We thank Reviewer 2 for the positive assessment and constructive suggestions. In response, we have strengthened the physical interpretation and broader context of our work, added a comparison between hybrid BNN calculations and those using the Rayleigh–Debye–Gans approximation, clarified the uncertainty and calibration methods, and expanded the discussion of relevant measurements of optical properties of BC-containing particles and potential applications of the hybrid BNN. Our response to each comment is provided below in blue text, while reviewer comments are reproduced in black.

---

## Reviewer 2 Assessment:

The paper describes a new hybrid Bayesian neural network model trained on discrete dipole approximation (DDA) simulations to better simulate the optical properties of particles containing black carbon. The authors compare the performance of the model with homogeneous and coated-sphere Mie theory and find that the Bayesian NN model performs better when simulating several optical parameters. One of the particularly interesting aspects of the study is the ability of the model to also determine uncertainties and therefore understand which input parameter space in DDA should be explored further.

The paper is well written and mostly clear, and the study is detailed and impactful. After the authors address the reviewers' comments, I think the paper should be published.

## General Comments:

1. The paper uses DDA as the truth, which is very reasonable; however, it would be interesting if the authors would also venture into providing some recommendations on what experimental laboratory or field studies should focus on in terms of "tuning" parameters, observational variables, scenarios or conditions, to fill similar gaps to those the authors identify in DDA simulation space.

   Based on a review of laboratory studies of the optical properties of BC-containing particles, we would recommend that future laboratory experiments target smaller aggregates ($N_{pp} < 100$) that are thinly coated ($V/V_0 < 5$). Previous laboratory experiments have shown that particles in this $N_{pp}$ range undergo moderate coating-induced restructuring when $1 < V/V_0 < 5$ (Corbin et al., 2023). Therefore, we expect particles in this region to have large morphological variability. This is supported by high epistemic and aleatoric uncertainty for particles in this region. Therefore, measurements of optical properties for these particles would further constrain model estimates and reduce epistemic uncertainties while better characterizing the resolved variability represented by the aleatoric component. Measurements that jointly constrain aggregate size, coating amount, morphology, and wavelength-resolved optical properties would be particularly valuable. We have included this discussion on lines 445-452 of the revised manuscript.

2. It would have been interesting to also add comparisons with the RDG approximation, at least for uncoated BC particles.

   In our original manuscript, we limited comparisons of hybrid BNN predictions to Mie theory in the main text, given that homogeneous-sphere and core-shell models are currently implemented in large-scale models. We have added a comparison between the hybrid BNN and the RDG approximation to Appendix Figure I1. This comparison shows that the hybrid BNN is more accurate than RDG approximation for $Q_{\text{ext}}$, with median errors between 0.02 and 0.05 for the hybrid BNN and between -1.4 and 0.46 for RDG approximation. The hybrid BNN and RDG approximation yielded similar predictions for SSA, with median errors between -0.01 and 0.02 and between -0.04 and -0.01, respectively. This comparison shows that the magnitude of scattering and absorption is underestimated by the RDG approximation, but the ratio of scattering to absorption is well-captured. Finally, the hybrid BNN is also more accurate than the

RDG approximation when calculating $g$ for coated aggregates, with median errors between -0.02 and 0.01 for the hybrid BNN and between -0.20 and -0.15 for RDG. We have added this comparison to Appendix I and added a brief discussion of the results to lines 326-330 of the revised manuscript.

3. Perhaps a stronger emphasis should have been given to the absorption because, in the atmosphere, the other optical parameters (for example, scattering, SSA and $g$) are dominated by particles other than those containing BC, owing to their lower concentrations; while it is the absorption properties of BC that are more radiatively important (e.g., to determine heating rates) despite their lower concentration due to their stronger absorption cross sections with respect to most of the other atmospheric particles. For example, in Figure 3, I would also add a plot with $Q_{abs}$ (instead of relegating it to Figure E1), even if that might be somewhat less sensitive to $D_f$.

   The reviewer is correct that total scattering is usually dominated by BC-free particles, while BC-containing particles have a strong influence on absorption. However, we emphasize $Q_{\mathrm{ext}}$, SSA, and $g$ because these quantities are commonly supplied to radiative transfer models, while $Q_{\mathrm{abs}}$ is typically calculated from $Q_{\mathrm{ext}}$ and SSA. We have expanded our discussion of the relative insensitivity of absorption to particle shape in the revised text (lines 316-325), which directs readers to Appendix H for more information.

4. Some additional references on previous work on these issues, especially in the realm of experiments and field studies, would be helpful.

   We have added citations to Cappa et al. (2012), Hu et al. (2021), Lack and Cappa (2010), Liu et al. (2017), Liu et al. (2020), and Shamjad et al. (2012) to the revised manuscript. These studies provide additional laboratory and field evidence for the influence of coating amount, mixing state, and particle morphology on the optical properties of BC-containing particles.

**Specific Comments:**

1. Line 30. Restructuring can also happen during cloud processing.

   We have specified that our reference to coating in this sentence could also apply to water during cloud processing.

2. Beginning of section 2. While Ensemble II "represents scenarios where liquid coatings induce restructuring via surface tension and capillary forces", what does Ensemble I represent? Is this more consistent with processes such as cloud-induced compaction, or cases where the coating (other than water), initially present on the particle, eventually evaporates?

   These particles generally represent particles with decoupled $V/V_0$ and $D_{\mathrm{f,c}}$. The reviewer is correct that this could arise from coating evaporation. This could also arise if coating accumulates on BC aggregates via coagulation or via nanodroplet activation, which can occur under high supersaturation conditions. We have added this clarification to lines 74-77 of the revised manuscript.

3. Title of section 2.3: Although I do understand the reason for this wording choice, "numerically exact" might seem a contradiction with the term "approximation" in reference to the numerical model used (DDA). Perhaps a different section title would be better.

   We have changed the title of this section to "Optical Properties of Simulated Black Carbon-Containing Particles" in the revised manuscript.

4. Line 120: "other similar models" such as?

We have specified in the revised manuscript that similar models include traditional neural networks which provide point-estimates of target variables.

5. The meaning of aleatoric vs. epistemic uncertainty is relatively clear; however, it would be even clearer if a simple example for each were provided.

   We have added a simple example which demonstrates aleatoric and epistemic uncertainty on lines 183-190 of the revised manuscript.

6. Line 159: Does theta include W and b (weights and biases)? If so, I find it confusing that later (in several following points in the paper) theta is referred to as "weights" only.

   We have changed references to $\theta$ as "weights" to "weights and biases" in the revised manuscript.

7. Lines 188-189: What is the reason for the miscalibration?

   The fitted values $\tau < 1$ indicate that the raw predictive intervals are systematically wider than required to achieve the observed validation set coverage, as described in Appendix C. These larger intervals may be associated with the independent-parameter approximation used for the Bayesian-layer parameters, which does not represent correlations between parameters, together with the nonlinear propagation from latent to physical space (Foong et al., 2020). We do not attribute the observed miscalibration to a single mechanism. Instead, we correct the observed miscalibration using post-hoc, per-target uncertainty scaling fitted on the validation set, following established approaches (Guo et al., 2017; Kuleshov et al., 2018). This scaling does not modify the trained model parameters and is applied only during inference. We clarify these possible sources of miscalibration and the subsequent calibration procedure on lines 174-175 and 207-208.

8. Line 192 and following: So, is tau just a tuning knob to match the validation targets?

   The reviewer's understanding is correct: $\tau$ is a post-hoc uncertainty-scaling factor, fitted on the held-out validation set after training by minimizing the Gaussian negative log-likelihood in latent space, following the process of Guo et al. (2017) and Kuleshov et al. (2018). The scaling is applied only at inference and does not modify the trained model parameters. Its purpose is to correct the small systematic miscalibration of raw variational-BNN predictive intervals described above. We have expanded the description of $\tau$ in the revised manuscript to highlight its post-hoc, calibration-only role (lines 207-217).

9. Line 208: How is the Gaussian noise chosen?

   We use Gaussian noise during inference as the model is trained using a diagonal Gaussian likelihood in latent space. The network predicts the mean and aleatoric standard deviation, and the standard normal noise is scaled by this predicted standard deviation during sampling. This makes the inference procedure consistent with the training objective. We have clarified the Gaussian sampling procedure and its relationship to the training likelihood in Section 3.3.2 of the revised manuscript.

10. Line 248: The distinction between $D_f$ and $D_{f,c}$ is important. Can the authors explain how the two are calculated? Related to this point, on line 250, it is mentioned that $D_f \leq 2$ represents uncoated or coated particles not significantly compacted. Is the compaction here referring to the core alone or to the full particle? If it refers to the core only, as it would seem more physically appropriate, then the coating could still result in a particle with an overall $D_f$ much higher than 2 even when the core is little or not at all compacted; for example, in the case where the coating is so thick that the information on the underlying core morphology is lost and the coating looks close to spherical on the outside surface. Some clarification on these points would be useful.

The reviewer is correct that compaction here is in reference to the BC core. The reviewer is also correct that either a compacted core with a thin coating or an uncompacted core with a thick coating can result in $D_\mathrm{f} > 2$. We have added text from 276-282 of the revised text clarifying this point. We have also specified on lines 275-276 that $D_\mathrm{f}$ and $D_\mathrm{f,c}$ are calculated directly from the structure of each particle using the structure factor, which is described in detail in Sorensen (2001).

11. Line 255: This asymptotic behavior is reasonable and has been shown in the literature before.

    We clarified on lines 285-292 of the revised text that Mie theory has demonstrated accuracy for aged BC-containing particles with surface morphologies that approach that of a sphere. We also note that the hybrid BNN is also accurate for these particles, but has lower errors for particles with $D_\mathrm{f} < 2.5$.

12. Lines 282 – 283: As written, this sentence seems to imply more of a numerical or mathematical reason for the low sensitivity of $Q_{abs}$ to $D_f$, just resulting from error cancellations in the calculations of $Q_{ext}$ and SSA. However, I believe the low sensitivity of $Q_{abs}$ to $D_f$ has a much more physical foundation, at least for uncoated particles; a reason that is also at the root of why the simple RDG approximation (based on compounded monomer number or total mass) for $Q_{abs}$ works relatively well as compared to $Q_{sca}$ and therefore also for $Q_{ext}$ and SSA. It is true that in the model here, $Q_{abs}$ is given by $Q_{ext}$ and 1-SSA to account for conservation of energy, but the physical lower sensitivity of $Q_{abs}$ is reflected, not caused by this mathematical relation.

    We have modified lines 320-325 to reflect the physical basis for the accuracy of core-shell and homogeneous sphere approximations for calculating $Q_\mathrm{abs}$. While $Q_\mathrm{abs}$ is often calculated from $Q_\mathrm{ext}$ and SSA, we agree that the appropriateness of equivalent-sphere approximations has a physical basis. This nuance is now reflected in the revised manuscript.

13. Line 298: This has indeed been shown to be the case in previous literature.

    We have noted that this effect has been demonstrated previously and referenced Sorensen's comprehensive 2001 review on the topic.

14. Line 299: Define $m_{eff}$.

    We have added a definition for $m_{eff}$ on lines 345-349 of the revised manuscript.

---

## References


Cappa, C. D., Onasch, T. B., Massoli, P., Worsnop, D. R., Bates, T. S., Cross, E. S., Davidovits, P., Hakala, J., Hayden, K. L., and Jobson, B. T.: Radiative absorption enhancements due to the mixing state of atmospheric black carbon, Science, 337, 1078–1081, 2012.

Corbin, J. C., Modini, R. L., and Gysel-Beer, M.: Mechanisms of soot-aggregate restructuring and compaction, Aerosol Science and Technology, 57, 89–111, 2023.

Foong, A. Y. K., Burt, D. R., Li, Y., and Turner, R. E.: On the Expressiveness of Approximate Inference in Bayesian Neural Networks, https://arxiv.org/abs/1909.00719, 2020.

Guo, C., Pleiss, G., Sun, Y., and Weinberger, K. Q.: On calibration of modern neural networks, in: International conference on machine learning, pp. 1321–1330, PMLR, 2017.

Hu, K., Liu, D., Tian, P., Wu, Y., Deng, Z., Wu, Y., Zhao, D., Li, R., Sheng, J., Huang, M., et al.: Measurements of the diversity of shape and mixing state for ambient black carbon particles, Geophysical Research Letters, 48, e2021GL094 522, 2021.

Kuleshov, V., Fenner, N., and Ermon, S.: Accurate uncertainties for deep learning using calibrated regression, in: International conference on machine learning, pp. 2796–2804, PMLR, 2018.

Lack, D. and Cappa, C.: Impact of brown and clear carbon on light absorption enhancement, single scatter albedo and absorption wavelength dependence of black carbon, Atmospheric Chemistry and Physics, 10, 4207–4220, 2010.

Liu, D., Whitehead, J., Alfarra, M. R., Reyes-Villegas, E., Spracklen, D. V., Reddington, C. L., Kong, S., Williams, P. I., Ting, Y.-C., and Haslett, S.: Black-carbon absorption enhancement in the atmosphere determined by particle mixing state, Nature Geoscience, 10, 184–188, 2017.

Liu, F., Yon, J., Fuentes, A., Lobo, P., Smallwood, G. J., and Corbin, J. C.: Review of recent literature on the light absorption properties of black carbon: Refractive index, mass absorption cross section, and absorption function, Aerosol Science and Technology, 54, 33–51, 2020.

Shamjad, P., Tripathi, S., Aggarwal, S., Mishra, S., Joshi, M., Khan, A., Sapra, B., and Ram, K.: Comparison of experimental and modeled absorption enhancement by black carbon (BC) cored polydisperse aerosols under hygroscopic conditions, Environmental science & technology, 46, 8082–8089, 2012.

Sorensen, C.: Light scattering by fractal aggregates: a review, Aerosol Science & Technology, 35, 648–687, 2001.