This is a potentially intriguing paper. It takes an existing database of bipolar volcanic signals (from Svensson et al 2020), and looks at the timing of eruptions relative to the onset of Dansgaard-Oeschger warmings (GI onsets) in Greenland. The main result is that eruptions in the database occur within 20 years of a DO onset more often than can be expected by chance. As a result it suggests that eruptions can help to push the AMOC (assumed to be the agent responsible for the DO event) over the edge.

If correct this is exciting. By no means all DO events have such a volcano near their onset, and large eruptions take place during glacials without causing a DO warming, so the conclusion would be reasonably stated as showing that eruptions are neither necessary nor sufficient for a DO event but they can precipitate the start of an event that would have happened eventually anyway. The authors don't really have a mechanism for why the volcano should cause such an acceleration of the event but just putting the idea out there would no doubt encourage new research and ideas.

But the issue is whether they have really shown what they claim. The statistics they use need to be looked at by a statistician but I am convinced that they are correct that the events logged in the Svensson et al (2020) paper do occur close to a DO warming more often than can be explained by chance. However, are those events a representative set of bipolar eruptions? I have my doubts, and I think the authors need to consider and (find ways to reject) an alternative.

Firstly the database contains only 81 events in 40 kyr of data. As the authors point out themselves this return rate of 1 every 500 years, is far below what is observed in the top 2400 years. Sigl et al (2015) found 81 bipolar eruptions in 2400 years (1 every 30 years) of which they classed 40 (1 in 60 years) as "large" (meaning a bit bigger than Pinatubo and a bit less than half the strength of Tambora). So we can already see that it is not surprising at all that there is a bipolar eruption within 20 (or 50) years of many DO warmings – if there is something special about the smaller subset considered by Svensson it needs to be that the chosen eruptions are large – the 1 in 500 year size of the 1458 or 1257 eruptions.

The present paper argues that they see fewer eruptions than Sigl because only very large eruptions exceed the background of the glacial sulfate and the lower resolution caused by thinning. However this seems unlikely to be the issue: while other chemistry changes a lot, the nssS background in the glacial (McConnell data online) is not very much higher than that of the Holocene. Volcanic eruption peaks are easily picked out even at sites with much lower resolution such as Dome C, and the layer thickness at WAIS Divide is still about 2 cm/year even at 50 ka age. Furthermore the peaks that occur near to terminations are by no means large ones. I didn't investigate all of them, and it's hard work to assess the peaks because the 2020 paper gives no scale for the y-axis but if we take the first 3 DO warmings where a bipolar volcano is identified within 20 years:

For the Bolling onset, eruption at 14705 years (GICC05), we can see in Svensson et al (2020) that it's a pretty modest peak in all records. It's barely visible at EDC, about the 6th largest peak between 14500 and 15000 years at EDML, while at NGRIP and WDC there are 8 peaks larger than this one in the 2000 years shown in Fig S2A of Svensson et al (2020). In WDC (using the online data) it rises about 50 ppb S (150 ppb sulfate) above the background. I estimated the flux from its width and the accumulation rate, and estimate that it's a modest eruption that would not meet Sigl's definition of a "large" eruption. In any case it is clearly not a 1 in 500 years eruption n any of the records.

For the onset of GI3, eruption at 27797 years, it's again a very modest peak. It's not really visible in any of the electrical records shown in Fig S4A of Svensson et al (2020) and it's a tiny peak in the only

Antarctic S record shown (WDC) – barely above the background (about 30 ppbS above the background), and a factor 10 smaller than several peaks in the 3000 years shown on the figure.

Finally for the onset of GI8, at 38232 years, it's again a modest peak in WDC, and in Greenland cores, the peak is less than half the size of a peak that preceded it by a century.

So, I would contend that the peaks near terminations are not especially large, and this cannot explain why they are the chosen subset of the much more numerous bipolar eruptions that exist. This then raises the suspicion that the reason there are more bipolar eruptions near GI onsets is something different: did the authors of Svensson et al (2020) find it easier to identify bipolar eruptions when they are near to DO onsets, or did they concentrate their efforts on the time periods around DO onsets.

The first issue (easier to identify bipolar eruptions near onsets) is obvious: these are the points where the relative age of Greenland and Antarctic cores are rather constrained by the existing methane matches that occur at DO onsets. This makes it much easier to be confident two eruptions in the two hemispheres that might match are actually at the same date. Hundreds of years before the DO onset, the relative age models are rather unconstrained and the authors of the 2020 paper would have felt less confident in calling two peaks near each other bipolar.

In addition the purpose of the 2020 paper was to look at the timing of events in Greenland and Antarctica at DO onsets. As a result, by their own admission they put a lot more effort into finding tiepoints around the DO warmings, which they did more confidently when they had layer counting. They reported "The bipolar layer counting is not continuous but is **focused on periods of abrupt climate variability** or high volcanic activity". They also noted "In addition to the published volcanic match points made for Antarctica, some 25 additional Antarctic match points have been identified in the present study to strengthen the synchronization **in the neighbourhood of Greenland abrupt climate change events**."

There is nothing surprising or bad about this, but it's just a fact that Svensson et al (2020) wanted matchpoints near the DO warmings and so they put more work into those areas, which are also the areas where it's easier to pin the relative dates down. As an example if I look at the section near the start of GI3 (27800 to 28500) I can see by eye at least 2 other potential bipolar matches of peaks in both WDC and NGRIP of similar size to the one identified at the onset, but where (correctly) the relative uncertainty in dating precluded marking them with certainty.

In conclusion, I believe that this is a topic worth studying, and I agree that the peaks in the Svensson (2020) compilation occur close to GI onsets more often than would be expected. However I am doubtful, using the information given in that paper, that the 81 bipolar peaks identified are an unbiased sample in time. I am not entirely sure how the authors can overcome my concerns. Probably the only way would be to run an algorithm that seeks possible bipolar matches within the prior (ie based only on methane) relative age uncertainty, so that an unbiased sample of possible bipolar peaks can be compiled and compared to the timing of GI onsets.

The authors tell us they are working on looking at the magnitude of the eruptions they identified and perhaps this will help them. It remains a fact that (based on Sigl's work) it is more likely than not that a modest bipolar eruption will occur by chance every 30 years and a "large" one every 60 years, so unless the authors can show that these ones near GI onsets are the rare "very large" ones, then they have not discovered anything very interesting.

In addition to this long and major comment above, I have some less important points:

Page 2, line 35 “by regarding” is strange wording, probably “by concentrating on”
Page 4 line 16 “Like this” should be “Thus” or “In this way”
Page 5, lines 1-5. I have now read this several times but I still can’t follow what you did. I get that you are seeking a point that exceeds the noise of the glacial data, but in detail the method is hard to understand. Did you start with the mean of the points within 80 years of the onset or before 80 years (and if so how far back?). For short stadials, is the set point 75% of the way back to the next GI or 75% since the last GI? Where is t_j – at the lower threshold or the upper threshold. I think you would need a section in the supplement that explains this clearly, with a diagram, and also explains why you chose this method over previous methods such as used by Rasmussen, or by those using Rampfit type methods. Having said that I accept that the points you have deduced are reasonable, so I am not objecting to the method but simply couldn’t follow it from the text provided.
Page 7, lines 10-15 essentially repeats what you already said at page 6 lines 5-8: needs redrafting.