The authors thank Anonymous Referee #1 for his/her considered comments.

Changes to the text and detailed corrections regarding typos and such will be provided at the end of this round of review, except where they affect the meaning of the text. More substantive comments are addressed here. Note unless otherwise indicated, abbreviated versions of our responses below will be inserted into the revised text.

Reviewer's comments are provided in italics

Lack of explanations : The authors provide an explanation for the suppressed variability of the western side of the jet stream (though I personally think that one key aspect is missing – see comment below). However, the more complicated part of the story, explaining the jet characteristics over the eastern ocean basin, is not discussed at the same level of detail. I think this is a shame and I encourage the authors to do more analysis to improve that part of the story.

We agree with this reviewer that it is important to disentangle the various possible contributors to deglacial jet changes on the eastern side of the North Atlantic (Natl). However, this can't be done to the same level as the western jet with the existing set of ensemble runs. The potential contributors that we have considered to the eastern jet changes include:

- Stationary and transient eddy responses to ice sheet topography changes,
- Thermal effects on latitudinal surface temperature gradients in the North Atlantic (and possibly sea ice) due to ice sheet area and height (over North America and Fennoscandia), greenhouse gas concentrations and orbital forcing,
- Indirect effects associated with the pinning of the western side of the jet to a particular latitude, and
- Indirect effects associated with ocean circulation responses to the boundary condition changes.

We will add to the revised text the following conclusions that we can draw based on our existing runs.

1. The presence of elevated and extensive ice sheets provides the dominant controls on the jet characteristics on both sides of the NAtl jet. When the ice sheets are fixed to their LGM configurations in FixedGlac, little change is seen in the position or distribution of the jet on both its western and eastern sides. One notable exception to this is after 4ka BP, when the eastern side of the jet shows a reduced frequency at its preferred latitude and an expansion of its range. This timing is coincident with abrupt warming events in the FixedGlac experiments that are accompanied with abrupt retreats in Northern Hemisphere sea ice extent.
2. The eastern side of the NAtl jet is more sensitive to the background climate state than the western side of the jet is. This is evident when examining the differences between the FullyTrans and FixedOrbGHG experiments. By present day, the western side of the jet is mostly unaffected by the orbital and GHG components being fixed to LGM values. In contrast, the eastern side of the jet is centred approximately 5° further south when orbital and greenhouse gas components are fixed to LGM.
3. The position of the western side of the jet has limited control on the position of the eastern side of the jet. This conclusion is arrived at by considering the difference between western and eastern jet response to the FixedOrbGHG experiment discussed above and by examining the PDTopo experiment,. For the latter, the position of the western side of the jet is approximately 8° further north than in the FullyTrans experiments before 20ka BP. In these same experiments

at that time, the preferred eastern position of the NAtl jet is only shifted approximately 3° further north compared to its position in the FullyTrans.

4. The ocean state has an effect on the position of the jet on the eastern side of the NAtl. The differences in the eastern position of the jet between FullyTrans and PDTopo from 4ka BP onward indicate that although their forcings are the same or very similar, the history of the simulation affects the position of the jet on the eastern side of the NAtl. Such a difference is not apparent on the western side of the jet.

What is not clear from the simulations presented here is the role of Eurasian ice topography and extent (ie versus that of the NAIS) and the relative role of background temperature versus temperature gradient. Given the length and content of the current paper, these last questions will be answered in a future study.

Lack of dedicated discussion section:The authors have decided to jump straight from the results section to the conclusions without having a proper discussion section where your #ndings are put in perspective with the existing literature. This makes it hard to get a sense for how your results di#er from earlier studies and how they contribute to our understanding of the atmospheric circulation during the deglaciation. The discussion section is in my mind the most important part of a paper, so its omission feels instinctively wrong and may give this paper less traction than it deserves.

We will add on a subsection to the Results and Discussion section where we will more clearly summarize our results in the context of the existing literature and discuss implications.

Page 4, line 23: Sentence starting with "These changes.." is incorrect. To the best of my knowledge, none of the papers cited here suggest that the subtropical and midlatitude jets entered a merged state over the N Atlantic at the LGM; see, e.g., Fig. 1 in Li and Battisti (2008) where there is a clear separation between the subtropical and eddy driven jets. Page 4, lines 26-28: None of the papers cited here investigated that explicitly.

We see that the citation here was misleading in that it appears to attribute the identification of merged jets to those papers, rather than just the data from which we concluded the change in distribution of heat and moisture transports. We will remove the citation altogether and make it clear that these are the authors' inferences based on those papers.

Nevertheless, we would argue that the separation between the jets at LGM is not as clear in all cases as the reviewer suggests. For example, in Figure 3 of Li and Battisti (2008), there is little difference in the Atlantic zonal wind profiles of the latitude of the peak winds at 200 hPa and at the surface differing at LGM (indication of the positions of subtropical and eddy-driven jets in Eichelberger and Hartmann (2007)). In Figure 2 of Lofverstrom et al (2014), the 800hPa winds across the North Atlantic during LGM are highly zonal, and there is little evidence of any separation between the subtropical and eddy-driven jets. Finally, in Figure 4 of Merz et al (2015), there appears to be a clearer separation between the subtropical and eddy-driven jets at LGM than during Present Day, or a less merged jet state, much as in our study. Delineating whether a jet is merged or not is problematic, since it is based on the separation of the distributions of the subtropical and eddy-driven jets. How much separation can there be and still be considered merged? The only definition for a merged jet that I have encountered is in Harnick et al (2014), who define the Zonal Jet Index as the anomaly with respect to monthy or seasonal climatology of the maximum value of the zonal derivative of the latitude with peak zonal winds. They define a threshold for the jet to be merged as being a negative value in this derivative with time that exceeds one standard deviation. Perhaps it will be more helpful for us to replace the text in question with the following. “These changes are consistent with what would be expected if the North Atlantic jet approached a more “merged” state at LGM, although not all of these studies show evidence of this. Irregardless, changes to the path of the jet are expected to result in changes in the distributions of heat and precipitation over Western Europe during this time.”

Page 4, line 32: “The timing..” meaning here is not clear.

We state in the previous sentence that the jet transition detected by Lofverstrom and Lora (2017) occurred at the separation of the Cordilleran and Laurentide ice sheets at 13.89ka BP. A transition at 13.89 ka BP would be unlikely to explain either the abrupt warming into the B-A (occurring nearly a millennium earlier), nor the abrupt cooling at the start of the YD (occurring a millennium later).

Page 5, line 5 and section 2.1: Perhaps nit-pick but this not technically correct. PUMA (the dynamical core of PlaSim) is indeed a dry primitive equation model. However, the extra layer of physics on top of the dynamical core makes PlaSim more than a primitive equation model. More correct to say that it is a simplified general circulation model or, better yet, an Earth-system model of intermediate complexity (EMIC).

Most importantly, we would like to point out that PlaSim solves the moist primitive equations, not the dry. We agree with the referee that PlaSim is more than just a primitive equation model but so is any current generation Earth System Model. Our main goal in pointing out that the foundation of the atmospheric model is based on the moist primitive equations is to illustrate that the dynamics have not been simplified beyond what is common in many Earth System Models today. Rather, the simplifications in the atmospheric model arise mainly in the parametrizations included: no treatment of volcanic or anthropogenic aerosols, only a single greenhouse gas species explicitly accounted for, etc. We will revisit how we describe the model in this manuscript to make the above clear, but we intentionally avoided the term EMIC. EMIC has become a vague term encompassing a wide range of models with different combinations of sophisticated and simplified components.

Page 5, line 29: The description here is not correct. The Gaussian grid is the 128 x 64 cell grid in real space that the data is outputted on ("Gaussian" refers to how the grid is generated). The primitive equation are partially solved in spectral space (wave space) and are thus transformed between grid space (on the Gaussian grid, in this case 128 x 64 grid points in lon x lat) and the spectral representation in wave space, which supports at most 42 harmonics in the zonal and meridional direction, respectively. (Hence the name T42, where the T is short for "truncation" or more specifically "triangular truncation").

We agree with the referee and will correct the wording in the text accordingly.

Page 6, line 1: Not sure if I understand how the LSG models works. Is it a dynamic model that only runs in the mixed layer? If yes, how can a realistic ocean circulation be established if there is no deep ocean? Do you parameterize fluxes between the deep ocean and mixed layer? If yes, how are these fluxes calculated? What is the depth of the mixed layer? Prescribed or dynamic? Studies have shown that the mixed layer depth was substantially greater at the LGM (e.g. Sherriff-Tadano et al., 2018), which can have profound implications for the ocean heat contents and energy exchange between the ocean and atmosphere.

LSG is a three-dimensional general circulation model for the entire ocean. It solves the primitive equations for the ocean under assumptions of large spatial and time scales, which filters out relatively fast components like Kelvin waves. Given these assumptions and that LSG solves its equations implicitly, the time steps used are much longer than would be commonly used for other dynamical ocean models (in this study, approximately 4 simulation days). However, in order to allow the model to respond more quickly than this to abrupt or short-lived changes at the ocean's top surface, a mixed-layer ocean model is used as an intermediary between LSG and the rest of the model. The mixed-layer model is fixed to a 50m depth, which corresponds to the depth of the top layer of LSG. LSG itself does not have a fixed mixed-layer depth, and fluxes between the deep ocean and the mixed-layer are calculated as part of an LSG integration. The mixed-layer ocean model is made to relax gradually toward the LSG solution over LSG's timestep via an applied bottom-boundary heat flux, but it is also free to respond to changing thermal forcings at its top boundary from the atmosphere and sea ice. The above details have been added to the supplement (or appendix).

Page 6, line 17: Please clarify, you update the boundary conditions every simulation year, but with 10x acceleration, meaning that you effectively only run every 10 years from LGM to PI. Is that correct?

As far as the forcings go, the referee's description is correct. However, since the model is run continuously forward in time under these accelerated conditions, the solution will not be the same as if we extracted one of every 10 years from an unaccelerated simulation. The differences are described in Appendix A2 and are most noticeable in phenomena with a decadal or longer response timescale. For example, if the ocean's mixed layer takes approximately 30 simulation years to fully adjust to a change in atmospheric boundary conditions, the forcings will have progressed 300 years in this time. Thus, the timescales of these responses will appear lengthened.

Page 7, line 12: "...temporal resolution of 100 years". This seems to conflict with the synchronous update of boundary conditions described above.

As stated on page 7, lines 1-3, not all of the boundary conditions were available at annual or decadal timescales. Where they weren't available, we interpolated their values linearly in time between available bracketing time points. Thus, all of the boundary conditions were updated at the start of every simulation year.

Page 7, line 5: Please clarify how this process works. You can't fit an even number of 0.5◦ grid cells in a T42 cell (which is around 2.8◦), so there must be some partial overlapping cells. Also, what does "effective higher-resolution grid cell length" mean?

We calculated the variance of higher-resolution grid cells within the T42 grid cells in the following manner. Writing the variance as the SUM(x^2)-(SUM(x))^2 and using conservative remapping as an area-weighted sum over overlapping regions in grid cells between the two grids (for a description of conservative remapping, see Jones, 1999), we conservatively remapped both the elevation and the square of the elevation from the higher-resolution grid to T42. We then squared the remapped elevation and took the difference as in the equation above to get the variance. The number of high-resolution grid cells contributing to this variance was approximated by the ratio of the total number of grid cells in the global high-resolution grid divided by the number of T42 grid cells. The effective higher-resolution grid cell length was then defined by dividing the area of the T42 grid cell evenly over the number of high-resolution grid cells contributing to the variance and taking the square root (assuming each of these grid cells are squares).

Page 8, line 22: Century should probably be millennium here (you discuss 21 ka – 20 ka and 1 ka to

1950), right?

Since the PlaSim simulations were generated with the forcings accelerated in time by a factor of ten, a millennium of forcing changes elapsed during the first or last centuries of the simulations. Thus, we were analysing 100 years of output and comparing them against a millennium of data from unaccelerated simulations.

Page 16, line 15: What standard metrics?

"Standard" here is intended to mean "commonly-used." The definitions employed here for jet latitudinal position and tilt arise from Woollings et al (2010) and a combination of Woollings and Blackburn (2012) and Lofverstrom and Lora (2017), respectively. However, variants on these definitions have been used in many of the papers discussed in the Introduction.

Page 17, line 5: Meaning here is not clear. Do you mean flat ice sheets (i.e., only accounting for the albedo effect)? Also, the ice sheet height is not the only thing influencing the circulation. As you say elsewhere, the spatial extent is also important.

The PDTopo experiment is defined with all forcings varying in time except the ice sheet thickness. The thickness of the ice sheets are fixed to present-day values, so the land elevation remains the same as during present day at all times. Nevertheless, the ice sheet area varies from an LGM extent to present-day (i.e an infinitessimally thin ice sheet). This allows the role of the ice sheet orography to be separated from the influence of its albedo and the rest of the forcings.

We do not claim that ice sheet height is the only thing influencing the circulation. Rather, the text says, "The component of the ice sheets that appears most important to this effect is their elevation." We support this claim by noting in the results of the PDTopo experiment (with time dependent ice area) that at LGM, the primary location of the jet is not shifted equatorward and is only slightly more focussed than during present-day. Since we know that the ice sheet is the primary control for the western side of the jet being equatorward-shifted and highly focussed via the results of the FixedGlac experiment (where all other forcings vary in time, but the western side of the jet remains predominantly in the same state throughout the deglaciation), this suggests that either the ice sheet orography or the combined effect of the orography and area are creating these effects. Since our DarkGlac experiment differed very little from the full-forcing runs due to extensive snow cover, we can not differentiate between these two possibilities. However, we can say that an elevated ice sheet is required for this effect .

Page 18, line 1: How did you arrived at this specific number (725 m)?

We determined the elevation threshold above which the jet appeared latitudinally restricted by empirical testing. The southernmost latitude of the ice sheet in eastern North America was identified by applying a mask for regions at or above the elevation value being tested and identifying its southernmost latitude value in this region. Then, as in Figure 12, the North Atlantic jet latitude for each month was plotted against the corresponding ice sheet minimum latitude. If the jet latitude ever exceeded the ice sheet latitude in any of the FullyTrans simulations, then that elevation value was rejected as the threshold.

Page 19, lines 3-10: This explanation is a bit too simplistic. I agree that the presence of the ice sheet constrains the jet latitude in the west, presumable in part because of obstruction of the flow by the *topography. However, the thermal gradient at the southern ice margin can influence the flow in a similar fashion (this is not mentioned here as far as I can see) – both the change in albedo at the ice sheet margin, and the adiabatic cooling of the flow by the implied elevation difference. The modern (PI) jet is also less variable in the western ocean basin because of the strong thermal gradient at the sea-ice edge. This is clearly a different mechanism than the presence of a big ice sheet, but the effect is similar.*

We reject the hypothesis that the albedo change along the southern margin of the ice sheet plays an important role in constraining the jet that we detect over North America, because there is no such constraining effect in the PDTopo experiment (See Figure 12). The PDTopo experiment includes a time-evolving, but infinitessimally-thin ice sheet, so the albedo change along the ice sheet's southern margin varies the same way in time in the PDTopo experiments as it does in the FullyTrans runs. Thus, the ice sheet must be elevated in order for this barrier effect to occur. Our analyses can not distinguish between whether the elevated barrier operates via a dynamical effect alone or also via a thermal effect associated with the temperature gradient along the southern margin of the ice sheet.

Page 21, line 25: Meaning here is not clear – this seems to be the definition of a shift in the jet latitude.

We agree that the wording used does not effectively bring out our point. What we intended to argue here is that the frequency maps calculated from the Trace-21ka data do not show any change in the preferred jet tilt between 14 and 13ka BP. These results are contrary to what is suggested by the results of Lofverstrom and Lora (2017), but there are two important differences in the methodology here from what they used. Firstly, although both studies examine the Trace-21ka data, the abrupt increase in jet tilt presented in Lofverstrom and Lora (2017) was detected at 250 hPa, whereas we analysed jet changes over 700-925hPa. Secondly, the jet tilt was defined in Lofverstrom and Lora (2017) as the difference in jet positions between 10-20°W and 70-80°EW, whereas we defined the jet tilt as the difference between 0-30°W and 60-90°W. When we alter our analysis conditions to match those of Lofverstrom and Lora (2017), we see that the preferred angle of jet tilt does not change around 13.9 ka BP. Rather, the frequency of time spent at this tilt and less tilted values decreases, while the range and frequency of tilts increases for more positive values. This combination of phenomena matches the abrupt increase in mean jet tilt presented by Lofverstrom and Lora (2017).

Page 22, line 16: What Figure is discussed here?

Figure 13, bottom left plot.

Page 22, line 25:I would encourage you to think a little bit more about this and try to give a mechanistic explanation for this phenomena. Doesn't have to be a full explanation, but at least something that adds a little bit more to the story.

We will expand on our discussion as discussed at the start of this document.

Figure 1: What ice sheet remained in North America through the Holocene and is 1.5 km thick?

Figure 1 plots the peak elevation of ice sheet-covered areas in North America and Fennoscandia. There is no 1.5km thick ice sheet over North America during the Holocene, but there are regions that have an elevation in excess of 1.5km that are covered by ice: glaciers in the Rockies, for example.

Figure 4: Panels showing LGM and past1000 are mixed up (LGM is shown in middle panels).

This is correct. We will fix the labels.

Figure 9: 10 successive years is a bit ambiguous because it can be done in at least two different ways: (1) a sliding mean where the input and output arrays have the same length; (2) form decadal averages where the input array is 10x longer than the output array. These methods will yield slightly different results. I doubt that the difference will be of su#cient magnitude to challenge your conclusions, but this type of information is important for reproducibility.

Neither of these methods were used to generate the plots of Figure 9 as no averages were performed. Instead, monthly jet latitudes and tilt were collected for every month in DJF for 10 successive years. Then, the fraction of time that the jet spent in each latitude/tilt bin was calculated by summing over the number of months with a jet in the given bin and dividing by the total number of months in the sample (= 30 months).

Figure 11: Caption appears to be wrong as you show latitude here, not difference in latitude across the N Atlantic.

The reviewer is correct. The caption will be changed.

Heather Andres and Lev Tarasov