Reviewer #1

The paper is reasonable as a description of a retrieval method, but it lack a demonstration that the algorithm is practically useful. The authors present the evaluation of simulated data, but do so in such idealised circumstances that the exercise provides little insight into the algorithm's overall utility. However, the presentation is decent and the central idea interesting and worth publishing, even if I'm not entirely convinced it is valid.

The paper presents a new concept for the retrieval of aerosol single scattering properties that have never been published before. As this concept is new, it has been presented in ideal situation to ease its understanding.

To be more specific in my critique:
• The paper must demonstrate the algorithm applied to noisy data. As it stands, the results shown are somewhat concerning as the algorithm can only precisely replicate the truth in the simplest case despite perfect input data. This implies significant (if not large) forward model errors that the paper neither discusses, quantifies, nor attempts to alleviate. Considering CISAR has already been applied to real data, this demonstration should be a straightforward matter of adding random noise to the existing experiment's data and adding an additional line to Figs. 5–9 for the retrieval under noise.

Adding noise to the simulated data is certainly useful as actual observations are not noise free. However, in that case, the performance of the method would need to be statistically demonstrated applying the retrieval on a large number of simulations. We do not believe that it would have ease the reading of the results. Similarly to adding noise to the data, we would have to consider actual observation conditions and include sections on information content specific to the selected radiometer and orbit. In order to avoid all these considerations and keep focused on the actual paper objective, we prefer not to start addressing these practical issues in this theoretical paper.

The following sentence has been added at the end of Section 4.4

A similar comparison has been performed against actual PROBA-V observations (Luffarelli et al., 2017). These comparisons a root mean square error between simulated and actual observations of about 0.03.

The following sentence has been added to Section 6.1, experimental setup.

An uncertainty of 3% is assumed in matrix S_y.

• Additionally, your experiments assume that the surface parameters are known a priori. That is unacceptable considering the title of the paper is 'joint retrieval of surface reflectance and aerosol properties'. If you promise a joint retrieval, you need to demonstrate it. If you don't intend to demonstrate it, redraft the paper to de-emphasise the surface terms while explaining why you built a joint retrieval and chose not to use it. Such a demonstration is important as the reader needs to understand how the optimal estimation balances uncertainties in aerosol against the surface. (I would show averaging kernels, but that isn't widely popular.)

This specific issue of surface retrieval has been addressed in a previous paper by Wagner et al. (2010). This paper demonstrates that an accurate retrieval of surface reflectance is possible with this joint retrieval approach. It requires however to include a complex scheme for the estimation of the surface prior information from the previous retrieval. Once again, we prefer to avoid these aspects as there were already published and deviate from the paper primary objective.

• From my experience teaching linear algebra, I would summarise the paper's central idea as, "Rather than assume aerosol optical properties for the retrieval, one should define a basis of aerosol types. The observed properties of any real aerosol are then some linear combination of the properties of those basis types."

This statement correctly summarizes the paper concept. However, the possibility to express the scattering albedo and phase function values as a linear mixture of the vertices properties need to be demonstrated. That is the main purpose of Section 6. The end of Section 1 has been re-written and reads now

The advantages of a continuous variation of the aerosol properties in the solution space against a LUT-based approach is discussed in Section (3). The proposed method expresses the scattering albedo and phase function values as a linear mixture of basic aerosol classes. The forward radiative transfer model that includes the Jacobians computation is described in Section (4). With the exception of gaseous transmittance, this model no longer relies on LUTs, and the RTE is explicitly solved. The inversion method is described in Section (5). Finally, the possibility to express aerosol single scattering properties as a linear combination is illustrated with simulated data representing various scenarios including small and large particles (6).

From that, I wonder if you could use an objective technique to select your vertices? Given that a large number of aerosol types are currently defined, there are various techniques within linear algebra from which a minimal set of types can be derived (e.g. Gram-Schmidt or one of the eigenfunction analyses). This could give you better fits in the ideal circumstances and would unambiguously resolve the question of how many vertices you should use (which is rather unsatisfying at the moment).

Please refer to our answer of Reviewer #2 main concern.

• A central assumption of the paper is that either the surface or the aerosol vary sufficiently slowly as to be effectively constant over the observation period. There needs to be some justification of this. I wouldn't accept a simple citation of existing work as the quality of such assumptions is highly dependent on exactly where and how you evaluate them.

This comment is fully justified. The speed at which surface radiative properties changes is also highly dependent on the spatial resolution. This method has been published in JGR and we do not see the need to come back on that work. It allows to stay focused on the paper objective. We however allow the aerosol optical thickness to vary quite rapidly.

• Various sections of this paper rely on a familiarity with Govaerts et al. (2010). A less brief summary of it's conclusions and how that method differs from this one would be useful to the casual reader.

The two main differences between Govaerts et al. (2010) and the current version are listed in Section 2. It concerns:

- The retrieval is not performed any more for a limited number of prescribed aerosol classes but with a continuous sampling of the solution space bounded by the vertices;
- The forward model solutions are not pre-calculated for a limited number of solutions and stored in LUTs. In this new model, the radiative transfer equation is solved online. It allows calculating the derivatives in any points of the solution space. With a LUT-based approach, it is also possible but very CPU intensive.

L95 While it is strictly true that optimal estimation requires state vector variables to be continuous, I feel that your argument here mis-characterises what is going on. As you admit, there are over 100 variables that affect the radiance seen at TOA scattered by aerosols. The prevalent approach is to assume that most of those variables are determined at the aerosol's creation. Diner et al. (2012) demonstrated that superior results can be obtained by assuming less when additional information is available (specifically, multi-angular observations).
To me, it seems obvious that attempting to retrieve either more variables or more physically motivated variables will result in a higher quality retrieval. The issue isn't that most retrievals assume an aerosol type. Their issue is that they use too few observations to assume anything better. To put it differently, your retrieval is superior because it uses more data and so can vary more variables, providing the OE algorithm with more freedom to find an accurate solution.

As explained in Section 2, there is an inconstancy in trying to apply an OE approach and the use of pre-defined aerosol classes. By doing so, it is assumed that the solution should precisely correspond to these classes. It imposes thus very strong prior information on aerosol single scattering properties. However, there is no uncertainty associated to this prior information, which is inconsistent. Secondly, in order to find the minimum of the cost function, the derivatives of all state variables should be defined in any points of the solution space. This is clearly not the case with the use of pre-defined aerosol classes. The method proposed here addresses specifically this limitation.

L 202 I have a problem with your terminology here. My experience is that in aerosol remote sensing 'single' or 'multiple' refers to the number of times that the light is scattered by an aerosol particle. Eq. (2) quantifies the reflection from the surface.

It also refers to the possible interaction with the surface not only aerosols.

I would like to see a justification for this other than 'it works in Sec. 6.' Firstly, I'm uncertain if you are only considering single aerosol scattering or not. While I'm quite happy that single-scatting properties combine linearly, I am suspicious of this being true for multiple-scattering. Secondly, I doubt that linearity would hold for large optical depths. I expect that your technique is a reasonable approximation, but you should directly assess the error introduced (i.e. don't just work it out after-the-fact from the quality of your retrievals).

Forward and backward simulation are calculated accounting for the multiple scattering contribution as can be seen in Eq. (1) term $I_m^{\uparrow}$.

The linear approximation used to express aerosol single scattering properties has no influence on the way the multiple scattering contribution is calculated. No linear approximation is performed on the estimation of $I_m^{\uparrow}$whatherever the optical thickness.

Tab 1. Are these values ever used within the retrieval to represent forward model error? If not, they should be.

Yes, there are. The following sentence has been added at the end of Section 4.4

A similar comparison has been performed against actual PROBA-V observations (Luffarelli et al., 2017). These comparisons a root mean square error between simulated and actual observations of about 0.03.

The following sentence has been added to Section 6.1, experimental setup.

An uncertainty of 3% is assumed in matrix S_y.

L314 I have no problem citing grey literature in general, but this is unacceptable. The context implied to me that this referred to a validation study, not an unreviewed conference poster. I assume that application of CISAR to real data is in Part 2 (which really should have been submitted alongside this paper) and this reference should be adjusted accordingly.

We acknowledge that the part 2 should have been ready at the same time of part 2. Unfortunately it has not been possible. We will now submit this part in the coming weeks.

L402 This paper contains no analysis of other OE techniques. You cite other studies, but they only comment on specific algorithms. I have no problem with the conclusion that retrieving SSA and phase function provides 'better' results than not doing that when multi-angular observations are available. Your use of multiple MODIS overpasses to emulate multi-angular observations is fine only in those parts of the world where the aerosol or surface is constant over a period of days. Those of us evaluating plumes or areas that suffer frequent cloud cover can't make that assumption and your rejection of such work is excessive.

Aerosols do not require to be constant during the retrieval in case data are accumulated over time to form a multi-angular vector. Only the surface radiative properties are assumed to be invariant. An example of AOT hourly retrieval from SEVIRI can be found in Luffarelli et al., (2016). Example of plume retrieval over bright surface can be found here : http://www.rayference.eu/1/cisar_seviri.php. Example of retrieval for polar orbiting instrument can be found in Luffarelli et al., 2017.

L425 Experiment F00 did not prove your assumptions are valid. It showed that they might work in one circumstance, and even if you had made a comprehensive evaluation I'd say you only showed that the assumptions were 'useful'. I might accept 'Experiment F00 demonstrated that such assumptions can produce accurate retrievals.' However, my earlier concern that your experiments are insufficient to demonstrate the utility of your method remains.

This result is fundamental as it demonstrates the possibility to express the aerosol single scattering properties with Eq. (8) and (9) in inverse mode. This demonstration has been initiated with a simple case and which is further elaborated throughout the experiments.

L433 This statement is far too broad. Replace 'which provide a limited number of independent observations' with something that expresses the limitation of your technique. As it stands, it reads like your algorithm could use any input data and that is plainly false.

This sentence has been deleted and will be moved in Part 2 of the manuscript.

L435 I agree that the choice of vertices is critical. Your empirical selection is unsatisfying and this choice needs more thorough consideration.

Please refer to the reply of Review #2 main comment.

• I have several issues with the title.
– Though the algorithm can technically retrieve surface properties, this is not demonstrated in the paper nor really discussed other than as a component of an aerosol retrieval. The author's previous paper discussed the surface at great length and, though I don't expect them to replicate that here, a paper on a 'joint retrieval of surface reflectance and aerosol properties' should discuss and demonstrate both.

This part will be specifically demonstrated in part 2. There is nothing really new or original concerning the theoretical aspects of the surface property retrieval in this paper with respect to our previously published results. It is the reasons why we did not elaborate these aspects in this theoretical paper. What we clearly stressed in Section 2 is the importance of jointly retrieved surface and aerosol properties as they are the corresponding radiative fields are tightly coupled.

– The authors rightly wish to highlight the unique consideration of multiple types within their retrieval. However, 'continuous variations of the state variables in solution space' fails at that. At first glance, the phrase was virtually meaningless as optimal estimation can only be applied to continuous variables. What the authors have done is define state space differently, such that the retrieval can consider variations along different axes. As an alter-natives, I recommend 'Retrieval of surface reflectance and aerosol optical properties through decomposition into representative types: Part 1: theory' or 'Retrieval of surface reflectance and aerosol optical properties by simultaneously considering a representative set of aerosol types: Part 1: theory'.

This remark is too obscure for me to be able to comment it.

– Regardless, 'variations' should be singular.

This typo has been corrected.

L98 Though MISR might not assign uncertainty to aerosol class, that doesn't mean you can't and that people haven't.

The sentence is written in the context of estimating retrieval uncertainty in an OE context as described in Rodger (2000).

L171 Are you sure Liu and Ruprecht (1996) is the most appropriate reference for your radiative transfer solver? That paper assumed spheres and spends a lot of time talking about the microwave.

Yes, the reference is the correct one, being referred to the Matrix Operator Method.

L207 Why use finite differences to calculate the Jacobian? It's slow and inaccurate while all the solvers I know provide analytic Jacobians nowadays. If your solver doesn't provide them, you should explore alternatives. My experience is that the discontinuities produced by finite differences make optimal estimation much more sensitive to the first guess and take longer to converge.

It would be so nice to have an analytic solution of the multiple scattering equation. Unfortunately, it is not the case.

p10 L3 Shouldn't the phase function be a function of angle too?

The typo has been corrected.

Eq.(8) Is τ a that defined in Eq. (3) or should it be an equivalent definition using τ' ?

This typo has been corrected.

L220 A superior sentence would be, "The relative RMSE between FASTRE and the reference model is in the range 1 – 3%." There's no need to hide the larger difference in the 0.44 channel.

The text has been corrected accordingly.

Eq. (11) The third and fourth terms are common, but they aren't from Rodgers (2000). You'll need another reference for them.

The following reference has been added: Dubovik et al., (2011)

Tab. 2 Why doesn't F0 have a value for Nf ?

These types are mono-modal.

L328 Delete the sentence starting 'The estimated single scattering. . . "

For which reason should we delete this sentence?

Sec 6.1 Precisely how many observations are you using and what instrument are you emulating?

We are not simulating a specific instrument. Observations are simulated in the principal plane every 5 degrees.

The following recommendations have been implemented in the text:

L22 In most of the cases an increase
L23 for an increase of in the fraction
L25 does not allow to full characterisation of the underlying
L35 product generation from archived
L54 Such an approach prevents
L74 that time in the EUMETSAT
L77 to 1. It This represents a severe limitation
L78 exceed such a limit.
L83 initiative to generate for a Climate Data Record (CDR) generation of

L91 coupling with the atmospheric scattering
L98 are defined as a prior knowledge
L105 considerable amount number of observations, such as those
L109 a large amount number of observations
L114 easily applied on to observations
L121 should actually only be applied on to the entire phase function Φ only. These
L131 Errant space after 'distribution'
L133 vary essentially mostly as a
L175 computes separately the contributions from single and multiple scattering separately, the
L194 represented by an external
Fig4 radiatively coupled with the
p10 L4 These The different vertices, representing fine and coarse mode aerosols, are
L219 As can be seen, in most of the bands the relative
L230 where most likely the atmospheric properties likely remain unchanged
L238 which increases the number
L241 thickness τ_V of for the respective aerosol vertices that are mixed in layer L_a used. Prior information
L242 consists of expected values x_b of the state variables x_b characterising
L243 regularization on of the spectral
L247 will be further herein referred to
L279 state variables, such as the
L280 weight, etc, are
L317 showing thereby therefore that the
L321 differs from the ones of these
L323 wavelength is actually limited
L332 as the straight difference
L351 Units shouldn't be italicised.
L377 The errors ε_τ is in this experiment F12 are further reduced with respect compared to
L378 manages however to
L405 allowing a continuous variation
L406 having the parameters describing the aerosol
L480 This reference is missing the page number. Also, the DOI has a space in it for some reason.
L512 This DOI also has a space.