General Comments: The analysis has been much improved since the last version, but it appears that some of the conclusions were hastily written, as they're not supported by the figures (see below).

Additionally, the text is still very colloquial and needs to be corrected to be more readable. There is a mixture of tenses, sometimes in the same paragraph. For example, Page 17, Line 4-5: "For our experiments we decided not to include atoms that were constructed from EDGAR or geostatistical data. We will use a pixel basis".

Thanks to the authors for adding the error analysis in addition to the sensitivity analysis. It's interesting that they don't quite agree (i.e. the L2 and L1 POS DIC are not all that different in errors), and this difference needs to be explained better in Section 5.2. As it stands, it's difficult to connect the sensitivity, smoothing and actual errors in the three subsections.

Specific Comments:

Page 3, Lines 13-14: It's incorrect to say that the particles "travel backward in time". Rather they sample the adjoint of the atmospheric transport, of which time is a dimension. This is not the only paper to make this simplification, but it's confusing given the role of diffusion, etc in transport. Also, the footprint is the surface influence on the measurement, rather than airmasses as it's stated here.

Page 3, Lines 20-22: This sentence at the end of the paragraph feels a bit out of place, since it's not an exhaustive list of the open problems in AIM, but rather two selected examples. I suggest to leave out these examples and state that this method addresses only the issue of representation of solutions in the inverse problem.

Page 4, Line 4-5. I think you mean that the dimensions of the two spaces are not the same. Are you making a claim about the sizes of n and m? Throughout the paper, you use the word "realistic", though that's not a precise word. The last sentence in this paragraph should be rephrased, since each norm will give a unique solution, irregardless of how "realistic" it is.

Page 4, Line 14: The parameters are sensitive to the noise, rather than the data, in the way that you've defined the terms on Line 12.

Page 5, Line 1: "best include this information". This seems to be a hanging fragment.

Page 5, Line 9: It may be clearer to the reader to explain what you mean by "the origin". I'm not sure what you mean here by "oscillations". Is it the dipole behavior in the optimized fluxes in the absence of smoothness constraints? This needs a bit more development to be clear.

Page 5, Line 15: "and vice versa" Do you mean rougher estimates with negative correlation? What would rougher mean? Larger dipoles, I assume, but I haven't ever seen

negative correlations in background covariance matrices personally, and I expect that this would cause instability in the estimation problem, which assumes positive definite matrices.

Page 9, Line 29: “often straightforward to calculate analytically” - this is true only for a small set of distributions, of which the Gaussian is the prime example. However, it’s not the case that most modern flux inversion techniques estimate the posterior uncertainty using these formulae, as the covariance matrices in question are of very high rank. Typically variational or ensemble techniques construct estimates using Monte Carlo methods.

Page 10, Line 12: “a smaller smoothing error results in a greater measurement error. The smallest total error is expected when both terms are approximately balanced.” I’m not sure what you’re referencing here. Is this a general principle? If so, please provide a reference, or give an example.'

Page 10, Line 25-26: What does it mean that assumptions are hard to guarantee? Are typical state vectors and models not sufficiently smooth/bounded/…?

Page 21, Line 29-30: Can you mark the location of the single large point source in Figure 8? The maps don’t make this obvious at all, and look like the flux field in Figure 3, rather than a single large point source. Maybe I’m misunderstanding?

Page 23, Line 31: Why would the L1 POS have a larger smoothing error? Earlier text points to the success of this technique for targeting pixel sources, so this result is confusing.

Figure 9: It’s not clear from this figure that L2 POS isn’t the best method overall, as the dipoles seem to be smallest in this figure, even though the overall MSE is larger than the L1 DIC POS method. Similarly, the smoothing error isn’t obviously better in the bottom panel than the top panel.

Page 26, Line 5-7: This conclusion is much too strong. For the regional fluxes, the standard inversion is as accurate as the L1 POS DIC inversion, particularly in the large emitting regions. It’s difficult to pick a clear winner between all of the different methods in this case.