I appreciate the effort the authors have put forth in addressing reviewer comments from the previous round of revisions. The newly included figures, comparisons with PCA, and cosine similarity analysis has helped reinforce the paper's results, and made for a more compelling read. I also appreciate the effort put forth in improving the figure legibility, as this is now greatly improved. Before I fully recommend the paper for publication, I did have a few minor comments for the authors to consider:

We thank the reviewer for their careful reading of the manuscript and for their additional comments and suggestions. They have further improved the quality and clarity of the manuscript.

General: I'd recommend doing one more read-through for grammar/spacing etc. since a lot of the text has changed. I won't list everything but small errors like a missing space on lines 235-238, also I believe it should be "scikit-learn", should be addressed before publication.

Thanks for the suggestion. We have reread the text and corrected a number of small typos throughout.

Figure 5: The tick labels on the far-right plot colorbar are not properly aligned.

We have updated the tick labels to be properly aligned.

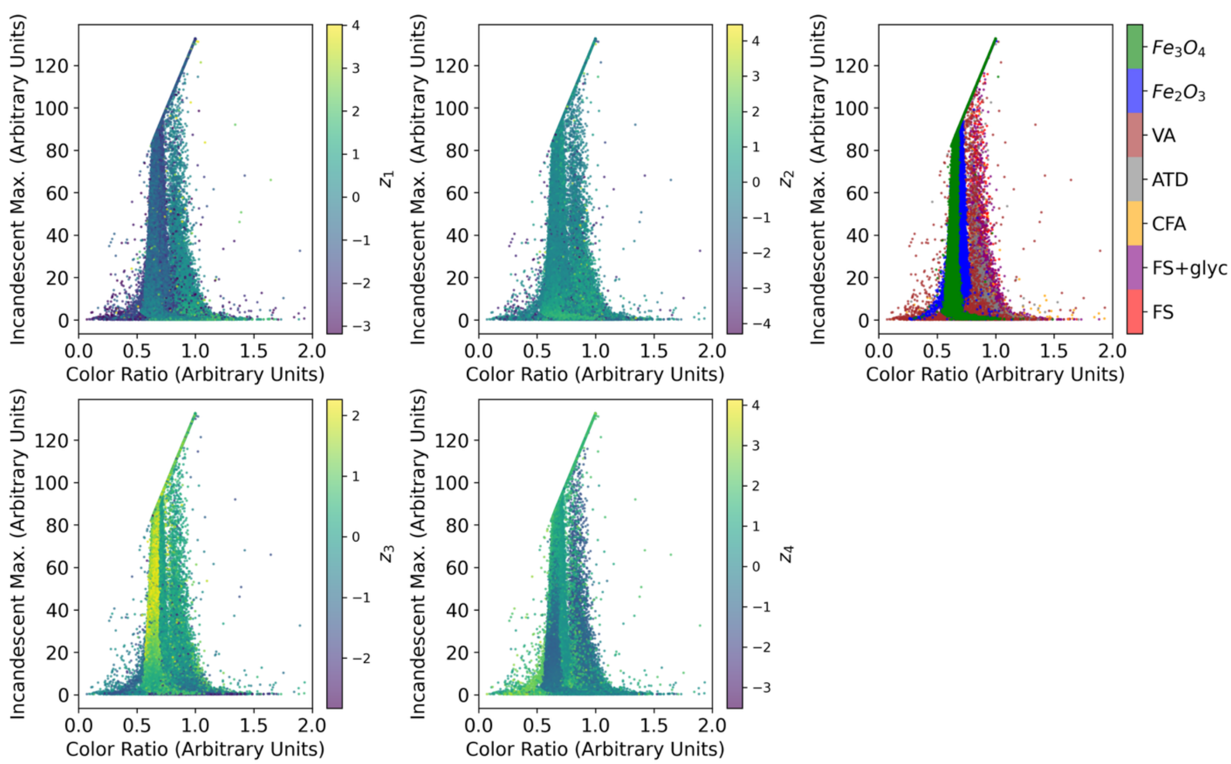
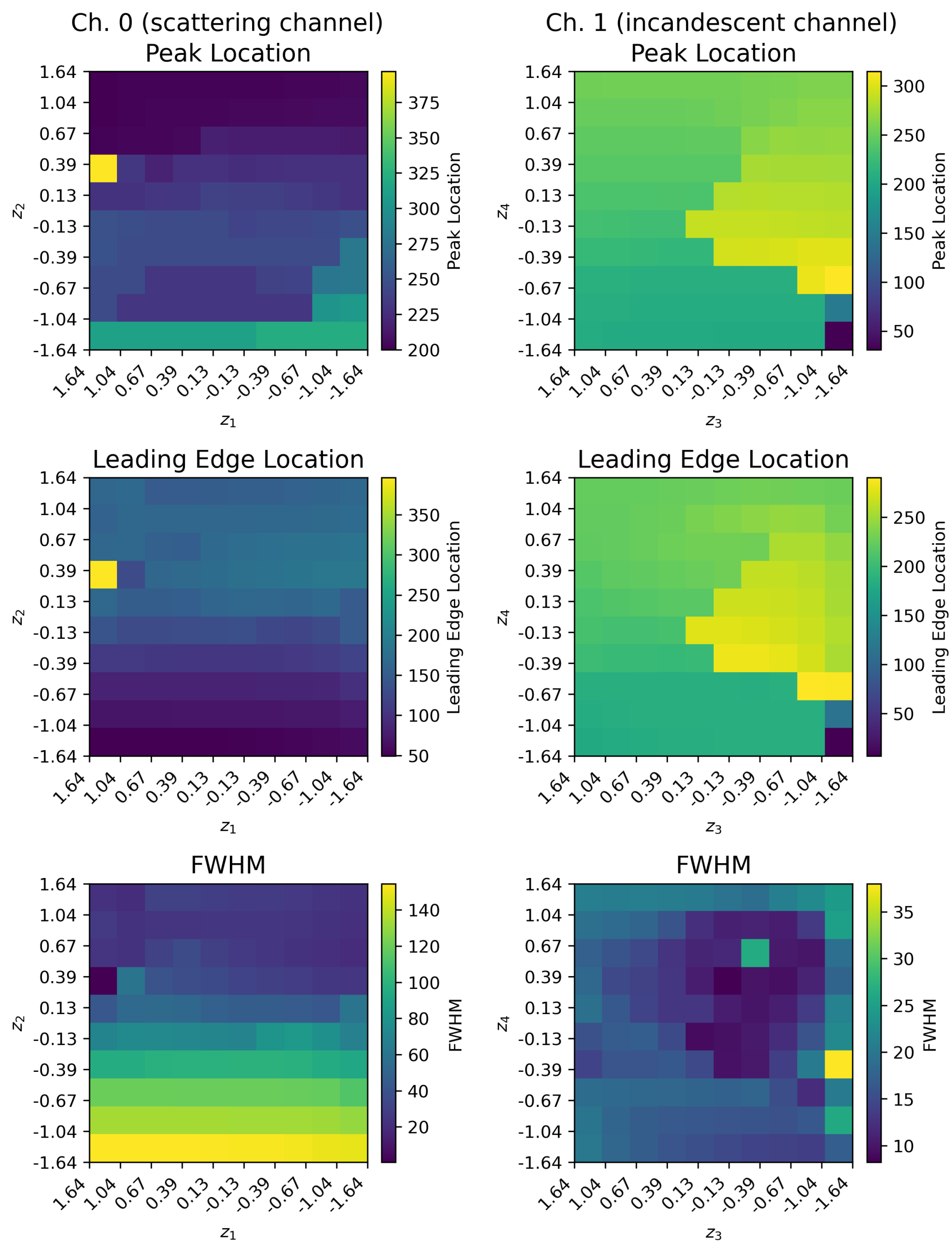


Figure 7: I like the addition of this figure, but I think it is lacking a bit of context. For instance, x/y-axis labels. Further, laid out this way, it feels like we should be comparing between the scattering and incandescent channels, but they colorbars have different scales which makes this challenging.

We have added x and y labels to the axes. We also flipped the orientation from horizontal to vertical to make it clearer that the comparison should be against the latent manifolds shown in Figure 6, rather than between the two channels. We have updated the caption of Figure 7 to reference Figure 6 in order to clarify this.



Also, what is the bright/dark feature at 1.64/0.39 in Ch. 0 and how should we interpret it? It really draws the eye here.

As can be seen in the latent manifold for Ch. 0, the reconstructed Ch. 0 signal corresponding to the latent values of (1.64, 0.39) is due to the second large peak in the signal, which is to the far right in the L-II time-series. Because the initial peak (which triggers the SP2 detection) is very small for this particular signal, the later peak is the one picked out by the peak finding algorithm from Scipy's signal analysis package as the predominant peak. Other reconstructed time series also show evidence of a second peak but they are not the dominant peaks in the time series, and thus are ignored in this analysis. This demonstrates how the VAE is able to pick up some of the more subtle shifts in the L-II signals, when compared with a more naïve application of a peak-finding algorithm that is traditionally used in analyzing the L-II signals from the SP2.

Section 5/Figure 10: This is an interesting analysis, especially the comparison with the PCA. I am curious though, due to the very high pattern correlation across both the 2D/3D VAE matrices, and the PCA matrix (which look basically the same with slightly different magnitudes), what is gained from a nonlinear approach over the PCA? Or is the PCA sufficient here? It wasn't clear what the takeaway was now that you have shifted from a solely nonlinear approach to PCA+VAE.

We added the following paragraph to the Conclusions in lines 436-442:

When comparing the linear and non-linear unsupervised machine learning approaches (PCA vs. the 2D and 3D VAEs), we found that the 2D and 3D VAEs did perform slightly better in terms of finding lower dimensional representations of the L-II signals that demonstrated higher within class similarity (Section 5) and improved downstream classification (Section 6). However, the advantage was relatively small, indicating that more interpretable PCA analysis could already provide many useful insights when analyzing L-II signals. More sophisticated unsupervised and semi-supervised learning approaches (as we comment on in the last paragraph of this section) could provide additional advantages over the VAE and PCA approaches, however, and warrant future research.

Section 2.3: On the generalizability topic, I was more so considering the fact that after training on the lab data, if that model is then used operationally and encounters aerosols outside the convex hull of training data, that the compressed latent vector may no longer contain relevant information, as the model has never seen cases such as these. We've run into issues like this before where we did our best to get the most robust set of training samples possible, but in real world situations, there is a chance to encounter something outside of the training set which the model does not necessarily handle properly. I was hoping to see some discussion around this and whether or not this would be an issue with the aerosol data.

This is the point that we are trying to make in terms of the utility of these approaches for outlier detection (e.g. finding populations that are not well-represented in laboratory data). We have added a caveat to the discussion to acknowledge this point in Lines 465 – 467:

However, because outliers may correspond to L-II signals that fall far outside the distribution the VAE was trained on, their latent embeddings can become unreliable. This highlights an important open question regarding the robustness of latent-space methods for outlier detection for the SP2, which warrants future research.