these additions.

```
This concludes my report.
```

We again thank the referee for the insightful review that helped improve our paper.