Answer and changes related to RC1

Dear Reviewer!

First of all thank you for careful reading and all good and constructive suggestions which improved the paper significantly (we hope). Before we answer all comments and questions, step by step and item by item, let us provide a brief overview of the essential changes:

We went deeply into the literature with special focus on volcanic sulfate aerosol (originating from the recent Raikoke and from the Sarychev volcanic eruptions), into our own Leipzig lidar data analysis, into the CALIPSO observations, we performed an extended HYSLPIT trajectory analysis and conducted further simulation studies to bring together much more solid information and substantial argumentation to provide a plausible link between the Siberian fires (in July-August 2019) and the polar smoke layer we observed later on in the North Pole region (since late September 2019). The focus is on the potential self-lifting process, on smoke aging and even on the consequences for the morphological and optical properties of smoke (such as the depolarization ratio).

We did all this very carefully and even if we present a large number of reasonable and plausible arguments for our chain of hypotheses we emphasize that our entire set of arguments still remains at a hypothetical level.

Nevertheless, there is no doubt! What we observed during the MOSAiC is clearly and unambiguously wildfire smoke. There is no way around. We observed this inverse spectral dependence of the lidar ratio (LR355 < LR532), and that is a unique fingerprint of wildfire smoke, and we observe this since 1998, in Canadian, Siberian, Australian smoke, … again and again.

And one of the convincing and motivating arguments to search for a much more massive aerosol source than the Raikoke volcano was: The Raikoke-related AOD at high northern latitudes was about 0.01 at 500 nm in autumn 2019 (according to simulations of the 2019 Raikoke as well as of the 2009 Sarychev volcanic aerosol impact and corroborated also by Raikoke aerosol observations at lower latitudes, discussed in Kloss et al., 2021). But, we observed AODs around 0.1 in autumn 2019 in the High Arctic, and thus an order of magnitude higher values. We were forced to find a massive aerosol source.

Some significant changes before we start the step by step reply:

- We show a new figure (Fig. 2, time series, MODIS AOD from 2000-2020), highlighting the extraordinarily strong 2019 fire season in central and eastern Siberia.
- We introduce a new table (Table 2), highlighting the unique signature (fingerprint) of aged wildfire smoke, namely the inverse spectral slope of the lidar ratio (LR355nm < LR532nm) together with high LR532. We compare these smoke features in the new table with the ones for mineral dust, volcanic sulfate and ash, Arctic, European, American haze, etc.
- We show a new HYSPLIT plot (Fig. 5) to better explain the link between the CALIPSO overflight over Siberia and the strong fires… downwind (west!) of the flight track.
- We show a new figure with the Polarstern track (Fig. 6, same as in Engelmann et al., 2021).
- We created a new figure (Fig. 7) to better show the influence of the polar vortex on all the observations.
- We introduced a new section (Section 3) on Siberian fires, self-lifting hypothesis, consequences for aging (leading to spherical smoke particles), and also on the potential miss-classification of smoke layers as sulfate layers when using the CALIPSO aerosol typing scheme (Sect. 3.1).

- We shortened the ozone-depletion/smoke/PSC section as requested, and combined the two figures of this section to one figure (Sect. 5, Fig. 17).
- We skipped several figures (layer-depth histogram, monthly mean extinction profiles) to keep the paper as short as possible.

We think, more cannot be done. Many parts of the paper are substantially improved, caused by the recommendations of the reviewers. But this kind of research must also tolerate hypothetical explanations and argumentation. We feel this approach, as we present it now, is justified.

Step-by-step reply: our answers in BLUE

Major comments

In the introduction, the authors mention that the burning season in 2019 was largest on record, they also employ superlatives like "tremendous environmental disaster" without a quantification of how the wildfires compare to previous years.

This is now improved, see new Figure 2

Little information is provided in the article on the Polarstern trajectory with respect to the stratospheric polar vortex. A figure showing potential vorticity and temperature evolution along the cruise is lacking. Such a figure would be very useful for interpretation of the measurements and especially for section 4 that addresses PSC, smoke aerosols and ozone depletion.

Thank you! We followed this idea, see new Figure 7 (presenting the scaled potential vorticity and temperature time series for different height levels). Before, we present the track of the Polarstern in Figure 6.

The fact that the observed aerosol layer originates from the Siberian wildfires is not demonstrated in the manuscript. As mentioned in the paper, several other sources of aerosols can be identified during that winter: volcanic aerosols from the Raikoke eruption, aerosols from pyroCb injection linked to fires in Northern America and PSC. The distinction between smoke and volcanic aerosols is made primarily from the spectral difference of the lidar ratio without consideration of meteorological processes that could further document the origin of the observed aerosol layer. Only one ensemble of trajectories is displayed for one day in the campaign in the manuscript. Such considerations on the evolution of meteorological evolution during the Arctic winter is needed especially in order to better differentiate smoke aerosol from PSCs.

We provide now a large number of plausible arguments and explanations that the Siberian fires were most probably the source for the smoke over the North Pole. We introduced a new section for an extended discussion (Section 3). But it remains impossible to 'demonstrate' this. It remains a hypothesis. All the fires in 2019 at high latitudes were of minor importance (according to the Kloss et al., 2021, paper). Thus, we have Raikoko aerosol and the Siberian smoke aerosol to consider.

We think any meteorologically based argumentation (including simulations of long-term air mass transport) would not help, is too uncertain. The temporal distance between the fires in July and August 2019 and the Polarstern observation in October 2019 is too large. Therefore we show Spitsbergen and Polarstern lidar measurements together in Fig. 12b. Spitsbergen lidar observations of the stratospheric aerosol are available from August 2019 (and thus for the fresh fire smoke from the Siberian fires) to January 2020, and thus we have even a long overlap between Spitsbergen and Polarstern lidar observations from October to January.

However, we know from many observations: If smoke enters the stratosphere it will be distributed quickly over large parts of the northern hemisphere, and this within a few weeks. This is already shown several times, e.g., for example recently after the strong Canadian fires (Baars et al, 2019) and much earlier by Fromm et al (2008). Fact is also that the decay of a smoke-related stratospheric perturbation takes more than a half year (Baars et al., 2019).

We believe, the key point is to show (and demonstrate): Do we have convincing argument that the Siberian smoke was able to reach the lower stratosphere (in the absence of pyroCb activity)? This aspect is discussed in large detail in Sect. 3. We develop a consistent picture that this was the case. The self-lifting process (solar absorption by smoke at high APT levels, warming of the air, and ascending of the warmed air within 3-5 days up to the tropopause) was able to lift smoke up to stratospheric heights so that the smoke was later on observed over Leipzig and Ny Alesund since August 2019 and over Polarstern since the late days of September 2019.

We discuss already in the Introduction section that the Raikoke volcanic eruption caused a stratospheric perturbation in terms of AOT of 0.015-0.02 at 500 nm during its maximum impact (around 10 August 2019). This is based on observations and Raikoke simulations (Kloss et al, 2021) as well as simulations of the similar Sarychev volcanic aerosol (Haywood et al., 2010, the eruption occurred 10 years before the Raikoke eruption, i.e., in the summer of 2009) and the Raikoke AOD decreased to 0.01 in October 2019 at high northern latitudes according to the simulations. To that time we observed AOTs of 0.1 (an order of magnitude higher AOT). So, we argue we clearly need another, a much stronger source for this massive stratospheric perturbation we observed during the MOSAiC campaign. The Raikoke sulfate aerosol cannot explain the strength of the aerosol load we observed. And all the other fire-related events in 2019 were even smaller than the Raikoke impact (as Kloss et al. , 2021, pointed out). All this is now given in the introduction, with the goal to clearly state: We need a strong source to explain the observed massive stratospheric perturbation. And after checking our lidar data at Leipzig and Ny Alesund and the strong increase in stratospheric AOD (towards 0.1-0.15 over Spitsbergen in the beginning of August 2019) we had to search for a strong fire area in July-August 2019. All this is described in the Introduction. All this is quite reasonable and consistent. But sure, it remains a hypothesis as we state that finally in Sect. 3.

Regarding the lidar ratio spectral dependence (LR355 < LR532) together with the high LR532: This is clearly our number-one argument to conclude: We observed aged smoke! There is no way around it. However, to corroborate this and to better convince the reader, we present a Table 2 with lidar ratio pairs (355, 532 nm) for aged Siberian, Australian, Canadian smoke, mineral dust, volcanic ash, volcanic sulfate, urban haze, Arctic haze, etc…. And the message is then clear. This was smoke what we observed during MOSAiC.

The methodology used for the analysis and retrieval of aerosols parameters is not described in sufficient quantitative details. The description relies mainly on references, some of which e.g. Ansmann et al., 2020 is still under review. For instance, section 2 does not summarize the method used for deriving main aerosols parameters such as the backscatter coefficient or the lidar ratio and their respective error. A table could summarize main characteristics of these retrieved parameters in terms of error and vertical resolution. A summary of the POLIPHON method used to derive the aerosol mass concentration is also needed. The GMAO method is used to retrieve the tropopause height but no explanation is given on why such method is better than the classical WMO one. Since tropopause height is generally more difficult to determine at high latitude than at lower ones, such explanation is needed. Also, a description of the method used to derive the bottom and top of the aerosol layer is lacking.

We considered all the remarks and improved Section 2 significantly along the suggestions.

The hypothesis of mixing between several particle types is not fully explored. For instance, there is no clear explanation of the quantification of less than 20% for the fraction of volcanic aerosol observed in Fall 2019 (page 7). In Figure 9, PSC are only identified over the smoke aerosol layer. Is there a possibility that PSC are also formed within the aerosol layer? Without knowledge of temperature history, it is difficult to conclude.

We skipped this part of the discussion on the mixing of volcanic and smoke aerosol and conclusion on the smoke and sulfate aerosol fractions.

Regarding the PSC influence: Yes, there were several cases with PSC development in the centers of the aerosol layer. We decided not to do any correction. It is better to live with the bias than to produce a new one by correcting the effect and introduce new uncertainties in this way. We had precise temperature information, because there were 4 radiosonde launches per day. So, we had always precise PSC formation temperatures.

What is the objective of section 3.3 (comparison with foregoing Aerosol studies)? This section cites a number of other studies analysing aerosol vertical distribution in the Arctic but not clear conclusion is driven from this section.

The goal was to show all these different measurements from 2000 to 2019 together with our MOSAiC observations to corroborate that our measurements fit very well into the High Arctic aerosol climatology, except the smoke layer with an order of magnitude higher extinction efficients. We leave the figure in (now Figure 13), but removed Section 3.3.

Specific comments

P2 L16 – 19: What parameters were considered from FIRMS and CAMS databases?

We skipped the respective sentences. However, one would find maps with fire spots (FIRMS) and information about the extraordinarily strong fire season of 2019 (CAMS).

P2 L34 – 35: Provide details on the mentioned simulations.

We do so now in Section 3, we used the radiative transfer model ecRAD of Hogan and Bozzo 2018, and we give precise information on all input parameters needed to simulate heating rates, and how we computed ascent rates as a function of heating rates by using the approach of Boers et al., (2010).

P3 L7 – 9: How do we know that the aerosol was trapped in the strong polar vortex?

We removed this statement. We do not know what is going on (in detail) below the vortex!

P5 L8 – 9: How were identified the PSC: by visual inspection? There is no explanation.

Yes, by visual inspection. We explain what we did to remove the impact, and how large the remaining uncertainty is in terms of AOT (about 5% or less). We did not make any further attempt because we believe we would introduce a (new) bias. It is better to see the remaining effect in the depolarization ratio and Angstroem time series and to get, in this way, an idea about the impact.

P6 L14: How are the bottom and top of the aerosol layer determined?

Again by visual inspection and Rayleigh signal fit to the measured signal profiles and then by using a threshold values of 1.1 for the 1064 nm total-to-Rayleigh backscatter ratio. We provide the details now in Section 4.1.

P6 L18 – 20: The HYSPLIT trajectories do not demonstrate that the smoke aerosol layer could have been trapped within the polar vortex. Figure 5 is not very clear: provide explanation for the colours of trajectories.

We improved the text. But we still mention that the aerosol was obviously trapped.

Colors are just used to distinguish different sub groups of trajectories. This is now explained.

P6 L23 – 24: How are determined error bars in Figure 6? Are they shown as one or 2 sigma?

This is one standard deviation, as usual.

P8 L28 – 31: How are the refractive index and SSA shown in Table 1 determined?

We state now in Section2: The single scattering albedo SSA, defined as the ratio of scattering-to-extinction coefficient, is finally calculated from the retrieved particle size distribution and complex refractive index characteristics with an uncertainty of $\pm$0.05.

P9 L22: Figure 10 is not well explained. Significance of the layers mentioned in the legend is not clear.

We skipped this figure.

P9 L23 – 24: If the aerosol layer was trapped in the strong polar vortex, how could it be influenced by smoke aerosol from lower latitudes? What about subsidence within the vortex? The situation of the lidar measurements with respect to the polar vortex is not clear and needs better description.

We introduced the new figure (Fig. 7) with the scaled potential vorticity (sPV) in (a) and temperature for different height levels in (b), along the Polarstern route. This was requested by the second reviewer as well. However, we leave out to discuss the impact on horizontal or vertical transport. We just mention that the vortex has a strong impact on the weather and air flow conditions below the vortex and widely suppresses meridional exchange.

Answers and changes related to RC2

Dear Reviewer!

First of all thank you for careful reading and all good and constructive suggestions which improved the paper significantly (we hope). It is usually often of great and complementary help when a reviewer is not an expert in the same field as the authors. You realized certainly that we are not working every day in the field of stratospheric ozone depletion and are also not just experts for polar meteorology.

Before we answer all comments and questions, step by step and item by item, let us provide a brief overview of the essential changes:

Motivated and mostly forced by the other two reviews, we went deeply into the literature with special focus on volcanic sulfate aerosol (originating from the recent Raikoke and from the Sarychev volcanic eruptions), into our own Leipzig lidar data analysis, into CALIPSO observations and HYSLPIT trajectory analysis and conducted further simulation studies to bring together much more solid information and substantial argumentation to provide a plausible link between the Siberian fires (in July-August 2019) and the polar smoke layer we observed later on in the North Pole region (since late September 2019).

Some significant changes before we start the step by step reply:

- We show a new figure (Fig. 2, time series, MODIS AOD from 2000-2020), highlighting the extraordinarily strong 2019 fire season in central and eastern Siberia.
- We introduce a new table (Table 2), highlighting the unique signature (fingerprint) of aged wildfire smoke, namely the inverse spectral slope of the lidar ratio (LR355nm < LR532nm) together with high LR532. We compare these smoke features in the new table with the ones for mineral dust, volcanic sulfate and ash, Arctic, European, American haze, etc.
- We show a new HYSPLIT plot (Fig. 5) to better explain the link between the CALIPSO overflight over Siberia and the strong fires… downwind (west!) of the flight track.
- We show a new figure with the Polarstern track (Fig. 6, same as in Engelmann et al., 2021).
- We created a new figure (Fig. 7) to better show the influence of the polar vortex on all the observations.
- We introduced a new section (Section 3) on Siberian fires, self-lifting hypothesis, consequences for aging (leading to spherical smoke particles), and also on the potential miss-classification of smoke layers as sulfate layers when using the CALIPSO aerosol typing scheme (Sect. 3.1).
- We shortened the ozone-depletion/smoke/PSC section as requested, and combined the two figures of this section to one figure (Sect. 5, Fig. 17).
- We skipped several figures (layer-depth histogram, monthly mean extinction profiles) to keep the paper as short as possible.

We think, more cannot be done. Many parts of the paper are substantially improved, caused by the recommendations of the reviewers. But this kind of research must also tolerate hypothetical explanations and argumentation. We feel this approach, as we present it now, is justified.

Step-by-step reply: our answers in BLUE

Ohneiser et al. present a summary of a persistent wildfire smoke aerosol layer observed during the MOSAiC field campaign in the high Arctic. The authors primarily focus on observations obtained from the "Polly" Raman lidar operated onboard the Polarstern, but also compare with other lidar platforms (e.g., CALIPSO and KARL). With the unique observations made possible by Polly and the MOSAiC campaign, the authors are able to provide quantitative descriptions of the optical and microphysical properties of observed aerosols for a region where such measurements are sparse. These measurements are made even more important by the unique atmospheric and aerosol conditions related to the 2019 Siberian wildfires and the unique polar vortex and ozone conditions of late-2019 and early 2020.

I am not a lidar measurement expert; in fact the editor has asked me to provide my views for portions of the manuscript related to polar vortex conditions and ozone depletion. However, I of course read through the entirety of the paper (multiple times), and so I feel comfortable saying that the paper is generally well-written. While I can not provide an expert assessment of the observational aspects and descriptions, as a "knowledgeable layperson", I'd say The authors mostly gave thorough and convincing descriptions of the observational data and their characterization. The paper is clearly cutting-edge and would be a perfect fit for ACP. However, I do have a few concerns, some more major than others; most of these are related to the aspects surrounding the polar vortex and ozone depletion in Section 4. Since Section 4 is a relatively small portion of the overall paper, I would say the revisions required to address my comments would be relatively minor to implement, but I do feel strongly that they are necessary changes.

General comments

(1) In reading through the paper, reaching Section 4 felt like a definite shift. Up to that point, everything felt (as a reader) mostly authoritative and backed up by quantitative results. In contrast, Section 4 felt very vague and overall speculative. One paragraph particularly stuck in my mind -- p14, L10-15 -- as it included many instances of language such as "were probably ..." and "were likely". Three "precise" research questions are posed early on in the section, and yet the remaining results and discussion do nothing to answer them and instead involve results that only hint at possibilities. Section 4 essentially says "the smoke aerosol layer could have been important for ozone depletion" simply because the PSC, smoke, and ozone depleted layers overlapped in the vertical (Figures 15 and 16).

We got the point and worked on Section 4 to meet the concerns. Section 4 is now Section 5 in the revised version. We re-arranged the text. See further explanations below.

Let us briefly start with this: There are many hypothetical arguments and aspects in the paper. This must be allowed as long we state that clearly. And we do so. For example, there is no clear answer to the question: What was the most likely source for the smoke we observed over the High Arctic? We try to answer this question in the new Section 3. There is also no clear answer to the question: Is it possible that the smoke contributed to the strong ozone depletion?

We think that nobody has an idea at the moment how efficient glassy organic particles are in heterogeneous chemical processes in the stratosphere and that nobody knows how these glassy particles can influence PSC formation! It will take years to investigate this properly! You need sophisticated laboratory studies for that. Modelling is insufficient.

Our job is simply to raise our hand to tell the ozone science community that there was a persistent smoke layer over the Arctic in the spring of 2020 when a record-breaking ozone depletion occurred. Please keep that into consideration in your future studies … .

By the way, we operate another lidar at Punta Arenas, Chile, and detected the Australian smoke from the tropopause up to 26-30 km height, from January 2020 to April 2021. We were sure, already in March-June 2020, that there will be a strong ozone depletion in the southern hemispheric spring season (September to November) because of the smoke…. sure, without any (precise) idea about the pathways…

There are many details that would have to be worked out in much more detail to answer the questions posed in Section 4. For instance, if the MOSAiC measurements were to be of use in answering these questions, then there would need to be details about the time-varying geometry of the vortex. While it's a fairly safe assumption that the vast majority of measurements (lidar, ozonesondes, etc) sampled air within the polar vortex, on weekly timescales the vortex and its edge are quite mobile (and similarly, so is the region of air cold enough to form PSCs). This means that measurements could sometimes sample air in different parts of the vortex, meaning some measurements would be less relevant than others in establishing where things were "in the right place at the right time".

First of all, this triggered us to present the new Figure 7 with the scaled potential vorticity and temperatures for different height levels along the Polarstern route from October 2020 to May 2020.

As mentioned, we re-phrased and rearranged the text to give the impression in Section 5 that we just want to point to a new aspect: smoke and ozone depletion. …. not more…

If we do not show that, others will do it. Furthermore, we see the same in the southern hemisphere: a strong ozone depletion height range, but this time right in the center of the Australian smoke layer.

As written, I do not feel Section 4 should be kept. I would suggest that it should be shortened and perhaps folded into the current Section 5, with content from Section 4 forming the basis of some discussion (i.e., providing motivating context for future work). At the very least, Section 4 should be shortened and rearranged so that the authors present the "possibility" before posing the big picture questions. The authors should also be clear (if they keep any of Figures 15 or 16) that conclusions cannot presently be drawn from their analysis.

This is done! We agree! We shortened Section 4 and rearranged the remaining text parts.

In my specific comments below I provide more detailed questions/comments related to the content of Section 4.

(2) Again, this is not my area of expertise so perhaps these are naive questions: The title of the paper and the introduction generally outline the assumption that the smoke aerosol layer measurements were tied to the Siberian wildfires, but are there not other potential confounding sources?

We changed the title! Removed 'Siberian'!

We agree, we should be more careful and therefore rephrased the introduction. But as you will see, there was the volcanic sulfate layer (with Raikoke as the source) and the smoke layer (from the huge Siberian fires as we hypothesize) and nothing else. There were many fires within the Arctic circle in 2019. But very strong and long-lasting fires were needed and probably also stagnant weather conditions to initiate self-lifting in order to end up with smoke in the stratosphere. Only if the smoke reaches stratospheric heights it will be distributed of the entire northern part of the Northern Hemisphere and will survive for months.

Can the authors provide more evidence to say that the Siberian wildfires were likely the dominant/overwhelming contribution to the observations discussed throughout the paper?

This motivated the new Section 3 on the Siberian fires, on self-lifting aspect, on smoke aging aspects, and even on the failure of the CALIPSO aerosol typing scheme to identify smoke, but instead to miss-classify that as sulfate aerosol.

The paper lacks some descriptive context as well: how severe were the 2019 Siberian wildfires in comparison to prior years? It is perhaps fortuitous that the MOSAiC campaign was able to sample aerosol conditions largely influenced by Siberian wildfires, but is there a quantitative measure of how unusual 2019 was that would better emphasize the importance of the MOSAiC measurements? These kinds of details may be present in the articles that the authors cite, but I think it's worthwhile for the authors to make them explicit where possible.

We include the new Figure 2 with MODIS AOD observations from 2000 to 2020. The maximum AODwas observed in August 2019 during the last 20 years.

Yes, it was fortuitous! And we (as aerosol lidar enthusiasts) were happy to observe this exotic aerosol layer throughout the winter half year of 2019-2020.

(3) There were a couple of times in the paper where a Figure is introduced and a portion of it is discussed, but then the discussion moves onto one or more different figures before later coming back to the initial one. An example of this is Figure 9: On page 9, line 17 Figure 9a is introduced, but then within the same paragraph the authors move on to Figure 10. The authors go back to Figure 9a, and eventually introduce Figure 11 before eventually mentioning Figure 9b. I recognize that this may just be personal preference, but I think that cases like these generally suggest that text or figures should be re-organized to better maintain the serial nature of the text.

I understand that in the final manuscript the authors will generally not have much of a say where figures will end up in relation to the text, but it is still a bit awkward for a reader to have to flip between multiple figures for a given piece of text (whether on paper or digitally). I wouldn't classify this as a major issue in this paper, but I still urge the authors to consider whether their figures or text could be restructured to make things flow more naturally.

We agree, kept these suggestions in mind, and did our best to improve the structure. We partly rearranged figures to avoid such a confusing back and forth in the discussion.

Specific Comments

P6, L17 and Figure 5: The information about the arrival height of 10km should be included in the figure. Also, what about the vertical information? How did these trajectories evolve in terms of altitude/pressure over this time period? Is this not important information for discussing the airmass origins?

We added the plot with the height information (lower panel of Figure 9, this was Figure 5 in the old version), and mention the arrival height of 10 km in the caption.

P10, L22: I do not believe the Lawrence et al ref here shows that the polar vortex collapsed on 20 April. In fact, their Figure 10 shows the vortex was unusally long-lived, lasting into May. Maps of the vortex on April 20 show that the vortex was undergoing a split, so it's probably fair to say the vortex began decaying around this time. However, if you click through further dates …., I think it's clear that a distinct polar vortex is present well into May. This comment is also relevant to other spots in the text where the authors mention the vortex collapsed in late April (e.g., P2, L11).

We improved the text accordingly.

P13, L14-15: This requires more nuance; it wasn't just weak planetary wave forcing, but also the very strong dynamical influence of downward wave reflection events during the 2019/2020 winter season (this is discussed in detail in the Lawrence et al ref).

We agree! But at the end we removed the long paragraph with all these meteorological statements to shorten the section. We did not see a need anymore to explain why the vortex was so strong.

P14, L3-9: While questions 2-3 are interesting and are worth further exploration, I'm not sure question 1 has much scientific merit in relation to the polar vortex. There are too many issues with timing and location and other confounding factors. Early in the season, the polar vortex was generally not extremely strong or weak. Furthermore, the very strong vortex conditions in Jan-Mar coincide closely with the dynamical downward wave reflection events mentioned above.

We removed question 1!

P14, L18-22: While true, I'm not convinced this is an appropriate apples-to-apples comparison. For instance, the 2019 Australian wildfires were severe enough to be comparable to a moderate volcanic eruption in terms of the impacts to solar radiation (as noted in the Khaykin 2020 paper already cited). I do not think that the Siberian wildfires have been established as being anywhere near as severe. There are other differences, but the point is, this kind of comparison runs the risk of equating inherently different events, which could be much more coincidental than this statement suggests.

We removed this part in Section 5. Only in the conclusion section 6, we left the remark concerning the ozone hole over Antarctica and the occurrence of the smoke at the same time.

P14, L23: How much of this introductory material is actually necessary for introducing the two figures of the section, which don't actually answer the "big questions" posed beforehand? The beginning of this section sets up big expectations that aren't met.

We re-arranged the text as already mentioned…. The 'big questions' are just briefly mentioned in the discussion on the potential pathways of the smoke impact on ozone depletion…

Figures 15 & 16: These figures appear to me to be mostly redundant. Figure 16 is mostly a coarse-grained version of Figure 15 with additional information about the ozone anomalies. It seems like this information could easily be combined into a single figure.

Yes, we prepared one figure but did not find a solution to 'efficiently' combine the information in both figures. But having one figure now, it is easier to see the effects. We indicate better the different months J: January, F: February, M: March, in both figures, (a) and (b) to facilitate the study of both figures.

P16, L16-22: This paragraph and Section 4 (which has major issues) seems like a missed opprtunity to provide expert guidance on how MOSAiC measurements could assist future studies that may attempt to answer such questions.

We think, Section 5 (the former Section 4) is sufficient to initiate, trigger, or stimulate further research. We are not experts for atmospheric chemistry, we are not members of the stratospheric ozone community, we are even not members of a well-organized stratospheric aerosol community. So we prefer not to stand up as a teacher to tell the pupil what to do next.

Answers and changes related to CC1

Dear Mike!

First of all thank you for your careful and detailed study of the manuscript and for the many, very 'direct' and constructive statements. We had your comments always in mind when we revised the manuscript. Before we answer your comments and questions step by step, we would like to briefly summarize what we did:

We went deeply into the literature with special focus on volcanic sulfate aerosol (originating from the recent Raikoke and from the Sarychev volcanic eruptions), and studied again many papers including the papers you suggested, we also re-checked our own Leipzig lidar data analysis, we again checked many CALIPSO observations and performed an extended HYSLPIT trajectory analysis and conducted further simulation studies (on self-lifting aspects) to bring together much more solid information and substantial argumentation to provide a plausible link between the Siberian fires (in July-August 2019) and the polar smoke layer we observed later on in the North Pole region (since late September 2019). The focus is on the potential self-lifting process, on smoke aging and on the consequences of slow self-lifting and thus efficient aging of the smoke particles for the morphological and optical properties of smoke. They become spherical before they reach the stratosphere and cause low depolarization ratios.

We did all this very carefully and even if we present a large number of reasonable and plausible arguments for our chain of hypotheses we emphasize that our entire set of arguments still remains at a hypothetical level.

Nevertheless, there is no doubt! What we observed during the MOSAiC expedition is clearly and unambiguously wildfire smoke. There is no way around. We observed this inverse spectral dependence of the lidar ratio (LR355 < LR532) together with the high LR532 around 85 sr, and that is a unique fingerprint of wildfire smoke! Note, that we observe such smoke fingerprints again and again, since 1998 (Wandinger et al., JGR, 2002). …. in Canadian, Siberian, Australian smoke, …

And one of the convincing and motivating arguments to search for a much more massive aerosol source than the Raikoke volcano was: The Raikoke-related AOD at high northern latitudes was about 0.01 at 500 nm in autumn 2019 (according to simulations of the 2019 Raikoke as well as of the 2009 Sarychev volcanic aerosol impact and corroborated also by Raikoke aerosol observations at lower latitudes, discussed in Kloss et al., 2021). But, we observed AODs around 0.1 in autumn 2019 in the High Arctic, and thus an order of magnitude higher values. We were forced to find a massive aerosol source.

Some significant changes before we start the step by step reply:

- We show a new figure (Fig. 2, time series, MODIS AOD from 2000-2020), highlighting the extraordinarily strong 2019 fire season in central and eastern Siberia.
- We introduce a new table (Table 2), highlighting the unique signature (fingerprint) of aged wildfire smoke, namely the inverse spectral slope of the lidar ratio (LR355nm < LR532nm) together with high LR532. We compare these smoke features in the new table with the ones for mineral dust, volcanic sulfate and ash, Arctic, European, American haze, etc.
- We show a new HYSPLIT plot (Fig. 5) to better explain the link between the CALIPSO overflight over Siberia and the strong fires… downwind (west!) of the flight track.

- We show a new figure with the Polarstern track (Fig. 6, same as in Engelmann et al., 2021).
- We created a new figure (Fig. 7) to better show the influence of the polar vortex on all the observations.
- We introduced a new section (Section 3) on Siberian fires, self-lifting hypothesis, consequences for aging (leading to spherical smoke particles), and also on the potential miss-classification of smoke layers as sulfate layers when using the CALIPSO aerosol typing scheme (Sect. 3.1).
- We shortened the ozone-depletion/smoke/PSC section as requested, and combined the two figures of this section to one figure (Sect. 5, Fig. 17).
- We skipped several figures (layer-depth histogram, monthly mean extinction profiles) to keep the paper as short as possible.

We think, more cannot be done. Many parts of the paper are substantially improved, caused by the recommendations of the reviewers. But this kind of research must also tolerate hypothetical explanations and argumentation. We feel this approach, as we present it now, is justified.

Step-by-step reply: our answers in BLUE

Ohneiser et al. (2021), hereafter abbreviated as "O21," is a provocative study of upper troposphere, lower stratosphere (UTLS) aerosols at high northern latitude between fall 2019 and spring 2020. It is provocative in that O21 observe UTLS aerosol nearly daily in that temporal span and conclude that the composition is wildfire smoke, the source is fires in a sector of Siberia in July 2019, and the transport pathway to the UTLS is diabatic heating/lofting.

The manuscript is also provocative in that Raikoke volcano (Kuril Islands) erupted in late June 2019 and polluted the UTLS with a mass of SO2 on par with or exceeding other eruptions that generated stratospheric clouds persisting for greater than a half year (E.g. Kasatochi, Sarychev Peak, Nabro, Kelut, Calbuco. See Solomon et al. (2016) for a tabulation.).

We checked the potential strength of the Raikoke sulfate aerosol in terms of AOD at 500 nm again, and we used especially the paper of Haywood et al (2010) for the Sarychev volcanic eruption (causing a very similar perturbation 10 years ago, in the summer of 2009), and came to the conclusion that the Raikoke-related AOD in October 2019 in the High Arctic was 0.01 to 0.015 at 500nm, but we observed 0.1, and thus an order of magnitude more, when we started the MOSAiC measurements! There is no doubt! We need a 'better' source for the aerosol we observed during MOSAiC. This is now clearly mentioned in the introduction. We must say, all the papers on Raikoke aerosol are a bit confusing because they provide observations of SO2 (yes this is related to Raikoke) but then they show extinction coefficients (and this is related to EVERYTHING, smoke, sulfate, background aerosol etc…). And we must say, in many papers the authors obviously just did not know what they measured. And the authors were probably miss-guided by CALIPSO lidar observations and CALIPSO aerosol typing. In the CALIPSO data base, you will find sulfate layers only (if we neglect the few pyroCb smoke cases). We (at TROPOS, Leipzig) compared our own Raman lidar observations (14 August 2019) with an almost direct over flight by CALIPSO (14 Augsut2019, same time): We saw the same layer features as CALIPSO, we measured the same depolarization ratios, they classified the layer as sulfate aerosol layer (based on depolarization observations), and our lidar ratio observations (LR355<LR532, and high LR532) clearly told us: WILDFIRE smoke. CALIPSO is a 'simple' backscatter lidar and we use an advanced, state-of-the-art Raman lidar, and at the end, we have to defend our unambiguous observations

(facts) and nobody asks: What may go wrong with the CALIPSO aerosol classification (assumptions) . Motivated by this mismatch, we introduce an extra subsection (Sect. 3.1) on the the miss-classification of smoke layers when using this quite simple CALIPSO aerosol typing scheme. Furthermore, we will write a paper on this topic (submission in September 2021).

If O21's conclusions are borne out, it will be a new insight into the polar UTLS and smoke transport to the UTLS. However, there is overwhelming observational evidence that Arctic UTLS in the second half of 2019 and early months of 2020 was blanketed by Raikoke sulfates.

The impression ….. 'overwhelming observational evidence' …. sounds strange in our ears. You mean you have rather clear arguments that all the aerosol we detected and measured in the stratosphere in 2019 … was Raikoke aerosol? We have clearly to state: We disagree! Our observations are in full disagreement with the 'overwhelming observation evidence'.

And concerning the 'evidence': In the Kloss et al. (2021) paper, for example, they show color-scaled OMPS-LP-derived AODs with high resolution up to 0.025 (ONLY) and all higher values are just given in the same color as for 0.025, simply in red. Why (?), because of saturation effects? This is a strange way of presentation, when keeping in mind that the AOD was obviously close to 0.1 in August to October over high northern latitudes (according to the Leipzig, Spitsbergen, and Polarstern lidar observations at 532 nm)! Furthermore, a rather inhomogenuous AOT distribution is shown by Kloss et al. with increasing AOT from low to high latitudes. This is in contradiction with model results which indicate a homogeneous Raikoke sulfate particle distribution from 45 to 90°N. The same was observed and modelled after the Sarychev eruption. Why was the volcanic sulfate aerosol this time (in 2019) inhomogeneously distributed from 45 to 90N according to the observations of Kloss et al., and after the Sarychev eruption it was homogeneously distributed, and thus in full agreement with the models? However, this inhomogeneous stratospheric aerosol distribution in 2019 is very reasonable when keeping the strong Siberian fires into consideration. But Kloss et al. (2021) totally ignored a potential contribution by the Siberian fire smoke.

Secondly, there is abundant evidence that the UTLS aerosol picture O21 describe over the July Siberia source sector is also dominated by Raikoke SO_2 and sulfates.

If the Raikoke sulfate aerosol was dominating (with an absolute maximum AOD of around 0.02 at 500 nm in mid August 2019 and later on with an AOD of the order of 0.01 in October 2019), how can we then measure AODs even exceeding 0.1? We observed stratospheric AOTs of 0.1 with lidar at Leipzig, at Spitsbergen, and at Polarstern in August to October 2019!!!! So, we clearly have a conflict, if we want to continue with our message: Raikoke aerosol dominated the stratospheric aerosol layer in 2019. The simple answer could be: All prediction of all the models are wrong, i.e., the Raikoke impact was totally underestimated. But then please tell us why the Raikoke effect on stratospheric AOD was about a factor of FIVE higher than the Sarychev volcanic AOD although the emitted SO_2 amount was just 20 % higher in comparison with the Sarychev SO_2 amount.

Thirdly, the lidar data
presented by O21 are more closely aligned with spherical sulfate droplets than smoke particles. These three points are elaborated on below.

Your phrasing suggests, stratospheric smoke particles are 'by law' nonspherical and thus produce significantly enhanced depolarization ratios. Yes, this is probably true for smoke particles

reaching the stratosphere by the fast pyroCB-lifting processes, then the depolarization ratio is significantly higher than zero. The emitted fires retained their nonspherical shape when reaching the stratosphere. After 3-6 months even these smoke particles would have developed a perfect spherical form by smoke aging processes and the depolarization ratio would be close to zero (as can be found in Baars et al., ACP, 2019). In contrast: if the smoke particles are lifted slowly into the stratosphere, by self-lifting processes, so that aging processes could be completed within the humid tropospheric environment (all this takes 2-4 days only) then the smoke particles have the chance to reach the stratosphere as spherical particles. And this seems to be the case here. The Raman lidar observations
of the lidar ratios at 355 and 532 nm clearly and unambiguously indicated SMOKE (LR355<LR532, LR532 very high), and at the same time the depolarization ratio was close to zero (clearly indicating spherical particles). But now, we have a problem with the CALIPSO aerosol typing scheme, because such a case of smoke self-lifting is not considered in the CALIPSO aerosol typing scheme. As a consequence, if particles are spherical they are assigned automatically as sulfate aerosol particles.

All this is now explained in detail in the new Section 3.

Point 1

Kloss et al. (2021) show that the Raikoke volcanic cloud dominated the high-latitude northern hemisphere from eruption through the early months of 2020. I consulted Chris Boone, ACE-FTS Project Scientist and co-Principal Investigator, to help qualify the 2019/20 UTLS plume further. ACE not only delivers aerosol extinction profiles but profiles of SO_2 as well. These were combined by Cameron et al. (2021) in an examination of several UTLS volcanic events including Raikoke; the results further qualify those of Kloss et al. and show the strong presence of Raikoke SO_2 and sulfates at high northern latitude in summer and fall 2019. In addition, ACE IR spectra have been used to identify smoke aerosol in connection with ACE Imager extinction profiles (Boone et al. (2020)). The same technique was applied to high-latitude northern hemisphere 2019/20 ACE data while invoking published sulfate IR spectra for comparison. The findings are summarized as follows. In July 2019, the lower stratosphere in the latitude region near 60 degrees north is stuffed with sulfate aerosols from the Raikoke eruption. …

We can confirm this. Our lidar observation at Leipzig (52N) on 23 July 2019 of 355 nm lidar ratios of 45 sr (typical value for freshly formed sulfate particles) and depolarization ratios at 355 and 532 nm close to zero suggest non-absorbing, spherical sulfate aerosol. Unfortunately, the volcanic AOD was too low, of the order of 0.005…, so that we could not determine 532 nm particle extinction and lidar ratio values to that time (before August 2019, when the stratospheric aerosol increased strongly).

….. Identification of the composition is accomplished by looking at the infrared spectrum of the aerosols and noting the coincident enhancement of SO_2 and ACE Imager aerosol extinction layers. In September/October 2019, there is a blanket of aerosols in the lower stratosphere in the latitude region near 80 degrees north. ….

Coincidence (SO2 and enhanced extinction) alone is not a convincing argument to us. Again, we confirm that there was a lower stratospheric aerosol layer around 80N in September and October 2019 (by our Spitsbergen and Polarstern lidars). And we can add, the stratospheric

AOD was close to 0.1 at 500nm, and thus a factor of 5-10 higher than the expected Raikoke AOD.

The blanket appears to be composed of Raikoke sulfate aerosols. …

…according to CALIPSO aerosol typing…, but not in agreement with our dual-wavelength Raman measurements of LR355<LR532 together with high LR532.

In February/March 2020, the aerosol blanket in the lower stratosphere near 80°N is still present. …

Yes, again, we can confirm this and the layer was still 10 km in depth.

…. SO_2 has decayed. However, the spectra associated with the Imager aerosol layers are consistent with sulfate. At no point did we find any evidence of biomass burning smoke playing a role in these stratospheric aerosols (Boone et al., 2020). Detailed support for these findings is available upon request.

We are not sure whether these spectra contain accurate information on the aerosol type. 'Evidence' and 'consistency' sound convincing, but is that sufficient to make solid statements on the aerosol type? Is our solid fingerprint (the measurement of an inverse spectral behavior of the lidar ratio as an unambiguous fingerprint for smoke) not much more 'convincing'?

Point 2

O21 hypothesize that their MOSAiC Arctic 2019/20 lidar signals are dominated by an impressive build-up of UTLS smoke that began between ~24-28 July 2019 in a zone within Siberia centered roughly at 60°N, 110°E (Figure 2). They present nadir satellite imagery combining fire hot-spot data, true-color imagery, and aerosol optical depth retrievals to show an intensification of burning and smoke concentration. They also display a single CALIOP curtain on 26 July to characterize the vertical structure and ascent of smoke layers (Figure 3). There are several concerns regarding Figures 2, 3 and their interpretation, listed below.

It is simply not possible to know from a single CALIOP curtain if ascent is taking place; there is insufficient information from such a quasi-instantaneous vertical snapshot. Additional information must be brought to bear.

Yes, sure!

There are technical issues with Figure 3 and its caption. The date of the CALIOP image is given as "2016." This is obviously an inconsequential typographical error.

We improved this.

But more substantially, the latitude, longitude coordinates labeled along the bottom are wrong.

We improved this as well. We apologize for all these mistakes!

The curtain displayed is actually situated west of the Figure 2 boxed focal point. It is west of the labeled coordinates by approximately 25° longitude. Compare Figure 3 with the actual orbital track and lidar data …

We now show in the new Figure 5 HYSPLIT backward trajectories, and indicate the CALIPSO flight track in this HYSPLIT map as a straight line and also show box 2 (with strong fires) defined in former Figure 2 (now Figure 3).

This is scientifically relevant for two reasons. One is that the vertical aerosol profile over the Siberia focus zone is unknown to the reader. Secondly, the aerosol profile on display is nominally upwind of the Siberia box, meaning that the history of those aerosols may be disconnected with processes occurring in the Siberia box.

The HYSPLIT backward trajectories in the new Figure 5 show that CALIPSO measured the smoke produced in box 2, because the track was downwind of box 2, and not upwind. Therefore, we are convinced that the CALIPSO backscatter curtain plot provides an impression of the vertical distribution of smoke originating from the fires in box 2.

But is that essential? The main reason to include the CALIPSO figure is to show a convincing example: Yes, there was a lot of smoke and the smoke was everywhere up to the tropopause and even above the tropopause.

Indeed, it appears from the CALIOP curtains linked above that the "Ascending smoke plumes (mostly in green) are visible up to the tropopause at 10-11 km height as well as in the lower stratosphere…" [Quoted from Figure 3 caption] are mostly assigned as sulfate by the CALIOP version 4 aerosol subtype algorithm.

CALIPSO had no choice or chance…, as we already discussed: The depolarization ratio was close to zero, and then there is only one solution: this is sulfate aerosol!

One can see from CALIOP curtains just upwind of the Figure 2 Siberia box, on days leading up to the O21 smoke AOD buildup, an assignment of sulfate subtype to "mostly in green" backscatter in the UTLS. Here is an example from 20 July …. Similar scenes are found each day thereafter leading up to the 24-28 July O21 period of focus . This multi-day, broad, high-latitude swath of aerosols primarily defined as "sulfate/other" conforms to the findings of Cameron et al. and our deeper investigation into ACE July 2019 profiles near 60^{o}N.

As stated already, we confirm this finding: Our July 2019 lidar profiles at Leipzig provide clear signatures for volcanic aerosols. This is not surprising because the extreme fires in Siberia accumulated between 19 July and 14 August (according to the very nice paper of Johnson et al, 2021), and caused a huge jump in the AOT (in the troposphere as well as in the stratosphere) NOT before the beginning of August 2019. By the way, even Johnson et al. saw a lot of structures in the stratosphere over their field site in Alberta in August (and later on), but unfortunately they restricted their discussion on the backscatter and ozone profile observations in the troposphere (this information is obtained by personal communication).

We were a bit disappointed when evaluating the Cameron et al. paper because the most interesting month (August 2019) is missing, probably because of missing observations because of saturation effects as a result of too high particle extinction coefficients.

It stands to reason then, that if Raikoke sulfates are blanketing high northern latitudes at that time (Kloss et al., Cameron et al., and our investigation), they would also be evident over the Figure 2 Siberia box before, during, and after the hypothesized smoke uplift. The CALIOP example below shows aerosol subtype findings consistent with the prior examples on 20 July over the Figure 2 box….

Again, CALIPSO aerosol typing in July is ok, the rest is not ok to our opinion.

Hence, if the July 2019 Siberia-zone smoke is the initial condition for the O21 lidar observations in fall and winter, the presentation surrounding Figure 2 and 3 is insufficient to make that case.

Yes, such a presentation is insufficient and remains insufficient. We agree that the text in the submitted version of the paper was misleading. We gave the impression that would be able to offer a solid explanation. To keep the answer short: The goal of the new Section 3 (with the discussion of Figures 2 and 3 in the old version) is now to offer a plausible way and a lot of reasonable arguments that the Siberian fires were most probably responsible for the smoke we observed during the MOSAiC expedition. However, we clearly state that all this is based on several hypotheses. As given now more precisely in the introduction, the motivation for the extended Section 3 on Siberian fires is simply: We observed a strong perturbation of the stratospheric aerosol layer by smoke and we realized that this perturbation is an order of magnitude stronger (in terms of AOD) than the Raikoke sulfate impact and we asked ourselves: What can cause such a strong impact on the stratospheric aerosol load? What was the source for this smoke? And the jump in stratospheric AOD toqrds 0.1 guided us to check the fire maps for July and August 2019.

Sure, there is no way at all to present a clear solution, that the smoke produced over Siberia was definitely (no doubt at all) the source of the smoke over the North Pole 3-6 months later on. Even, modelling and air mass transport analysis would not help. But we have the Spitsbergen observations from the beginning of August 2019 to January 2020, and thus a coherent link from the fire months (July and August) to October, and even an overlap (October to January) with the Polarstern lidar observations. All this is shown in Figure 12 in the revised manuscript.

Point 3

O21 acknowledge that the stratospheric aerosols detected by MOSAiC lidars were spherical according to aerosol depolarization measurements (Page 7, line 21). In principle, there seems to be less uncertainty as to the shape, composition, and depolarization of volcanic liquid sulfates in contrast to biomass burning smoke. Lidar measurements of tropospheric and stratospheric smoke converge on the idea that smoke is depolarizing (e.g. Burton et al., 2015; Fromm et al., 2008), suggesting some amount of asphericity. The first reported lidar observations of stratospheric smoke emphasized the unmistakable signal of depolarization (Siebert et al., 2000; Fromm et al., 2000). More recent stratospheric smoke papers, cited within O21, consistently report depolarization by smoke particles. Hence O21 would be establishing a new finding--stratospheric smoke with essentially no aerosol depolarization. Arguing for this peculiarity, O21 mention an aging process and a collapse of black carbon core. Neither process is described, and no publications are cited. Given the weight of evidence for the overarching presence of Raikoke UTLS liquid sulfates during the O21 reporting period and the MOSAiC depolarization results provided therein, the arguments for particle aging and black carbon core-shell collapse must be made more substantively.

This was another motivation to introduce the new Section 3. The first version of the text was six pages long to bring together all necessary information on particle aging, the influence of smoke aging on morphological properties and finally on the optical properties of smoke particles as given in Ansmann et al. (2021). But this would be too long in such a MOSAiC paper. So, we decided to provide a compact Section 3 only. All the published knowledge about stratospheric

smoke (or better UTLS smoke) and resulting depolarization ratios measureable with lidar is linked to pyroCb-related smoke lifting. And this fast lifting of smoke by pyroCb convection prohibits particle aging and the development of a spherical shape before the smoke particles reach the stratosphere. They keep there nonspherical shape and thus produce significant depolarization of laser light. This is the present status of knowledge, and this is considered in the CALIPSO aerosol typing scheme, and this is the reason that any stratospheric aerosol layer with low depolarization ratio is classified as sulfate aerosol. In the Burton et al (2015) paper the smoke reached the dry upper troposphere, where favorable conditions (dry, no condensable gases) were given to slow down the particle aging process so that the particles remained nonspherical.

But now (in the case of the extreme Siberian fires) we had to introduce the self-lifting process. Otherwise there would be no smoke in the stratosphere, because as you told us (and we checked that also), there was no pyroCb development. And self-lifting takes some days, and this time is sufficient to finalize the aging process. And at the end of any aging of smoke particles is the spherical shape, the evolution of a perfect spherical shell indicated by the low depolarization ratios. And as we show in the paper of Baars et al. (2019), even the stratospheric smoke is aging and reaches the spherical form, but this takes months… before the depolarization ratio goes down to zero again. By the way, we observed the same for the Australian smoke over Punta Arenas. In January and February 2020, the smoke depolarization ratio was significantly enhanced and then dropped immediately to zero in March-April 2020, at least at the lower heights of the stratosphere.

In conclusion, this is the first time that we have smoke in the stratosphere (and there is no doubt that is was smoke because of the invsere lidar-ratio spectral dependence) and that this smoke was not depolarizing. This is quite a very new aspect. And it is therefore not surprising that such a unique event is not considered in the CALIPSO aerosol typing scheme. But the other way around, it is impossible that the UTLS aerosol layer we observed consisted of sulfate particles. All the observed facts (lidar ratios, Angstroem exponent) would be in severe contradiction with this conclusion. And even the simple Mie models would be in trouble to produce lidar ratios of 55 sr at 355 nm and 85 sr at 532 nm for typical size distributions and sulfuric-acid water droplets.