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.