General remark:

The paper is well written and focusses on an interesting atmospheric topic: Long-range transport of pyrocumulonimbus-related Canadian fire smoke in the stratosphere observed over France in August 2017.

An original and unique measurement strategy is selected: Two lidars at Palaiseau and Lille and one mobile system that measured the smoke during a travel from Palaiseau to Lille.

The highlight is the retrieval of the spectrum of the particle linear depolarization ratio measured at three wavelengths. There is a strong decrease of the depolarization ratio from 25% in the UV down to less than 5% in the near IR.

However, the discussion of the results is partly confusing and must be improved.

Major revisions are necessary.

Detailed comments:

The title is misleading. The focus is on lidar observations (in France) of aerosol layers several kilometers above the tropopause, and not below the tropopause (upper troposphere). So: … smoke aerosols in the lower stratosphere …. would be correct. Furthermore, the title does not indicate where you made these observations: …. Long range transport of Canadian fires smoke towards Europe observed over northern France… , and not, for example, in eastern Asia or in the Arctic. Please improve!

P1: The abstract must be updated after the requested changes.

P2, L2: All abbreviations must be explained when they are mentioned for the first time (in the main text). The abstract is a stand-alone text, and does not count in this respect.

P2, L25. After the foregoing paragraph, a paragraph is missing in which the literature is reviewed which is already available regarding the Canadian smoke period in August 2017: Khaykin et al. (GRL, 2018), Ansmann et al. (ACP, 2018), Haarig et al. (ACP, 2018). If there are more, please mention them as well. Such a literature of foregoing work is (always) required in the Introduction. What is already done and thus known? What is our new contribution?

P3, L1…. : Again, please explain: LILAS, IPRAL, PLASMA, MAMS etc…

It is quite confusing that Section 2, that contains first observations and in-depth analysis of all the satellite observations of the smoke including CALIOP (space lidar) measurements, …. is then interrupted by a 'dry' lidar methodology section. It would be better to have the technical section first (as section 2, or as an Appendix) and then all the observations and discussions in the follow-up sections, continuously in sections 3, 4, 5…. without a break…

P3, L28: With the AOD of the smoke layer and the vertical extent of the layer (from the lidar observation), the layer mean particle extinction coefficient can be determined. These values should be presented as well.

P4, L13: Please explain EOS AM-1

P4, Section 2.4: Isn't that a similar study of OMPS UVAI maps as already presented by Khaykin et al. (2018, in the supplementary part)? Should be mentioned.

P5, L3: I see a clear and strong jump in the aerosol load from 11 to 13 August 2017 over the northern parts of western North America (Canada), in Figure 4. This is a clear and convincing indication that the pyrocumulonimbus cluster activities on 12 August were most probably responsible for these aerosol features. But this not explicitly written.

P5, L10: explain AIRS!

P5, L10-25: The same holds for Figure 5, a clear and sudden increase in the CO concentration from 11 to 13 August is visible. Furthermore, this 12 August event seems to be the reason for the extreme smoke load over Europe on 21-22 August (Figure 5f) as reported by Ansmann et al. (2018). This should be also mentioned.

P5, CALIPSO section 2.6: You already discuss the linear depolarization ratio, but the definition and explanations are provided later (in section 3). This is one of the reasons, why you better start with a more technical section right after the introduction.

P6, L3: Mixed-phase clouds can produce depolarization ratios from almost zero (mainly drops) to 40% (mainly ice crystals), so why should they show depol values of 26-35% only?

P6, L8: Why does an increasing trend in the depolarization ratio indicate aging of smoke particles? First of all, I do not see a trend in the noisy CALIPSO measurements. And second, the strong depolarization can be simply explained by the irregular shape of the smoke particles…, and why should the irregular shape change with time in the stratosphere where dry soot particles seem to dominate.

P6, section 3: Interrupt! A complex section is given, which is only interesting for lidar experts. But the paper is written for the research community dealing the atmospheric smoke and aerosols.

In Sect. 2.1, the lidars are explained (Raman lidars). And in Section3, we learn that you make only use of the Klett method. This makes quite a significant difference. The lidar ratios at 355 and 532 nm are not directly measured, you need to assume that the lidar ratio is height constant (and equal to the obtained layer mean value), so the solutions of the backscatter and extinction coefficients at a given wavelength are not independent of each other, the profiles are rather similar. And at the end, you use the non-independent extinction and backscatter solutions in the lidar inversion retrieval to obtain the microcphysical properties. This inversion algorithm however needs independently measured backscatter and extinction coefficients at 355, 532 and 1064 nm (backscatter). Please mention and explain this 'contradiction' in more detail. What are the consequences for the inversion uncertainties?

P7, L7: You calculate the signal means (within 500 m thick layers) at aerosol layer top and base. Please be precise. You probably use signals below the base and above the top only? What happens if there are traces of smoke above and below the layer, what uncertainty causes the use of noisy Rayleigh signals above and below the aerosol layer in the lidar ratio retrieval? All in all, be using the simple Klett method the uncertainty in the lidar ratio values (at 355 and 532 nm) must be at least 30%. All this needs to be clearly stated. We need overall uncertainties in the presented lidar ratios at the end of section 3.1.1.

The same for section 3.1.2: We need overall uncertainties in the depolarization values at the end of the section. If you have already 10% uncertainty in the volume depolarization ratio at 355 and 532 nm, the uncertainties in the particle linear depolarization ratios will be close to 20% for the particle depolarization ratio (especially at 355 nm). How trustworthy are such very large 355 nm particle depolarization ratios of 28% you present later on in the article? Even in the case of desert dust (rather irregularly shaped particles) the particle depolarization ratio is usually 25% or less. Please comment on this.

P9, Sect. 4: First question: The given standard deviations (in the text and in Table 1) indicate the uncertainty or the atmospheric variability (in time and in the vertical profile) or the uncertainty in the retrieval? It is not clear to me. Please state that clearly.

Table 1: I miss the information about the height (base, top) of the smoke layers in the stratosphere in the table. I would recommend to provide the layer-mean extinction coefficients as well.

P10, L15: If you have an uncertainty of 20% in the 355 nm backscatter value and 10% in the volume depolarization ratio then it is hard to get an uncertainty of 15% in the particle depol ratio. I ask this to force you to re-check the results concerning the very large 355 nm particle depolarization ratios of 28%.

P10, L29, Figure 8 and P11, L5: The only reliable profiles are the Klett solutions for the backscatter profiles. The extinction profiles and the EAE profiles are estimates and are strongly based on the used height-constant (layer mean) lidar ratios. It should be made very clear that the solutions of backscatter and extinction are so similar (or better identical in the profile characteristics) because of the use of the Klett method and the assumption of a height-constant lidar ratio.

In the discussion of the Angstrom values (AE) you may know that EAE = BAE +LRAE, as given in Ansmann et al. (JGR, 2002 on ACE-2 observations in Portugal). And that fits very well in your observational cases. However, if the assumed LRAE is height constant, and the BAE is height constant, EAE can only be height constant. But in reality, EAE and LRAE may vary. This should be mentioned. Without the Raman lidar approach you cannot provide information on the 'real world' EAE, extinction, and lidar ratio profiles. They are based on an assumption, and not on measured facts. So it is misleading to show profiles of extinction and EAE without saying that these are just retrieval products heavily controlled by the assumption of a height constant lidar ratio.

P10, L30: 'The UTLS aerosol layer was between 17 and 18 km'. This is one of the sentences that forced me to ask: Why do you call that an UTLS layers. The layer is clearly in the lower stratosphere.

P11, L13: Figure 10 is questionable and probably not representative to explain the long range transport including transport times. The CALIPSO Figure 6 tells us that the smoke layers were mostly at heights of 14-17 km over North America and the Atlantic, sometimes even below 14 km. And at these heights the wind speeds were much higher than at 17-19 km (in Figure 10). We need more trajectories! Especially for the CALIPSO heights (12-17 km) before we can make conclusions on the travel duration (in the discussion section 5). And one effect makes the use of trajectories difficult: The ascent of the layers from day to day by absorption of sunlight! This is not considered in any trajectory modeling. So trajectories are of limited use here. But 20 days of travel from Canada to Europe? That would be quite new, in view all the papers on Canadian smoke and long range transport. Even Khaykin et al. (2018)

writes that the smoke needed 21 days to travel around the globe (at midlatitudes) to be back over Canada again, and that the smoke needed only 10 days or less to reach Europe.

P11, section 4.2.2: Again the question: How trustworthy are the solutions of the inversion method when only based on Klett optical properties, so that the basic information are backscatter profiles? Please comment on that.

Discussion section:

P12, L26-28: I have my doubts that one singular event (the major pyrocumulonimbus cluster on 12 August) leads to so many smoke layer that you observed between 24-31 August. To my opinion, layers observed 10 days later maybe linked to the 12 August event, and maybe even layers observed on 24 August. But all other smoke layers later on were most probably triggered by other reasons and causes. And note that the absorbing smoke layers ascended during the travel (because of strong absorption of solar radiation). Khaykin et al. (2018) mentioned, 2-3 km per day during the first days. So layers can easily reach the lower stratosphere after 3 days when initially injected to 8 km height ….

P13, L1: In the discussion of the very large 355 nm particle depol values, please check: There is a corrigendum note to the Burton 2015 paper (you will find it on the ACP page). In this corrigendum note it is stated that the 355 nm depolarization ratios were only 20.5% for the smoke case (and not 24% as erroneously obtained in the beginning of the data analysis and published in Burton et al. (2015) paper). So, again, the 28% you get at 355 nm are really 'outstanding' and must be re-checked.… Find out, for example, how large the impact of the Klett backscatter profile is in the retrieval of the particle depol ratio at 355 nm… by using plus/minus 10% backscatter coefficient profiles.

P13, L5-L22: You provide the suggestion (Figure 11) that the depolarization ratio increases with travel time. This is surprising, because Nisantzi et al. (ACP, 2014, Injection of mineral dust into the free troposphere during fire events observed with polarization lidar at Limassol, Cyprus) find the exact opposite and show that in a figure: Decreasing depolarization ratio with transport time. However, their study is exclusively based on tropospheric smoke. And you combine tropospheric and stratospheric observations in your analysis.

Who is right? Maybe both! But first of all, one has to clearly distinguish aerosols in the troposphere and in the stratosphere. As mentioned, tropospheric as well as stratospheric smoke depolarization ratios are considered in your Figure 11. To my opinion, in this way you compare apples and oranges, and therefore the conclusions maybe wrong, at least are 'dangerous'. In the troposphere, coagulation and interaction with gases and moisture can take place, as you discuss so that aging leads to growing particles. According to measurements they get coated (Dahlkoetter et al, ACP 2014), they get coated with liquid stuff, so they get more and more spherical. And that influences the depolarization ratio. The depol ratio decreases, in accordance with the Haarig et al. (2018) results. In the stratosphere all this is less probable, and coating of soot particle is practically suppressed because the moisture and all the gases are not given in the stratosphere. It seems to be that the particles in the stratosphere remain unchanged during the travel. That is what I see also from the noisy CALIPSO observations in Figure 6 and what is also shown in the Haarig paper. But maybe I am wrong. Please provide us with a convincing explanation why aging leads to an increase in the depolarization ratio!

Disregarding my personal opinion, as a consequence of the statements above and the rather different particle aging processes in the troposphere and stratosphere…: If you want to show Figure 11 than you need to clearly indicate the stratospheric values, may be by very different symbols, e.g., by big open circles. We need a strong contrast between tropospheric and stratospheric values… The lower values in the Figure are obviously tropospheric depol values (apples) , and the higher ones are stratospheric values (oranges) . Furthermore I have my doubts about transport times of more than 14 days…. The related trajectories are always very uncertain and thus questionable.

Final point: I think the primary goal of the discussion section should focus on the comparison with other findings, especially with Haarig et al. (2018), and also Burton et al. (2015, upper tropospheric smoke case). As mentioned several times, your results are based (to a large fraction) on the use of the Klett method. And this includes the determination of the particle depolarization ratios because the Klett backscatter profiles are needed. This can be a significant reason for deviations, especially in the case of the products at 355nm.

P13, L23: Regarding the laboratory studies: Miffre et al (2016) used Arizona Test Dust (ATD). This dust is not a natural dust component. It is 'manufactored' by a company and ATD particles have rather sharp edges. As a result, ATD shows larger depol ratios then usual desert dust particles.

Regarding the Jaervinen et al. (2016) study: These authors made their observations by one wavelength (488 nm, and partly at 552 nm) only! ..and not at 355 and 532 nm, as you mention it. But they showed their observations as a function of the size parameter (size mode diameter /wavelength) so that Mamouri and Ansmann (2016) estimated the probable depolarization ratios for 355, 532, and and 1064 nm. The results (depol ratio for fine dust at 355, 532, and 1064 nm derived from the Jaervinen study) are shown in Haarig et al. (2018). Almost all your discussion information on page 13 (L23-34) was discussed in the same way in the papers of Mamouri and Ansmann (2014, 2016). So, please add these papers to your references.

P14, L37: Regarding the effective radius (from the lidar inversion retrieval), please compare with Haarig et al. (2018). Is there agreement? Please discuss the comparison.

P14, L39: Again, please avoid mixing of tropospheric and stratospheric effects of potential particle growth. Müller et al (2007b) focusses on tropospheric smoke only when discussing aging effects.

P15, conclusion section: The results of the paper are summarized only in this section. But concluding remarks are not given. So maybe, create some concluding remarks, some outlook ideas…