Reply to Anonymous Referee #4

Corrections: Section 1.2 (pg 2330, Line 19): "... provided only that the requisites PDFs are known." -> requisite

Fixed .. thank you.

Conclusions/Discussion (Page 2342): "To mitigate the problem of dimensionality in Bayesian retrievals, we described an algorithm for objectively distilling the relevant information content from N channels into a smaller number (M) pseudochannels while also regularizing the background (geophysical plus instrument) noise component. In the present demonstration, M = 3 and N = 1. In the application of this method to TMI data described by Petty and Li (2013), M = 9 and N = 3." Shouldn't M < N ? This is confusing. Also, it sounds like you used M=3 in the present demonstration, but M = 1 ... I think M and N need to be switched here.

Yes, M and N were inadvertently reversed .. now fixed.

Section 1.2
Where do these "candidate solutions" come from? At some point, somewhere in a given retrieval algorithm, there is a modeled relationship between the radiances and the geophysical parameter(s) one is interested in. Whether it's a radar-derived precipitation rate (e.g., a Z-R relationship) co-located with radiance observations, or a CRM database of profiles with forward modeled radiances. Consequently, the model bias that one is so eager to disconnect themselves from gets buried somewhere or, worse, over-constrains the retrieval problem by under-populating the solution space. Given a (theoretical) perfectly co-located and beam-matched radar observation for each feedhorn on a given sensor, one is still wholly limited by both (i) the sensitivity range (and instrument error) of the radar; and, (ii) the physical relationships between geophysical parameters (gas, precipitation, surface, multiple-scattering, clutter noise, etc.) and the measured reflectivities. Given this idealized scenario, one now has a basis for a "pretty good" retrieval algorithm, _but only for the cases that the radar(s) could observe_. In the case of TRMM, for example, this would mean a very large percentage of precipitation occurrence (e.g., light precipitation) would never be retrieved with skill. Could one improve upon this by performing a similar "dimensionality reduction" on the radar observations (or whatever source observations)? The reflectivity at each range gate is a measurement, although not truly independent of the preceding ones due to path-integrated attenuation and, possibly, multiple-scattering effects (see Battaglia, for example).

These are all valid issues affecting *real* retrievals using *real* data. The present paper deliberately sidesteps ALL of the physical issues highlighted above by the reviewer and focuses exclusively on issues of sampling and weighting and how these issues are mitigated via the proposed dimensional reduction. This is in fact exactly why we elected to use highly idealized fake data in our experiments rather than real data, so as to avoid getting bogged down in those unrelated (but important) complications. By the way, in our JTech papers (now accepted) that apply our technique to real data, the explicit assumption is that we're not trying to retrieve "reality" but rather using the radiometer to retrieve whatever the radar *would* have reported. Specifically, we're training the TMI to retrieve the PR-based 2A25 product, including the latter product's defects.

Section 3.1 (pg 2335, Equation 3): This example may strengthen your argument: I was playing around with a simple example of equation 3, and noticed that if one simply increases the number of channels – without adjusting sigma_i – "s" also naturally increases. So if one adds additional radiometer channels to the typical "Bayesian" retrieval, the weight (w=exp(-s)) rapidly decreases. The act of adding a single channel will, because of the threshold w > 0.01, will result in potentially worse retrieval quality.

While the point is technically correct, I believe it's one of somewhat peripheral interest, for the following reasons: (1) the choice of minimum weight w can easily be modified to suit the number of channels used. (2) If one were to simply precede equation (3) with 1/N and then take the square root, it would then be a measure of the Euclidean distance between the observation and the candidate solution, and s would then be fairly insensitive to the number of channels N. For fixed N, these additional operations are somewhat irrelevant and can be absorbed into one's choice of w.

Figure 4: I realize this is still "background" stuff, but what's the deal with the near-zero retrievals when the true rates are as high as 3? It would be interesting to have a color-coding (or shading) to indicate what the sigma value is for each point. Are there cases where the retrieval is near the 1:1 line, but the sigma values are really large – i.e., a good match for the wrong reasons?

I believe the answers to the above questions are readily found in Fig. 3. Specifically, Fig. 4 can be thought of as a hybrid of the four panels in Fig. 3 (plus additional values of sigma not depicted). Clearly the near-zero retrievals with true rates near 3 are associated with the largest values of sigma (panel d in Fig. 3). Also, there are indeed a couple of points (for R ~ 0.8-0.9) that fall near the 1:1 line in Fig. 3d, but they're only a tiny subset of the total.

Figure 5, 6, & 7: It's hard to tell what the actual retrieval skill is on these plots, particularly at low precipitation rates.

Actual retrieval skill is arguably irrelevant in this paper, since we're utilizing "fake" data that cannot saying anything useful about actual skill when similar techniques are applied to real data. The far more important point, I believe, is the marked improvement in the overall behavior of the retrievals when directly comparing Fig. 5 with Fig. 2, Fig. 6 with Fig. 3, and Fig. 7 with Fig. 4. I hope the reviewer agrees that these comparisons are compelling.

Figure 7: Same sentiment as my comment about figure 4.

See above.

Other Comments and Recommendations
For a very long time, the community has been recycling poor (statistically) "matching" algorithms, and, we keep putting lipstick on the pig by improving the various bits and pieces without changing the actual framework. Even worse, perhaps, is that the retrievals obtained from these algorithms get propagated into various climate datasets, degrading the potential knowledge obtainable from past and present precipitation retrievals.
The present method here, while not necessarily mathematically new, presents an important (and easy to implement) approach to improving upon this long-standing problem. Future retrieval approaches would be wise to utilize the method presented here to improve upon the dimensionality problem, and isolate the variables to be retrieved – or, alternatively, determine those that cannot be isolated.

I agree with the reviewer, and thank you for the supportive comments. The two papers now accepted in *JTech* (see references below) will hopefully persuade others as well.

A few things I would have liked to seen in this paper:

(1) Application to real observations (I realize space considerations are an issue, will this be a subject of a future publication?)

Yes:

Petty, G.W., and K. Li, 2013: Improved passive microwave retrievals of rain rate over land and ocean. 1. Algorithm description. In press, *J. Atmos. Ocean. Tech.*

Petty, G.W., and K. Li, 2013: Improved passive microwave retrievals of rain rate over land and ocean. 2. Validation and intercomparison. In press, *J. Atmos. Ocean. Tech.*

(2) Additional eigenvalues (M > 1) and a physical relationship between values of the Mth eigenvalue and precipitation rate (or whatever variable it's actually sensitive to, that's never clearly stated .. despite matching to precipitation rate in the training/val database. It could be that, for example, cloud ice is strongly correlated with precipitation rate, and the first eigenvalue happens to be the sensitivity to that. Which is "okay" in the sense that it ultimately gives you what you want, but that limits one to a certain set of microphysical processes in retrievals – i.e., you might miss warm rain altogether).

In the papers cited above, M=3, so we're retaining sensitivity to more than just one physical signature (or perhaps also to non-linearities in the physical signature).

(3) Retrieval skill. Visually it's easy to discern that at high precipitation rates, the proposed algorithm performs well. At low precipitation rates (what GPM is purportedly designed to retrieve), it's difficult to discern on the figures how well or poorly it is doing. A log scale in precip rate would be an easy step to accommodate this visual inspection, a slightly more involved step would be to assign a skill to the retrieval or clearly denote variance in a different way. I don't have an immediate good idea about how to communicate that clearly.

We use a logarithmic depiction of skill in part 2 of the above-cited papers. As mentioned before, in the present paper, we don't want to get too hung up on the details of skill at various "rain rates" for this admittedly idealized "fake" data set.

(4) Dealing with extreme and/or uncommon events. It was mentioned in the beginning, but I didn't notice any additional discussion of this important aspect of retrievals.

The problem of uncommon events is that you are less likely to find a match at all, and this problem is worse for higher dimensional matches. In our part 1 paper cited above, we describe an algorithm for first searching for a match with M=3 and then, failing that, falling back to M=2, etc. Eventually a match is always found, though perhaps with less contribution information from the various pseudochannels.

(5) A final comment about non-linearity – there's very little discussion of non-linear relationships between the transformed TBs and the precipitation rate. It appears that you are arguing that by reducing the off-diagonal elements of the covariance matrix, that you are mitigating the non-linear response. It's not clear to me that this is what is occurring. Could you discuss the effects of non-linear relationships in the present approach?

Non-linearity isn't addressed in this paper at all, and the off-diagonal elements of the covariance matrix are assumed here to arise from physically based cross-correlations between channels as opposed to non-linearities. The "fake" rain signature is purely linear, and the background noise is purely Gaussian (i.e., no curvature of the noise cloud). Non-linearity poses additional problems for the dimensional reduction in that if you spherize the background noise (as we do in this paper), you might not actually get as much improvement in the effective sampling density as you do in the purely linear case. But while this might reduce the overall performance gain, it arguably does not eliminate the benefits of the dimensional reduction altogether.