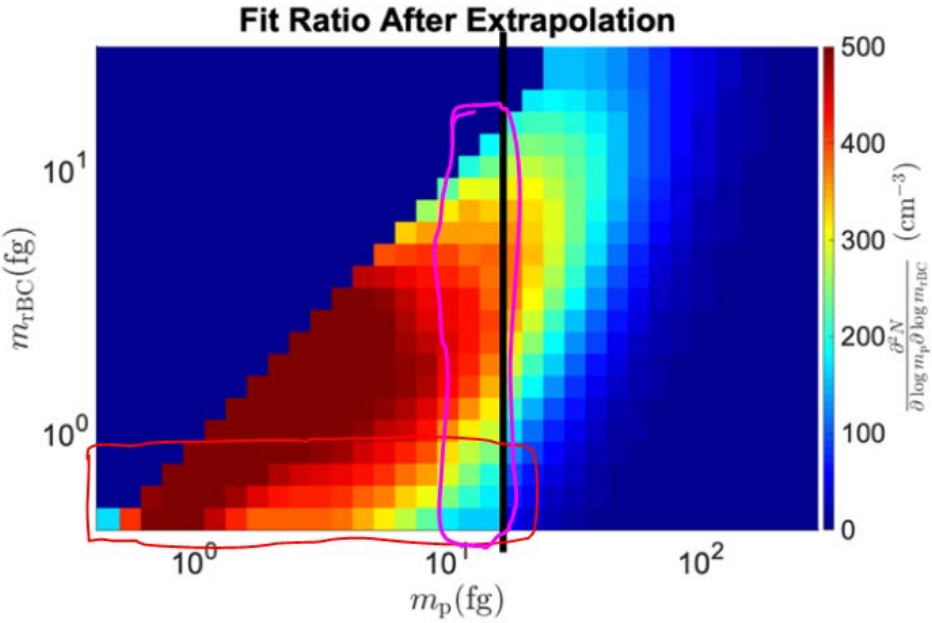
General remarks:

The introduction provides a good overview on reasons why a better understanding of the mixing state of atmospheric black carbon (BC) particles with other particulate matter is required when it comes to e.g. their climate impacts through aerosol-radiation and aerosol-cloud interactions. This prepares the ground for the single main topic of this study.
In a recent lab study, the combination of a coquette particle mass analyzer (CPMA) and a single particle soot photometer (SP2) was shown to be a useful and reliable means to quantify the 2-dimensional mixing state distribution as a function of total particle mass and BC core mass, which also works for complex BC aerosols with high mixing state diversity on a single particle level. Furthermore, such measurements were not accessible with comparable data coverage and accuracy to other previously applied experimental methods. The study here takes the combined CPMA-SP2 approach for the first time to the field, specifically to an air pollution hotspot where the environmental impacts of aerosols are of great interest. This certainly makes the resulting data set on BC mixing state and mixing state diversity on single particle level worth publishing.
Unfortunately, the manuscript was hastily prepared and not really at the level expected for a high quality scientific journal. Reasons for this judgment are given in below comments and include:

- Very imprecise and sometimes wrong statements concerning the mixing state index "chi", which is the very central topic of this study.
- Insufficient sensitivity analyses and discussion on how the results for the mixing state index "chi" may be affected by the detection limits of the applied instruments.
- Week discussion and interpretation of the results.
- Poor language at many instances, definitely below the level that can be expected from joint forces of the whole author list.

Major comments:

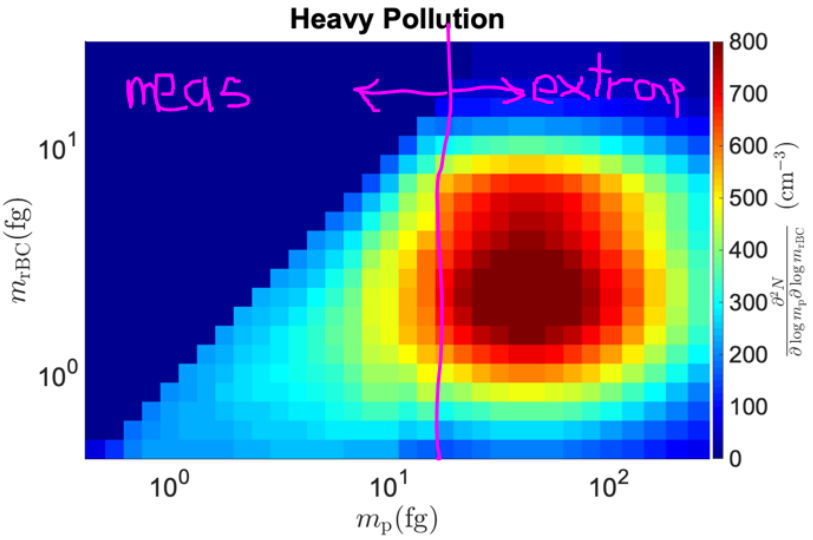
1. P6, L24: I am puzzled about the statement "In this study, because only the rBC-containing particles are detected, it is not possible to detect a completely external mixture." – I would expect that the formalism was spitting out chi=0 (indicating completely external mixture for the particle types under consideration), if the coupled CPMA-SP2 measurements would provide m_rBC = m_p for every single particle (or more precisely: when calculating the limes towards this mixing state as the equations contain multiplication of "zero times infinity" or division of "zero by zero" for a perfectly externally mixed population).
2. Truncation of the BC particle mixing state 2D-PDF due to detection limits of the applied instruments is often unavoidable, and so it is in this study. To my judgement, the authors have done a fair job in addressing the impact on the resulting mixing state parameter from the upper limit of measured total particle mass (at least for 3 out of the 4 examples; separate comment on the latter follows). However, they have not at all addressed the impact of the SP2 lower detection limit for rBC mass in a single particle. The 2D-PDFs very nicely illustrate that for total particle mass equal to the SP2 lower detection limit only externally mixed BC particles can be measured, while internally mixed BC with, consequently, smaller BC core remains undetected. From a pessimistic perspective, the nearly complete mixing state information for BC containing particles is only available for either a narrow range of BC core sizes (roughly indicated by red square), or for a narrow range of total particle mass (roughly indicated by the magenta square).



How does the overall mixing state diversity parameter compare with its value calculated for the above two constrained size ranges? Does this hamper absolute interpretation of the chi values or is this only a minor issue?

Personally, I do not really have a feeling for the sensitivity of chi to such truncation effects. However, the fact that chi appears to be weighted by particle number and that particle number is dominated by small particles in the range of the lower SP2 cut-off would suggest that truncation at the lower SP2 cut-off could have a much more dramatic effect on the resulting chi value than the truncation at the upper CPMA cut-off in many cases (the heavy pollution period is obviously the counter-example).

3. Taking the previous comment a step further would raise the question whether a number-weighted parameter is suitable when it comes to e.g. BC lensing effect, where we care in first order approximation about the BC core size range where the BC mass size distribution peaks. Would it also be useful to additionally compute the particle diversity diameter exclusively for a narrowish range of rBC core mass around the modal size of the rBC mass size distribution (which will require some extrapolation above the upper CPMA cut)? This might also provide a mixing state information that is more readily comparable across different studies with similar instruments as they only have to cover this range well, whereas the actual lower/upper instrument cut-offs would loose out in relevance.
4. The extrapolation of the mixing state PDF looks reasonable for the example shown in the SI (and likely for the low/medium pollution winter and the summer cases). However, showing the heavy pollution period mixing state results without indicating the measured/extrapolated parts of the 2D-PDFs is really misguiding the reader:



While the measured part still provides valid information on the fact the BC is generally highly internally mixed, the mixing state diversity parameter chi, which is vastly dominated by the extrapolated part of the 2D-PDF, can hardly be meaningful in absolute terms.

5. P10, L6-13: While not being wrong, the discussion in this paragraph is rather weak. In the end the observed mixing state is a result of various factors including mixing state at emission, average atmospheric residence time of the BC particles, representative aging rate (in the sense of coating acquisition rate), etc.
    a. For the high pollution winter case the authors simply say: high secondary aerosol production → more condensation → more rapid coating acquisition → higher MR. This is incomplete.
    b. For the low pollution winter case the authors simply say: observed thin coatings may indicate fresh emission. This is also incomplete.
    c. The statement "due to the presence of large numbers of large rBC containing particles with their m_{rBC} between 1 fg and 10 fg, the bulk MR value increased rapidly" does not really make sense (BC core size alone cannot explain anything about MR).
    d. Further down (P10, L30) the following statement is made for the high pollution period: "This large fraction of thickly coated rBC-containing particles was expected due to the coal burning in winter." – How does this relate to above arguments and where the coal burning BC mixing state knowledge come from?
6. Presentation of results shown in Fig. 9 (last paragraph of Sect. 4.3): The potential evolution of MR during aging after emission has strong implications on which type of conclusions can or cannot be drawn. This is not properly acknowledged in the current discussion which is, therefore, flawed.
7. P12, L16: The following statement is made here: "This means although the rBC mass loading was low in summer however a more homogenous distribution of coatings or internal mixing state was present." – What is the expected process that causally links BC mass concentration with BC mixing state homogeneity? (Or do you have cross-correlations between BC mass concentration and other parameters that causally relate to BC mixing state diversity in mind?). The point raised in this comment also questions the meaningfulness of Fig. 12c.

Minor comments:

8. P1, L25 (abstract): The "or" in the following statement sounds confusing: "χ of rBC-containing particles was highly positively associated with increased bulk MR, rBC mass loading or pollution level in winter,…"

9. P1, L29 (abstract): The following statement is difficult to understand and hardly self-explaining without having read the manuscript: “The same level of bulk MR corresponded with a higher χ in summer than in winter and this tended to suggest a limited formation of coatings on rBC largely depended on primary sources.”
10. P1, L32 (abstract): What is the mechanism that links ambient temperature with BC coating thickness?
11. P1, L34 (abstract): The statement “The mixing state of rBC-containing particles should also depend on the coating formation mechanism, both primary source influence and secondary coating formation mechanism should be considered in interpreting the rBCcontaining particles mixing state in the atmosphere.” should rather be made early in the abstract, i.e. before providing interpretations of observed BC mixing state degree and variations.
12. P2, L18-21: the Shiraiwa 2007 study cited here in the context of relationship between BC mixing state and its droplet activation does not provide any activation measurements. Schroder et al. (2015) or Motos et al. (2019) are examples of recent studies which directly investigated the activation of BC particles in ambient clouds as a function of their size and mixing state.
13. Sect. 2.4 on CPMA-SP2 coupled system: Are small uncharged particles passing the CPMA in addition to the charged and mass selected particles an issue (upper cut-off size for uncharged particles depends on rotational speed)?
14. P8, L12-14: Measurements of major chemical components of atmospheric aerosols rarely provide “zero”. Instead, concentrations can drop by several orders in magnitude or below the lower limit of quantification. Therefore it is not very useful to report observed value ranges as “from zero to ….”.
15. Figure 3: The period marked as “light pollution” looks like what is described as “clean” in the first paragraph of Sect. 4.1. Actually, a few lines further down (P.8, L20) I came across the definition “light pollution means NR-PM1 < 100 ug/m3”, which is not congruent with the shading shown in Fig. 3.
16. P9, L5: Here it says: “The volume equivalent diameter for rBC-containing particles (Dp) and the rBC core (Dc) is also calculated by assuming a density of 1.2 g/cm3 for rBC-containing particles and a density of 1.8 g/cm3 for rBC cores.” – Strictly speaking, different densities should be assumed for core and coating then it comes to estimating Dp from mp. Furthermore, an average density of 1.2 g/cm3, which is lower than that of BC, inorganic salts and likely a fair fraction of the organics, appears to be a bit low. However, doing it more precisely would likely have little effect on the resulting Dp values.
17. P9, L9-11: The shift of the rBC core size distribution to larger sizes appears to be quite dramatic (it would definitely look dramatic when shown as mass rather than number size distribution).
    a. Are these rBC core size distributions based on the polydisperse SP2 measurements or on the (extrapolation dominated) coupled CPMA-SP2 measurements? I would only trust the latter for the high pollution period (unless that one was jeopardized by coincidence issues).
    b. How do these findings on rBC core size distribution compare with existing literature (if any)?
18. Sect. 4.3: The claim is made here that the CPMA-SP2 combination more accurately provides the MR than SP2 only data (combining incandescence signals with LEO-fit based optical sizing). This claim seems fully justified. A quite interesting question would be: how comparable are the mixing state parameters chi inferred using these two alternative methods (for a subset of the 2D-PDF accessible to both methods)? – You would have the data set to look into this if you wish to do so.
19. P10, L1: The statement “Where Ni is the number concentration for rBC-containing particle i” does not really make sense.

20. Figure 6: In the second to last minor comment made just above I suggested to compare the chi values resulting from the coupled CPMA-SP2 and SP2 only approaches. It is great to see that a similar comparison is done for the bulk MR in Fig. 6 and associated discussion. However, no information is provided whether the m_p and m_{rBC} ranges considered for this were constrained to the jointly accessible ranges or whether the MR from the two approaches are based on different m_p and m_{rBC} ranges. This kind of information is required to put the level of (dis-)agreement between the two in proper context. The same remark applies to Figs. 7 & 8.
21. P12, L10: The statement "…while the bulk MR metric cannot be used to predict the rBC mixing state in summer…" should probably read "…to predict the BC mixing state diversity index…". (bulk MR is a metric for BC mixing state)
22. P12, L12: The statement "…the rBC-containing particles were still well mixed." should probably rather be about the "mixing state diversity".
23. P12, L18: "…with the source apportionment study that about 64% of rBC-containing particles were from primary sources in winter experimental period…" – What are secondary sources of BC-containing particles!?
24. P12, L20: "…rBC-containing particles may have coatings externally mixed with the rBC cores…" – What is an externally mixed coating!?
25. P12, L26: "In summer, the primary emissions are less compared to winter, and SOA contributes more in Beijing" – In absolute or relative terms?
26. P12, L30: "However, this does not mean the amount of coatings on BC will be necessarily higher in summer because the higher temperatures and enhanced dilution may cause the primary semi-volatiles to favour the gas phase." – What would you conclude based on your study: is the fractional contribution of primary non- and semi-volatiles ever substantial when MR is high, be it winter or summer? Do we at all care about the phase partitioning of primary semi-volatiles when it comes to BC mixing state?
27. P13, L1: "The results here suggest that the increase of coating content such as the bulk MR above 4 could importantly increase the internal mixing state of rBC-containing particles" – Please choose a more precise wording. Do you rather refer to the mixing state diversity?
28. Section 5: The general importance of know BC mixing state including mixing state diversity on single particle level is motivated. This study provides likely unprecedented mixing state diversity information for BC particles. However, the important question whether the observed diversity is at a level where errors made in different applications by assuming an average mixing state are small or large is unfortunately not address. This would be desirable, though possibly a topic to be follow-up in future studies. As is, the current Section 5 could just as well be moved to the introduction as motivation for doing such measurements.
29. P14, L12: "…and makes the bulk MR may act as a predictor of mixing state…"
    a. Requires language editing.
    b. Do you mean: "…a predictor of mixing state diversity"?
30. P14, L15: "The slightly higher χ in summer indicates that internal mixing is preferred for rBC particles…" – Higher chi – for BC particles only – does not imply internal mixing of BC. Instead it says something about the particle-to-particle variability of MR.
31. Supplement, Fig. S1: it would be instructive to point out at least the obviously multiply charged particles in some way. In fact, having a full inversion scheme at hand would very easily allow to split the measured 2D-PDF into contributions from singly, doubly and triply charged particles.

Technical comments:

32. "Coating thickness" should be explicitly defined as it is a purely operational definition rather than being meaningful from a morphology point of view. Specifically for this study: "coating thickness" refers to a "concentric spheres equivalent coating thickness" derived from direct

measurements of total particle and BC core mass (and assumed BC and coating densities). Or would it actually be possible to largely omit the term "coating thickness"? "BC mixing state" would for example seem more appropriate in the subsection title 4.3 and the first sentence of this subsection because it is all about the "BC mixing state expressed with the non-BC to BC mass ratio of the BC-containing particles".

33. P1, L12 (abstract): repetition of "in the atmosphere"
34. P1, L30 (abstract): requires language editing (missing "to"; adjective vs adverb)
35. P2, L2 (abstract): Starting the sentence on this line with "This…" does not appear to be appropriate.
36. P2, L6: Should better read: "…component of atmospheric particulate matter and…"
37. P3, L13: "a polluted"
38. P4, L1: Please provide the year when the field experiments were done.
39. P4, L26-29: Please fix this sentence.
40. P5, L28: Please fix this sentence.
41. P5, L30: Using "N_{BC}" instead of solely "N", would be more precise as the 2D-PDF exclusively includes BC-containing particles, whereas BC-free particles are not considered.
42. P6, L3-4: It should probably read "dlogm_p" and dlogm_{rBC}", however, I am not a mathematician.
43. P6, L5-6: A more precise wording would include "inverted" and "retrieved from the measured…"
44. P6, L13-15: Please fix this sentence.
45. P8, L4: "approximate" instead of "appropriate"?
46. Figure 7 (and several of the following time series graphs): It would be helpful to indicate the different periods ("heavy pollution", etc.).
47. P9, L9-10: Please fix this sentence.
48. The caption of Figure 9 appears to be a revolutionary evolution.
49. Figure 9: What is included in the "MR=0" class? Everything with MR less than a small but non-zero value? Or in other words: Finite CPMA transfer function and uncertainties of CPMA and SP2 make it impossibly to accurately quantify MR=0 (even if such particles were truly existing). How is this dealt with?
50. P11, L4: The sentence "As the CPMA-SP2 system only detects the rBC-containing particles and the number of species set here is 2, i.e. rBC and non-rBC material." is incomplete. Please fix.
51. P14, L10: "…, which illustrated there was more internal mixing during this period…" is obsolete as the complete and more precise statements follow in the next sentence.
52. Supplement, Eq. 9: the superscript for the parameter $p_i^?$ appears to be missing.

References:

Schroder, J. C., Hanna, S. J., Modini, R. L., Corrigan, A. L., Kreidenwies, S. M., Macdonald, A. M., Noone, K. J., Russell, L. M., Leaitch, W. R., and Bertram, A. K.: Size-resolved observations of refractory black carbon particles in cloud droplets at a marine boundary layer site, Atmos. Chem. Phys., 15, 1367–1383, https://doi.org/10.5194/acp-15-1367-2015, 2015.

Motos, G., Schmale, J., Corbin, J. C., Zanatta, M., Baltensperger, U., and Gysel-Beer, M.: Droplet activation behaviour of atmospheric black carbon particles in fog as a function of their size and mixing state, Atmos. Chem. Phys., 19, 2183–2207, https://doi.org/10.5194/acp-19-2183-2019, 2019.