Dear Editor,

Please find our answers to the reviews (second round, Mike Fromm, Ref.#1, Ref.#2).

First of all, we thank all reviewers for careful reading, their interest, and for making further critical remarks and good suggestions.

Our answers in BLUE:

Editor comment: Extraordinary claims need extraordinary evidence, and the self-lofting of smoke from the lower troposphere to the stratosphere falls into that category for me.

Let me (Albert Ansmann) start with the following remark. We shouldn't mix the major goal of the paper which is the occurrence of a thick smoke layer over the North Pole region (from our MOSAiC campaign). The self-lifting aspect is of a second-order importance. The self-lifting aspect was introduced to explain our statement: if we detected smoke, what was the most obvious source for it? So, the reader expects an answer, even if we have only one possible explanation to our finding- yet presented it as hypothesis.

The message of the paper is simple We observed a smoke-dominated aerosol layer over the North Pole region throughout the winter half year, and that's it! And there is no doubt that we observed a smoke-dominating aerosol layer because of the unique potential of a dual wavelength Raman lidar to measure the lidar ratio (LR) at two wavelengths and thus to provide this unambiguous optical fingerprint of smoke, the inverse LR wavelength dependence. This is highlighted in Table 2 and discussed in the Sect.4.1. There is no way for an alternative interpretation. And again, there is no better way to identify smoke than with dual wavelength Raman lidar. However, without the self-lifting aspect, the explanation for the "smoke source" might be missing by a reader. So we need this inspiring aspect to further initiate a scientific discussion.

Back to the paper content, we did a rather comprehensive and careful discussion on the self-lifting. There is nothing to add. Our revision consumed several weeks of preparing convincing – from our opinion- results, generation and revision of figures, and a careful and critical discussion. We clearly left open to the reader to accept or not accept our argumentation. The paper will trigger new research regarding the source of the smoke that was measured over the North Pole region.

Regarding Mike Fromm comments, we were glad to find out that our arguments in a revised version are much more convincing comparing to the original submitted manuscript. However, there were some points for additional discussion.

The main points of Mike Fromm's review to our opinion are, (a) we ignored the volcanic sulfate contribution to the observed aerosol. This point is now improved, Sect 4.1, 4.2; (b) that MODIS AOT maps (Fig.1 and 3) are biased by clouds. Here we provide missing information in Sect.3 now, that we carefully analyzed the data to avoid a cloud contamination bias and provide references that studied the cloud impact during moderate pollution events with the conventional MODIS retrieval that we used (Chudnovsky et al. 2013a, 2013b); and (c) that presentation of observations showing the stepwise ascent of smoke, day by day would be nice. Unfortunately, the latter is not possible, and we explain this point based on simulations (see explanations in Sect 3). Any heating of an absorbing layer in the troposphere will immediately lead to incoherent structures in the aerosol layering, because of much stronger heating at the top than at the center or base of a given layer. This behavior is different for troposphere and stratosphere (in the stratosphere coherent structures may be retained). Instead of the requested observations, we present a first self-lifting simulation figure (new Fig. 6).

Let me first summarize all the improvements we did in our secondly revised version:

- **Figures 8, 11, 12, 13, 14, 15, and 18 are now colorblind-safe as recommended by Ref #2. We used blue, green and black/dark-gray colors, only. Fig 10 is not improved because all curves are close together. Colors are unimportant then. See more information at the end of this letter.**

- **We followed all suggestions of Ref #1, regarding Raman method (<20km), least squares fit, smoke particle density, depolarization ratio, etc, and provide proper references. See step by step answers at the end of this letter.**

- **We discuss the impact of the Raikoke volcanic sulfate aerosol on the observed aerosol optical properties in more detail (Sect 4.1, 4.2). The Raikoke aerosol fraction was most probably about 10-15\%, and AOT of the order of 0.005 in Dec 2019 to Feb 2020. We mention that also in the abstract and in the conclusion section now.**

- **We present a new Figure (Fig. 6) with first results of our self lifting simulations in Sect 3, and provide a discussion including the consequences of vertically inhomogeneous heating (highest values at smoke layer top, lowest at layer base) which leads probably to diffuse rather than coherent aerosol structures …. which is reflected in the CALIPSO observations.**

Below please find a point by point response to Mike Fromm comments:

Summary Statement

O21's defense of their Siberian smoke hypothesis is admirable yet unconvincing. Hinging all their conclusions on the position that one and only one conclusion can be drawn from a peculiar wavelength dependence of lidar ratio places this paper at odds with a mountain of evidence that something other than smoke was in the stratosphere in 2019 before, during, and after O21's hypothesized smoke incursion scenario. If one considers the combination of stratospheric aerosol abundance, omnipresence, and longevity O21 show, this event is thereby in a category with ultra-strong pyroCb events and big volcanic eruptions. It is imperative then for O21 to offer more than a hypothesis if this work is to merit publication in its present form.

This argumentation was already stated during the first round of review. We are afraid that Mike Fromm does not accept lidar-based results and facts that are evident to all those who work with lidar. The basic Mike concern is: How one can be confident that he/she detected aged smoke?

We discussed this point comprehensively. We improved accordingly the text. We introduced a new convincing Table2. Unfortunately it seems that we start from the beginning. Below we clarified our point more comprehensively.

The fact is that this '*peculiar wavelength dependence of lidar ratio*' **(lidar ratio=LR) clearly and unambiguously points to a smoke-dominated source of aerosols. There is no way around this conclusion. Simply- impossible. Obviously, lidar experts can judge and acknowledge this enormous value of multiwavelength Raman lidars. And in the case of the MOSAiC smoke layer, the difference of 30 sr (LR355= 55sr, LR532= 85 sr) is high and does not leave room for a significant (e.g., 50%) contribution of volcanic sulfate aerosol. We cannot ignore it. We state and discuss this result clearly in Sect. 4.1 and 4.2. There is no room for any alternative combination. Modeling of particle optical properties would show that volcanic aerosol (from young to aged), urban haze, dust, and marine particles are unable to produce such a 'peculiar' LR spectral behavior. For readers who are not directly working with lidar data this result was clarified again in our revised version. And we thank Mike for this comment making our revised version more multidisciplinary, better understanding each other scientific language.**

Importantly, we observe these aged smoke layers since the Lindenberg Aerosol Characterization Experiment 1998 (since more than 20 years, see special issue in JGR on LACE98, 2002), and we summarized our smoke findings including this clear optical fingerprint for aged smoke, for the first time, in the paper of Mueller et al, (JGR 2005), and we detected so many layers since then, and never observed any other different exception cases of this nice LR spectral behavior of aged smoke. The pattern was the same for all our studied cases through many years of research.

Considering that- may be our explanation was lost across the revised paper – since Mike did not see it clearly, our discussion was improved and this point was clarified there again (on the ignorance of the volcanic sulfate impact and *'a mountain of evidence that something other than smoke was in the stratosphere in 2019 before, during, and after O21's hypothesized smoke incursion scenario').*

Regarding "a mountain of evidence" pointed out by Mike Fromm. It does not match the reality. 'Evidence' already indicates: We (as a whole scientific community) do not know exactly. To be clear again (as during the last round), I do not see this mountain! And why do I not see this mountain? There is no optical method (satellite passive remote sensing) that permits us to decide CLEARLY and UNAMBIGUOUSLY: This is sulfate aerosol! We discussed that already exhaustingly during the first round (first reply letter).

Motivated by this view of Mike Fromm, I discussed this point in a quite long e-mail conversion with Corinna Kloss (Kloss et al., 2021). She based her work on OMPS observations and came up with her constructive Raikoko ACP paper in early 2021. OMPS allows to retrieve TOTAL extinction (AOD), which is great, but not enough to distinguish between smoke and sulfate particles. Obviously guided by the volcanic aerosol modelers (co authors), Kloss et al (2021) exclusively focused on the Raikoke eruption and nothing else.

However, we understand and aware that we shall include also the impact of the Raikoke sulfate aerosol in the discussion. We improved this now in Sect. 4.1 (last paragraph) and 4.2.

In particulate, if we assume a volcanic sulfate LR of 40 sr (for both wavelengths, see Table 2) and a smoke LR355 of 60sr and a smoke LR532 of 100 sr, and an extinction-related Angstroem exponent of 1.5 for sulfate aerosol, and 0.75 for smoke aerosol, and finally assume a 15% contribution of smoke to the total AOT at 532 nm, only then we can reproduce the found overall LR values of LR355 of 55 sr and LR532 of 85 sr.

But LR355 of 60 and LR532 of 100 sr is already an extreme LR pair, if we would take 60sr and 90 sr, we could reproduce the measured values of about 55sr and 85 sr only with a smoke fraction of 5-10%.

All this is explained now in the final paragraph in Sect. 4.1.

And the final critical comment of Mike Fromm is …… *If one considers the combination of stratospheric aerosol abundance, omnipresence, and longevity O21 show, this event is thereby in a category with ultra-strong pyroCb events and big volcanic eruptions. It is imperative then for O21 to offer more than a hypothesis if this work is to merit publication in its present form…*

This ONE hypothesis of self-lifting explains so many optical features and measurements including the rather low particle depolarization ratio. In this regard, we do not think that we shall include additional arguments to our statements. We want to be on the safe side as much as possible. Any further (speculative) argument will not help. It remains hypothetically. Better to have just ONE, if that already is in consistency with our lidar observation of lidar ratio and depolarization ratio.

Importantly, we do not present our hypothesis in a dogmatic way. We leave it open to the reader to accept our hypothesis or not. In this regard, we truly believe that our paper will be extremely interesting to atmospheric science community initiating multidisciplinary discussion and new research in this direction. We have several views on the observed phenomenon of pollution transport in Polar regions. And we need to be ready that some of the results will transform our understanding of the different chemical and physical processes in the atmosphere.

In this regard, we did not change much in the discussion. Important addition is a simulation of the smoke self-lifting process (new Fig. 6). This will help to get an idea about self lifting times scales and that full particle ageing (after 24—48 hours the aging process is usually completed at tropospheric conditions) is really possible during the slow lifting over days before reaching the UTLS region.

Again, our major argument and a fact based on our measurements is: There was a dense smoke layer over the North Pole area…. and the logical consequence then is: Please tell us: What was the source? And if you do not know exactly: Please provide your opinion (hypothesis)? Only that is expected by a reader … and exactly that was done (comprehensively) in our paper.

If there would be a clear link (between the MOSAiC smoke layer and Siberian fires in July/August 2019), we would be happy to provide it. But we did not find any other possible source and stick to Siberian fires argumentation- which seems logical. This fire event was clearly remarkable and very large on a

spatial extent: by the end of the July 2019, the size of the fires reached 2,600,000 hectares. While Siberian wildfires are common during summer, record-breaking high temperatures and strong winds have made 2019 fires particularly devastating (https://www.bbc.com/news/world-europe-49224776). We offer a careful and detail discussion with many arguments and figures and argue that perhaps Siberian fires were the source of smoke in the Polar region. And all details provide a consistent picture. If later one, scientists find a better explanation (triggered by this paper), that will be great! We do not state, that we are right. We carefully state again and again: this is just a hypothesis, but for - the most convincing one so far.

The MOSAiC lidar data are an invaluable resource that will illuminate multiple, exciting findings. It is of course exciting to contemplate a new UTLS pathway for smoke to pollute the stratosphere in such a big way. But it would be equally exciting to learn that stratospheric smoke and volcanic sulfate might both give somewhat similar lidar-signal patterns. Given that there is overwhelming support for a hemispheric volcanic sulfate plume in 2019/20, the MOSAiC lidar-data analysis would be fundamentally improved if Raikoke influence was put on an equal footing (at least) with the hypothesized Siberian smoke explanation.

After *'mountain of evidence'* **we get confronted with** 'overwhelming support for a hemispheric volcanic sulfate plume in 2019/20'. **However, although bombarded by these overwhelming arguments, we remain biased by our solid, clear, and unambiguous observation that a smoke-dominated aerosol covered the North Pole throughout the winter half year of 2019/20.**

And regarding similar lidar-signal patterns: The relevant literature and own simulations of particle optical properties 'teach' us …….. that aged LIQUID sulfuric-acid-containing water droplets of volcanic origin with a typical and realistic accumulation-mode size distribution and simple, well-known refractive index characteristics for sulfuric-acid containing water droplets (with single scattering albedo close to 1.0) are UNABLE to produce lidar ratio of 40-50 sr at 355 nm and at the same time lidar ratios of 70-90 sr at 532 nm. Impossible! All the papers on volcanic aerosol (which partly include modelling results) show this clearly. Especially for aged volcanic sulfate layers (several month after formation), there is almost no way to produce significant differences between LR355 and LR532. We corroborated this simulation-based finding by our dual-wavelength Raman lidar measurements in 2009 (Mattis et al, JGR, 2010). These Mattis et al. volcanic LR355 and LR532 values are given in Table 2. We found LR355 and LR532 of about 40 sr. And in the case of these highly light-absorbing smoke particles with complex core-shell configuration and partly in glassy state, the story is quite RATHER different.

And regarding… *'Given that there is overwhelming support for a hemispheric volcanic sulfate plume in 2019/20, the MOSAiC lidar-data analysis would be fundamentally improved if Raikoke influence was put on an equal footing (at least) with the hypothesized Siberian smoke explanation'.*

What can we answer? After 'mountain of evidence' and 'overwhelming support…, now 'equal footing'. No comment from our side. We are exhausted!

As mentioned previously, we try to emphasize now more clearly that Raikoke sulfate aerosol may have contributed by 10-15% to the observed particle extinction coefficient and AOT at 532 nm, and that this is in agreement with the expected sulfate AOD of 0.005 (or max values of 0.01) in Dec 2019-Feb 2020 at high northern latitudes.

Major Concerns

O21 are standing by a concise claim that all of the MOSAiC UTLS lidar-detected aerosol they show is wildfire smoke. They give no quarter to other compositions (such as Raikoke sulfate). This is a huge challenge, which not only requires convincing the reader that only smoke can explain their measurements, but also to explain the fate of non-smoke particles in the UTLS that were undoubtably abundant at high latitudes in summer 2019.

We re-checked the text to avoid the impression that we completely ignored the volcanic aerosol. We explained that already in this reply letter above.

The new Fig. 5 shows that all trajectory overpasses of Box 2 occurred before or at the onset the ramp-up of Siberian AOD (Fig. 3). The 7-km trajectory is centered in Box 2 on 17 July, 4 days before the onset of the large- AOD episode declared by O21. The 9-km trajectory parcel moves rapidly over Box 2 on 20 July, when AOD is unremarkable.

The analysis of the trajectories presented by Mike From is incorrect. The red trajectory (arriving at 3 km height) was clearly within BOX 2 on 22 and 23 July 2019 (and thus during the heavy fire days) and even within BOX 1 (the hot spot fire region). The fires were intense from 19 July to 14 August 2019. Then the smoke had 3-4 days to ascend before CALIPSO came along and measured the lofted smoke. In these 3-4 days, the smoke can be lifted by 5-8 km as our simulations show. The new Fig. 6 corroborates that.

BOX 1 and BOX 2 were arbitrarily defined to better guide the reader to the most intense fire places. But there was also smoke to the left of BOX 2 with monthly mean AOD around 0.8. So, even the 7 km backward trajectory clearly shows… that the air mass was close to the ground (below 3000m) until 22 July. So, there were four days to accumulate smoke since 19 July 2019.

Finally, yes, the 9 km trajectory was not in direct contact with emitted smoke plumes. We state that in the manuscript. So, we need self-lifting to get smoke into the 9 km height range. Obviously, this was the case (from 3-9 km in about 2-6 days, according to Fig.6).

Regarding 'ramp-up', this was not a ramp-up event… (like a pryoCb event). It was a smooth change in the layering structures that occurred in the end of July and beginning of August. Even over Leipzig (51 N), and thus not only over latitudes >70N, we saw a significant change in the aerosol geometrical (and optical) properties, between 29 July and 5 August 2019, from the occurrence of only sharp and geometrically thin layers (usually of 500 to 1000 m vertical extent) to more complex conditions with a smooth and thick layer structure (with vertical extent over several kilometers) just above the tropopause in addition to the sharp layers above 15 km height. All this will be shown in a new paper in preparation (Ansmann et al.: Misclassification of stratospheric Siberian fire smoke as Raikoke volcanic aerosol in 2019 by the CALIPSO aerosol-typing scheme). This paper will be submitted in August/September 2021.

For the explanation of self-lofted smoke to be of consequence, conditions within Box 2 on 20 July would have to have been primed by large, low-altitude AOD some days prior.

The smoke needs 2-6 days to ascend from 3-4 km height (typical injection height) to 9-10 km height. So, to explain CALIPSO observation on 26 July, we need smoke at low level on 21-23 July (and not before 20 July).

Another aspect that became clear to us based on the simulations: The heating of a given dense smoke layer is vertically inhomogeneous. For example, at cloud-free conditions, strong heating occurs at the top and much lower heating occurs at the center and even much lower heating occurs at the base of a given layer. This means, you will find immediately incoherent structures. There is no way to lift the smoke layer as a whole to large

tropospheric heights. And exactly these diffuse (almost random) structures are visible in the CALIPSO lidar observations.

This feature becomes different for stratospheric heights (for stable layering conditions, very different profile of pot. temp. compared to the profile in the troposphere). Smoke plumes reaching the stratosphere (e.g., by pyroCB activity) and forming stratospheric layers can keep their layer structures for a long time during the ascent by roughly 3-7 km during a 10000 km travel, as we showed in the Ohneiser et al. (2020) paper on Australian smoke observed over southern Chile for layers arriving in the beginning of January 2020.

Per O21's analysis of Fig. 1 and 3, such a condition did not exist. Hence it is unclear why that CALIOP curtain is shown in support of their premise. Moreover, in my original review I explained that CALIOP curtains looked like the 26 July one every day before that for several days. Thus, there is a consistent picture of ubiquitous UTLS aerosol in place that cannot be attributed to Siberian smoke. Whether the ambient UTLS aerosol is smoke from previous, unrelated injections or Raikoke sulfate, it must be confronted in terms of what happened to it such that it apparently vanished and was replaced by self-lofted Siberian smoke.

Yes, there was Raikoke aerosol in the stratosphere before 26 July, nobody is telling the opposite. But in the beginning of August there was quite a big change in observable aerosol properties. Please, start to accept this fact!

The CALIPSO lidar observation are in full agreement with the simulations that lead to diffuse layering structures in the free torposphere, and therefore we show this CALIPSO observations. These observations perfectly support our hypothesis. So we are puzzled by Mike Fromm comment stating that 'there is a consistent picture of ubiquitous UTLS aerosol in place that cannot be attributed to Siberian smoke'. **We agree, if we discuss all the sharp layers above the tropopause (with 500 to 1000 m vertical extent). Most of them seem to be related to Raikoke aerosol. But all the diffuse layers (with 2-4 km vertical extent) in the UTLS regime seem to be of different origin. In the new paper (Ansmann et al., misclassification … when using the CALIPSO aerosol typing scheme), we will show an example (Leipzig, Polly lidar, 14 August 2019, layer from tropopause to 4 km above tropopause, layer mean LR355=70sr, LR532=105sr), so again…. clearly smoke! But depol was low as measured by Polly and by the CALIPSO lidar (near Leipzig overflight on 14 August 2019, within the Polly measurement window) so that CALIPSO announced: Sulfate aerosol!**

O21 base their hypothesis (and an upcoming paper) on Boers et al. and de Laat et al. Boers et al. was a theoretical prelude to de Laat et al., laying the framework for the Solar Escalator paper. On its own, the Boers et al. paper stands as a still unproven mechanism for lofting smoke from the lower to upper troposphere. de Laat et al.'s position, that pyroCbs did not occur on Black Saturday, has been contradicted by observations given in multiple publications (BOM, 2009; Cruz et al., 2012; Dowdy et al., 2017).

Pumphrey et al. (2011) proved that stratospheric enhancements of Black Saturday emissions were detected on the day after the pyroCbs. If BOM, Cruz, Dowdy were in error, and the Boers/de Laat mechanism was solely responsible for the stratospheric smoke plume documented by Pumphrey et al. and Siddaway and Petelina (2012), it would be reasonable to predict that O21's hypothesized aging, and its impact on particle depolarization, would drive the post-Black Saturday lidar landscape.

The paper of Boers et al. (2010) is clearly the first paper that points to the smoke self lifting potential. I studied all papers in this field already in 2017.

On the other hand, there is no doubt that the Black Saturday smoke (in 2009) reached the stratosphere by pyroCb activity. And CALIPSO depolarization ratios were clearly enhanced, a clear sign for the impact of pyroCb activity.

CALIOP measurements of the stratospheric plume would exhibit the same contradictory signals as claimed by O21. I.e. the Black Saturday stratospheric smoke would embody nil depolarization and thus be dominantly mis-classified as sulfate. This is not the case. A perusal of CALIOP backscatter curtains of ~1.5-month-old Black Saturday smoke reveals native measurements of enhanced depolarization. In fact, it is likely that the enhanced depolarization was a factor in the CALIOP version-4 feature classification scheme. The layers are regularly labeled as "cirrus" in lock step with classification of the layer as cloud composed of ice. This is in spite of the fact that the layers are above 20 km altitude. An example of one such scene is given here: https://tinyurl.com/caliopsmoke

Agreement at all. The consequence of fast lifting by pryoCB convection is that smoke has no time for aging and developing a spherical shape. So, depol is enhanced.

O21 are encouraged to survey additional CALIOP aged Black Saturday smoke detections from March 2009. They reveal other spurious classifications (such as volcanic sulfate) mixed with cirrus. It is evident that the best explanation for these features is smoke from Black Saturday. (See Siddaway and Petelina and Pumphrey et a;. for maps of the advected Black Saturday plume in the tropics.). As with the boreal 2019/20 situation, the lesson is that no single remote-sensing instrument probing the UTLS is sufficient for unambiguous characterization of particulate composition. This is why total reliance on MOSAiC lidars for characterizing three seasons' worth of aerosol observations requires several complementary data items, and the context provided by publications such as Kloss et al. (2021) and Cameron et al. (2020) in addition to Johnson et at. (2021).

We do not agree with your comment. We totally agree that the combination of many techniques is always of advantage, to create a trustworthy, consistent view on the observed (puzzle-like) features and processes. However, we disagree when all these techniques only provide 'evidence' and 'consistency' but not unambiguity or clarity! ... as the dual wavelength Raman lidar is able to provide. This makes a big difference. We rely on our good measurements and (my) almost 40 years of experience in the field of lidar aerosol field observations.

The argument made by O21 regarding the published sAOD (e.g. Kloss et al.) falling short of the MOSAiC 0.1 value is without much merit. There is little doubt that sAOD in Kloss et al. is probably biased low, in part for the reason given in O21—saturation. The sAOD values shown therein, peaking at about 0.025, are also an artifact of the broad aerial/temporal averaging applied. Hence they make a poor point of comparison with individual lidar profiles. That being said, it is straightforward to see in CALIOP data that the stratospheric aerosol at high latitude prior to the hypothesized Siberia incursion, far exceeds sAOD=0.025. Take for example a CALIOP curtain on 22 July, with an aerosol layer over North America with native level stratospheric backscatter exceeding .003/sr. Applying a conservative lidar ratio of 50 gives extinction exceeding 0.1. https://tinyurl.com/gtdot1

The volcanic aerosol was inhomogeneously distributed over the Northern Hemisphere, but the general Raikoke impact and the general trends are clear in Kloss et al. 2021 paper, and show an increasing AOT with increasing latitude, which is in full contradiction with model results (Sarychev, 2009, Raikoke, 2019) of a homogeneous spread of the volcanic aerosol over the northern part of the Northern Hemisphere and not an increase of AOT towards the North Pole. All this was already discussed last time and is written in our revised manuscript (already last time).

The MODIS AOD analysis in Fig. 2 is impressive but inconclusive. No accounting is given of any significant difference in peaks. Several additional peaks are also quite impressive. Might one conclude that in those years a similar, scalable impact on the stratosphere was predictable? Were any observed? It should be straightforward to do so with the available ground-based lidar and satellite remote sensing data sets. Another caveat is that MODIS AOD is severely low-biased in the presence of high-concentration aerosol plumes (Figure 7; Fromm et al., 2008). It is akin to a saturation bias; thick aerosol is classified as cloud. The import here is that there is a huge unknown in any MODIS AOD analysis focused on extraordinary plumes. The Siberia smoke situation in July/August 2019 was indeed extreme, but the true quantifiable extreme here and in many other cases is unknowable based solely on MODIS AOD retrievals. Hence it is unclear how quantifiably unique the Siberia 2019 smoke situation was. That being said, if indeed the Siberia 2019 smoke was lofted to the UTLS, it should be elementarily possible to follow the lofted smoke plume with satellite data such as CALIOP. If O21 can show observations of day-to-day, stepwise escalation of optically dense smoke from its initial placement to the tropopause and beyond (in accord with their preliminary theoretical calculations), this could be a compelling argument to include in the present thesis.

Now we enter another new field- MODIS AOD bias in the products.

As mentioned previously, we present a simulation of self-lifting (Fig 6), we cannot show observations of coherently ascending layers. Vertically inhomogeneous heating of the smoke layers lead to turbulent aerosol structures. This is written in the manuscript (Sect. 3).

Alexandra Chudnovsky (our co author) wrote: Very important critical comment. However, I do not think that AOD maps were so largely biased by cloud-contaminated pixels. Based on my experience and works, the AOD retrieval is biased during low pollution events (no AODs are generated) and for high/thick, dense dust storms (misinterpretation with clouds). For the Eastern USA- during smoke events (pollution transport following Canadian fires), the AOD data was provided and clouds were perfectly masked. This area as we know is largely covered by clouds and only 40-50% of yearly data is available for the analyses. For Siberian smoke- I looked at daily RGB and AOD maps. The latter show relatively high retrieval rate- the smoke was not so thick with a nice and spatially continues pattern. For cloud contamination I would expect spurious /spatially uneven pattern.
Here we deal with a moderate pollution event, of a very large spatial extent, clouds were masked and the algorithm – I would say- aggressive- meaning that also "good pixels" are excluded largely reducing the data base. I looked at all AOD and RGB images during July 20- August 20- 2019 and selected high AOD values- comparing it to cloud masked ones. For cloud adjacent pixels we have uneven spatial pattern – and still it did not bias the general view of montly averaged AOD. For some pixels- the AOD was not retrieved although one can see yellowish smoke above clouds on RGB images. I would say that we rather underestimate the strength of AOD spatial pattern during this event.

Some publications on the impact of clouds on the quality of AOD retrieval in several publications. For example- Chudnovsky et al. 2013a; 2013b; 2014; Kloog et al. 2014; Rogozovsky et al. 2021;

<u>Final remark</u>

Thank You Mike for all the questions and critical remarks, for your genuine interest in our work, spending so much time to read the different versions of the manuscript and comments writing. We learned a lot!

However, at the end I should add: I am working in the aerosol lidar field (plus passive remote sensing of aerosols from space and from ground, AERONET) since almost 40 years, and prepared about 50 papers as main writing author, and I must admit, this paper is one of the most exciting, sound, concise, and well-organized paper I was ever involved.

So now, we think, it is time to give readers a chance to 'learn' more about aerosols in the North Pole range in the winter of 2019/20, and as we now know, it was some kind of a mix or Raikoke and Siberian smoke aerosols.

Ref #2

To the extent possible, the authors should make an effort to ensure their figures are colorblind-safe. By my reckoning, Figures 7, 9, 11, 13, and 17 use color combinations that actually cause ambiguities and issues with distinguishing figure elements under different color deficiencies. Other figures could also be revisited. For example, in the figures with only 3-4 colors (such as Figures 10 and 14), there are better color choices that can be used than red, blue, and green. There are many web resources available (e.g., colorbrewer2, Adobe Color's accessibility tools, etc.) that can be leveraged to pick safe choices.

We considered this aspect now. See, Figs. 8, 11, 12, 13, 14, 15, and 18 (because of new Fig.6, new numbers). We try to use just blue (for 355 nm results) and green (for 532 nm results) and now, for 1064 nm or for PSCs, we select the color of dark gray to black. We checked several web pages and think this combination is ok…. Most important is to avoid red, when using green and blue.

Ref #1

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

The authors have substantially improved the article after the reviews. They addressed in detail the major points of concern raised in my review. The article can thus be published after following minor revisions have been considered:

P4 - L29: The sentence "The Raman lidar method was exclusively used to determine particle backscatter and extinction profiles" is not completely correct since it is mentioned a few lines later that "At heights above 20 km, the backscatter and extinction properties of the stratospheric background aerosol are determined from the elastic backscatter signal profiles (alone) by assuming a particle extinction-to-backscatter ratio (lidar ratio) of 50 sr".

We improved this: Raman lidar method is exclusively used for smoke (<20 km), and higher up the Fernald method is used.

P5 - L11: what is the reference for the fact that dust cause depolarization ratios around 0.3 at 532 nm?

We now use Gross et al. (2015) here, this reference was already included in the paper.

P5 - L25: The Baars et al (2016) reference is mentioned for the least-squares regression analysis used to determine particle extinction and extinction-to-backscatter ratio profiles, but I did not find an explanation of the method in the mentioned reference.

The references are now properly given: Pappalardo et al., 2004, and Russo et al., 2006.

P6 - L9: Some more explanation is required about the value of the smoke density used to obtain smoke mass concentrations.

We extended the text and provide some more information from the Ansmann et al (2021) paper, which is now published.