We thank Reviewer 1 for the constructive comments and suggestions. We have addressed all comments in our response and feel these revisions improve the manuscript. The reviewer comments are copied here in italics, our response is in plain text, and revised statements from the manuscript are underlined. We also submit a "marked up" version of the revised manuscript.

Reviewer 1

In this paper, the authors present measurements and modeling of BC mixing state in the springtime Canadian high Arctic. Measurements were collected using Single Particle Soot Photometer (SP2) and Ultra-High Sensitivity Aerosol Spectrometer (UHSAS). The authors reported that measurements of aerosol mixing state provide important constraints for model estimates of direct radiative effect. The dataset and the associated analysis and modeling results are valuable to atmospheric and environmental researchers and the topic is fitting well within the scope of ACP. I suggest some revisions to improve the clarity and scientific merit of the current manuscript, after which I recommend publication.

Specific comments:

1. Particles measured by SP2 and UHSAS were combined to determine the number fraction of rBC particle. These two instruments have different size-cut and different measurement sensitivity. It would be worthwhile to include some details on the influence of the instrumental size-cut on the overall results, instrumental inter-calibration and any size adjustment for combining data from two different instruments. For example, if we have a 100 nm particle (mobility or some sort of equivalent size), do the SP2 and UHSAS both measure it as the same size. If not, how do the authors combine the measurements from two instruments to construct the distribution of the number fraction of rBC particles as a function of size? For the size distribution shown in Fig. 4A, how much are we missing below the measurement window (100 nm)? Is it possible to get a closure in Fig4A (i.e., total particles = bare-BC+ coatedBC+ non-BC)? They have some limited discussions on these, but it's hard to follow and get the whole picture. Details would be helpful for general readers. Details can go to the SI.

Regarding instrumental inter-calibration and any size adjustment for combining data from two different instruments, both the SP2 and the UHSAS measure optical size using the scattered light intensity from an IR laser (1064nm light for the SP2 and 1054nm light for the UHSAS). To convert the scattering signal to a physical size, the refractive index of the particle is needed and the two instruments differ slightly in the refractive index used. To address the referee's comments, we have added additional information on the uncertainties in converting optical measurements to size in Sections 2.2.2 and 2.3.

Regarding the size cut, the detection ranges of the two instruments are given in Section 2.2.1 (SP2) and in Section 2.3 (UHSAS). Very few particles had rBC core diameters larger than 650nm, so the mismatch in detection range for large sizes is unlikely to be significant. If there were particles with rBC cores smaller than the lower detection limit of the SP2 and very thick coatings, these would end up in the size range of the UHSAS but not be accounted for here as “rBC containing”. However, we have tested a bounding case (See supplemental Figure S1) where small rBC components (including those below 75 nm) were assumed to have very thick coatings (double the median value). In this case, no substantial difference in the DRE was observed compared to the median coating-thickness case (see Supplemental section).

Revisions to the main text:
In Section 2.2.2: In the Mie model we used a refractive index of 2.26–1.26i for rBC (Moteki et al., 2010), which is appropriate for rBC from fossil fuel combustion. Other studies have used this value in analyses of particles from urban areas (Laborde et al. 2013) and from biomass burning (Taylor et al. 2014). In the Mie model we used a refractive index of 1.5-0.0i for the coating species. The value of 1.5-0.0i is appropriate for dry sulfate and sodium chloride (Schwarz et al., 2008a, 2008b). Previous studies have shown that varying the refractive index used for the rBC coating material from 1.45 to 1.59 had a minor impact on coating thickness (Taylor et al. 2014a).

In Section 2.3: Further details of the instrument operating principles can be found in Cai et al (2008). The UHSAS was calibrated using polystyrene latex spheres (PSLs), which have a refractive index of n=1.59. Kupc et al. (2018) investigated the size dependence of UHSAS measurements on the assumed refractive index of the particles being measured. For particles with diameters smaller than 600 nm and real refractive indices of n= 1.44–1.58, the diameter measured by the UHSAS varied by +4/−10% (relative to the refractive index of ammonium sulfate, n=1.52). For this study, a shift in the UHSAS size distribution by 10% would change the fraction of particles containing rBC by less than 3% which has a negligible impact on the DRE calculations.

In Section 3.4: Fourth, our measurements of BC mixing state only apply to BC particles in a limited size range. The size distributions in Figure 4 imply a substantial fraction of BC and non-BC number concentration exist outside the size range of our measurements. We attempt to account for this limitation in our model analysis through consideration of upper and lower bounds on coating and BC-containing particles.

2. The discussion in Sec. 3.1 is limited. The authors should consider expanding the discussion in this section. There is no mention of Fig. 4B. What’s the implication of fitted line in Fig4B? Can the authors propose any parameterization based on this fitting in Fig.4B, which can be used for constraining/evaluating model results in a similar environment where there are no measurements? The campaign-average mixing state is mostly focused in this section. Were there any changes in mixing states depending on air mass trajectory (relatively-fresh vs. aged) in that very clean environment?

We have added several statements expanding on the measurements of BC mixing state during the POLAR6 campaign. We have copied these below, but first discuss the specific points brought up by the reviewer.

Thank you for catching the lack of reference to Figure 4b. We had mistakenly referenced a Figure 3b (which does not exist). We have corrected this error.

A parameterization based on the fit line in Figure 4b (or perhaps the results in Figure 3) is a really interesting idea; however, we feel it is beyond the scope of the current manuscript. The main reason is that it is not clear from the observations what underlying microphysical processes may control the fraction of BC containing particles (e.g., coagulation time scale, condensation rate, cloud processing). While it does not explitictly include physical processes, the fit line in Figure 4b is a polynomial fit that best matches the observations. As this may be helpful for future studies we have included the form of the fit in the caption to Figure 4b. In addition, we have added a couple sentences discussing the fitted line:
The solid line in Figure 4b is the best fit to the experimental data shown in Figure 4b. A polynomial function was used because this gave a good fit to the experimental data. The coefficients for this fit are provided in the caption to Figure 4; however, we note this fit is derived only from measurements in the springtime Arctic and is not based on any underlying understanding of the chemical or physical processes involved. With this polynomial fit, we…

And in the caption to Figure 4b:
The solid black line is a polynomial fit to the observations following the form (with x in nm): $y = 1.25 – (1.52*10^{-2})*x + (8.05*10^{-5})*x^{2} – (2.00*10^{-7})*x^{3} +(2.31*10^{-10})*x^{4} – (1.00*10^{-13})*x^{5}$.

Overall, we did not find substantial variability in observed mixing state parameters across flights. This is likely due to relatively long transport (~days) of most aerosol sampled during the campaign. We have added the following statements:
Across all flights (and altitudes), we do not find substantial variability in measured coating thickness. This lack of variability is likely due to atmospheric processing over several days during transport of the air mass to the Arctic region. In the SI, we plot coating thickness for each flight in Figure S3. The low degree of variability across flights justifies the use of campaign averages in our model analysis.

*3. Details measurements of aerosol properties are essential to improve model predictions and provide better constraints on the model results. However, based on the discussions in Sec. 3.3., it is not clear how much constraints are getting added by the detailed size-resolved BC mixing state measurements. It is clear that fully-externally or fully-internally mixed assumptions are not very realistic ones. However, external*1.5 bounding case vs. two constrained cases (fBC-*

*constrained and rshell-constrained) showed a similar level of uncertainty on the estimated DRE. For example, In page 12, L-26-28: "the fBC-constrained DRE is 0.3-0.4 W m-2 more positive than the external*1.5 mixing-state assumption, while the rshell constrained mixing state is 0.3-0.4 W m-2 more negative". Here, the two constraints cases provide two different results. If we don't know which measured case is a better representative one, then the details measurements are not adding that much of additional values compared to some bounding case. It would be worthwhile to include some discussions on this and current limitations and future directions that should be focused more.*

This is an important point to make in our paper, and we have added several statements explicitly addressing this comment. While we are unable to fully constrain the model to observations of BC mixing state, we are able to reduce the range in estimates of the DRE due to BC mixing-state assumptions using these measurement constraints. For instance, the range of estimated DRE of the pan-Arctic mean using the fully internal/external bounding assumptions is 0.31 W m^{-2}. Using the measurement constraints, this range is reduced to 0.14 W m^{-2}. This reduction of a factor of two in the range of the DRE due to uncertainty in the mixing state of BC is an improvement over only using bounding mixing-state assumptions. The remaining range of 0.14 W m^{-2} is due to the interaction of mixing state with the uncertainty in the ratio of scattering to BC mass. Thus, this study demonstrates that the uncertainty range in DRE can be reduced by only constraining BC mixing state (and not aerosol mass concentrations); however, this approach is limited to the representation of the scattering-to-BC mass ratio. If the model does simulate the observed scattering-to-BC mass ratio, than the uncertainty range in DRE would be further reduced.

To clarify this point, we have added the following statements:

In the abstract:

We find that the pan-Arctic average springtime DRE ranges from -1.65 W m^{-2} to -1.34 W m^{-2} when assuming entirely externally or internally mixed BC. This range in DRE is reduced by over a factor of two (-1.59 to -1.45 W m^{-2}) when using the observed mixing-state constraints.

In Section 3.3:

The range in the pan-Arctic mean DRE using the bounding mixing-state assumptions (*external, allCoreShell*) is 0.3 W m^{-2} (or about 20%). Conversely, using the measurement constraints on BC mixing state (r_{shell}*-constrained, fBC-constrained*) results in a range of pan-Arctic mean DRE of 0.14 W m^{-2} (or about 9%). Thus, using the measurement constraints on BC mixing state reduces the range in DRE from assumptions of BC mixing state by a factor of two.

In Section 4:

The estimated range of the DRE using the measurement-constrained mixing states (-1.45 to -1.59 W m^{-2}) is approximately a factor of two less than the estimated range when using the bounding mixing-state assumptions (-1.34 to -1.65 W m^{-2}).

If the BC mass fraction in TOMAS does match the observations, both mixing state parameters (shell thickness and fraction of BC-containing particles) can be constrained simultaneously and the DRE in both cases would be identical.

*4. The denominator of Eq. 1 [(r_core3)/(r_shell3+r_core3)] is confusing. If I understand correctly, to get the total volume of a coated-BC particle, we need to add core volume (r_core3) and shell volume (r_total3 - r_core3), where r_total= r_core+2*r_shell. In that case, [r_shell3+r_core3] would not provide the total volume of a coted-BC particle.*

This is not quite accurate (though we appreciate where the confusion arises from). We are using the term r_shell to refer to the shell thickness (the distance from the outer edge of the core to the outer edge of the particle). The total radius of a coated particle is equal to simply the core radius plus the shell radius (r_total = r_core + r_shell) and not twice the shell radius. In other parts of the paper, we discuss the diameter of the coated particle, which is calculated as dp_total = d_core + 2*r_shell. The "2*r_shell" term is used here as the SP2 measures shell radius (often referred to as "shell thickness"). In addition, we feel discussing "shell diameter" may be misleading as it may not be clear to a reader if this term does or does not include the core diameter.