Dear Editor,

We would like to thank you and the two reviewers for your constructive comments and suggestions to improve the clarity of our manuscript. We have made changes to address these comments and suggestions. The following are the main changes:

1. More information has been added to the abstract to increase clarity, quantify the results better and summarize the comparison between the MF and OE methods.
2. Details have been added about the treatment of aerosols and surface albedo.
3. Explanations are supplied for the behavior of the retrievals as a function of the different parameters.
4. More quantification is provided for the differences between the two retrieval techniques.
5. Effects of changing the *a priori*, *a priori* error and simulation spectral resolution are described.
6. New Tables and Figures have been added to provide more detail.

Point-by-point responses to the comments are provided below. The reviewer comments are in blue, our responses are in red (line numbers refer to those in the revised manuscript), and modifications to the original manuscript are highlighted in yellow.

Vijay Natraj
On behalf of all co-authors

This paper provides some analysis about how aerosols properties affect CH_4 retrieval, which will attract a lot of interests from the audience of this journal. However, it is suggested that more specific analysis about the aerosol model are needed and the main points about aerosol impact need to be emphasized in both abstract and main part.

We thank the reviewer for the excellent suggestion. We have added some sentences (lines 27–30) in the abstract.

The presence of aerosols causes an underestimation of CH_4 in both the MF and OE retrievals; the biases increase with increasing surface albedo and aerosol optical depth (AOD). Aerosol types with high single scattering albedo and low asymmetry parameter (such as water soluble aerosols) induce large biases in the retrieval.

We added Tables 1 and 2 and provided a description of the aerosol models (lines 173–178).

Table 1 lists optical properties for four basic aerosol types (dust, water soluble, oceanic and soot). Table 2 shows the corresponding properties for three aerosol models that are defined as mixtures of the basic components from Table 1 (WCRP, 1986). We employ the

Henyey-Greenstein phase function (Henyey and Greenstein, 1941), where aerosol composition is determined by two parameters: single scattering albedo (SSA) and asymmetry parameter (g).

We also added more description of the aerosol impact in the main text (lines 305–312). Further, we added Figures 7b and 7c.

Since the retrieval bias is large for high SSA and low g, the water-soluble aerosol type (Table 1) and the maritime aerosol model (Table 2) can be expected to induce greater biases in the retrieval. In order to compare the impacts of SSA and g in further detail, retrieval results due to a ± 5% change in SSA and g for the three aerosol models from Table 2 are shown in Figures 7b and 7c. Note that for the maritime aerosol model, the SSA is set to 0.999 for the +5% scenario to ensure physicality. It is clear that (1) the maritime aerosol model induces larger retrieval biases than the other aerosol types, and (2) the retrieval results are more sensitive to changes in g than those in SSA.

Moreover, in the two retrieval algorithm used in this study, no aerosol loading is included. I'm just wondering if AOD or other aerosol parameters are retrieved simultaneously with XCH_4, such as adding AOD in the state vector of OE retrieval, will the retrieval bias be improved? If any preliminary results could be shown, it will be interesting.

We appreciate the reviewer's suggestion. The issue is that the MF method does not permit retrieval of AOD; it has traditionally been intended to provide a quick detection of CH_4. The OE method, on the other hand, is more flexible and does allow aerosol retrieval. We did not add AOD to the state vector since one of the methods was incapable of handling it, and we would not be able to make a meaningful comparison. In this work, we instead study the aerosol impact indirectly, by including it in the simulations but not in the retrieval. Through this process, we demonstrate that the MF method has larger biases for diffuse sources. We indicate in the abstract (lines 24–27) that the AOD is not included in the state vector.

Using a numerically efficient two-stream-exact-single-scattering radiative transfer model, we also simulate AVIRIS-NG measurements for different scenarios and quantify the impact of aerosol scattering in the two retrieval schemes by including aerosols in the simulations but not in the retrievals.

The reviewer makes an important point, though. We modify/add the following sentences (lines 344–351) indicating that the MF method is also not optimal for scenarios with aerosol scattering.

For scenarios where scattering is ignored, the two retrieval techniques seem to be complementary, with differing utilities for different enhancements. On the other hand, when RT models that account for scattering are employed, the MF technique is suboptimal. Further, MF retrievals rely on accurate characterization of the surface albedo, especially when the aerosol loading is large. Finally, the MF method does not retrieve concentrations, which are necessary to infer fluxes. Therefore, the OE technique is in general superior due to its ability to support simultaneous retrieval of aerosols, surface albedo and CH_4 concentration.

We also add two sentences (lines 37–40) in the abstract and add/modify some sentences (lines 362–364, 369–371) in the summary.

However, when aerosol scattering is significant, the OE method is superior since it provides a means to reduce biases by simultaneously retrieving AOD, surface albedo and CH_4. The results indicate that, while the MF method is good for plume detection, the OE method should be employed to quantify CH_4 concentrations, especially in the presence of aerosol scattering.

The MF method shows smaller bias ratios at large CH_4 concentrations than the OE method; it is, therefore, the optimal method to detect strong CH_4 emission sources when scattering effects can be ignored in the retrieval.

Therefore, when scattering effects need to be considered, the OE method is the appropriate choice. Indeed, the MF method was intended for plume detection. OE enables accurate quantification of XCH_4 in the presence of aerosol scattering.

Furthermore, the section 3 has less close relationship with the topic of this paper, the authors are suggested to think it more.

We feel that this section belongs in the paper. Section 3 provides a comparison of MF and OE retrievals from a real AVIRIS-NG measurement. These results provide heuristic information about the relative performance of the two techniques. However, there are some difficulties in comparing these retrievals. Further, understanding the retrieval effects of ignoring aerosol scattering is easier when we employ simulations. Therefore, both methods of comparison are useful and illustrative. We add/modify the following sentences (lines 225–231):

While these results provide heuristic information about the relative performance of the two retrieval techniques, it is difficult to compare the CH_4 enhancement directly between the two methods since the background CH_4 concentration used in the MF method cannot be quantified exactly. Further, evaluating retrieval biases due to ignoring aerosol scattering is not trivial when real measurements are used. Therefore, we simulate synthetic spectra (see section 4) using the 2S-ESS RT model to study the impacts of aerosol scattering as a function of different geophysical parameters by varying them in a systematic manner.

Specific comments

1. In the third paragraph of Introduction, I suggest the authors to add more the description about how to retrieve CH_4 concentration from satellite measurements, especially the advantage of hyperspectral imaging in CH_4 retrieval. I think the description about atmospheric correction has less relationship with the topic of this paper.

The reviewer is right. The description of atmospheric correction did not flow well with the rest of the introduction. We removed that paragraph and added a paragraph on satellite retrieval of CH_4 concentrations (lines 68–84).

Satellite monitoring of CH_4 can be broadly divided into three categories: solar backscatter, thermal emission and lidar (Jacob et al., 2016). The first solar backscattering mission was SCIAMACHY (Frankenberg et al., 2006), which was operational from 2003–2012 and observed the entire planet once every seven days. It was followed by GOSAT in 2009 (Kuze et al., 2016), and subsequently the next generation GOSAT-2 in 2018 (Glumb et al., 2014). In between, the TROPOMI mission was also launched in 2017, which observes the planet once daily with a high spatial resolution of 7×7 km^2 (Butz et al., 2012; Veefkind et al., 2012). CarbonSat (Buchwitz et al., 2013) is another proposed mission to measure CH_4 globally from solar backscatter with a very fine spatial resolution (2×2 km^2) and high precision (0.4%). Thermal infrared observations of CH_4 are available from the IMG (Clerbaux et al., 2003), AIRS (Xiong et al., 2008), TES (Worden et al., 2012), IASI (Xiong et al., 2013), and CrIS (Gambacorta et al., 2016) instruments. These instruments provide day/night measurements at spatial resolutions ranging from 5×8 km^2 (TES) to 45×45 km^2 (AIRS). GEO-CAPE (Fishman et al., 2012), GeoFTS (Xi et al., 2015), G3E (Butz et al., 2015), and GeoCarb (Polonsky et al., 2014) are proposed geostationary instruments (GeoCarb was selected by NASA under the Earth Venture - Mission program), which when operational will have resolutions of 2–5 km over regional scales. The MERLIN lidar instrument (Kiemle et al., 2014) scheduled for launch in 2021 will measure CH_4 by employing a differential absorption lidar.

2. Line 171-172: How to do normalization for measured radiance? Add some description about this, please.

We have added some description to explain how and why the normalization is done (lines 197–199).

The normalization is done by calculating the ratio of the radiance to the maximum value across the spectral range, such that the values fall between 0 and 1. This is a first order correction for the effects of surface albedo.

3. Line 181: Is the typical XCH4 background of 1.822 ppm shown by the authors here related to the background covariance matrix and mean radiance used in MF method? Some reasons are expected here. By the way, it is better to mention the background covariance matrix and mean radiance in MF retrieval of CH4 plume case here.

For a real AVIRIS-NG measurement, the mean radiance and background covariance matrix used in the MF method are taken from a reference region close to the CH_4 plume source. For the OE method in this case, a typical XCH_4 background of 1.822 ppm is used. This typical value is obtained from annual mean data tabulated by the NOAA Global Monitoring Laboratory. For synthetic MF retrievals, we compute the mean radiance and background covariance matrix using simulations for the same typical XCH_4 background of 1.822 ppm at different values of the surface albedo. We indicate how we obtained the typical value for the OE model in lines 208–211.

In the OE method, results are shown as a multiplicative scaling factor compared to a typical XCH_4 background of 1.822 ppm. This value is the globally averaged marine surface annual mean for 2014 (Ed Dlugokencky, NOAA/GML (www.esrl.noaa.gov/gmd/ccgg/trends_ch4/), the year corresponding to the AVIRIS-NG measurement being studied.

4. In the OE retrieval in section 3, what is the definition of the a priori value of XCH4? What aerosol model do the authors use? I think some parameters about aerosol model are expected here.

We have added the following description (lines 170–178) to provide the information the reviewer requested:

The *a priori* values are within 10% of the true values; *a priori* errors are assumed to be 20% for all state vector elements. The retrieved results are shown as the column averaged mixing ratio (XCH_4, in ppm). Aerosols are not included in the state vector for both the real and synthetic retrievals. They are, however, considered in the forward model for the synthetic simulations. Table 1 lists optical properties for four basic aerosol types (dust, water soluble, oceanic and soot). Table 2 shows the corresponding properties for three aerosol models that are defined as mixtures of the basic components from Table 1 (WCRP, 1986). We employ the Henyey-Greenstein phase function (Henyey and Greenstein, 1941), where aerosol composition is determined by two parameters: single scattering albedo (SSA) and asymmetry parameter (g).

5. In section 4.3, the authors show the variation of OE XCH4 retrieval bias with SSA, g, AOD, surface albedo and XCH4. Which parameters affect XCH4 retrieval bias most? From aerosol parameters, which type of aerosols, such as smoke, dust or sea salt, causes the largest or lowest bias in XCH4 retrieval? These information will attract the audience's interest and provide guidance to correct aerosol impact in future XCH4 retrieval algorithm.

We have added the following sentences (lines 302–312) to describe the effects of SSA and *g* on the retrieval. We also added Figures 7b and 7c.

This behavior can be explained as follows. At higher SSA values, there are more multiple scattering effects (that are ignored in the retrieval). On the other hand, larger values of *g* imply greater anisotropy of scattering (preference for forward scattering), leading to a reduction in multiple scattering effects. Since the retrieval bias is large for high SSA and low *g*, the water-soluble aerosol type (Table 1) and the maritime aerosol model (Table 2) can be expected to induce greater biases in the retrieval. In order to compare the impacts of SSA and *g* in further detail, retrieval results due to a ± 5% change in SSA and *g* for the three aerosol models from Table 2 are shown in Figures 7b and 7c. Note that for the maritime aerosol model, the SSA is set to 0.999 for the +5% scenario to ensure physicality. It is clear that (1) the maritime aerosol model induces larger retrieval biases than the other aerosol types, and (2) the retrieval results are more sensitive to changes in *g* than those in SSA.

The AOD and surface albedo impacts are compared in lines 321–324.

The CH_4 bias induced by differences in the surface albedo is not as large as that due to AOD variations, but albedo effects are noticeable at large AOD. Figure 8d shows the sensitivity of retrieval biases to changes in AOD and surface albedo, again demonstrating the greater impact of AOD than surface albedo in the retrieval.

6. In OE retrieval, the a priori error of XCH4 will affect the retrieval bias as well. Maybe the authors could check its impact.

The effect of changing the *a priori* error is described in lines 330–333.

Similarly, the XCH_4 difference is less than 4 ppb when the *a priori* error changes from 0.05 to 0.5 (Figure 9b). Compared to the bias of about 923 ppb induced by neglecting aerosol scattering for this scenario, it is clear that the impacts of the *a priori* and *a priori* error are very small.

Technical corrections

1. Figure 9a and 9b have some overlaps with the same XCH4. There is no need to express them using two figures.

Our objective is to show two regimes, one where the OE method has lower bias ratio and the other where the MF method performs better. There is a crossover region between ~1.5–2 where both methods produce similar results. We believe that using two figures shows this behavior clearly. Some of the details might be lost if they were combined into one.