Reply to the anonymous referee 2

The authors thank the anonymous reviewer for his comments and suggestions to improve the paper. Please find hereafter our point-by-point responses to comments and suggested corrections. We also acknowledge him deeply for his carefully reading and the time spent to list all the point concerning grammatical and vocabulary errors. It increases a lot the readability of the paper.

- There are other studies related to this one by A. Stap et al. They have investigated the errors due to cloud heterogeneity on aerosol retrieval algorithms for partially cloudy scenes, also developed for the POLDER radiometer. These could be mentioned in the introduction. F. A. Stap, O. P. Hasekamp, C. Emde, and T. Röckmann. Multiangle photopolarimetric aerosol retrievals in the vicinity of clouds: Synthetic study based on a large eddy simulation. Journal of Geophysical Research: Atmospheres, 121(21):12914-12935, 2016. 2016JD024787.
F.A. Stap, O.P. Hasekamp, C. Emde, and T. Röckmann. Influence of 3D effects on 1D aerosol retrievals in synthetic, partially clouded scenes. J. Quant. Spectrosc. Radiat. Transfer, 170:54 - 68, 2016.

Thank you suggesting these very interesting publications. We add them in the introduction:
In case of partial cloudy scenes, shadow, cloud enhancement of the clear areas by neighboring clouds can modify the retrieved aerosol properties. Errors on the retrieved aerosol properties are dependent of the cloud distribution, optical thickness and spatial resolution (Stap et al., 2016a; Stap et al., 2016b).

And in the conclusion section :
Further that assessments of cloud heterogeneity uncertainties, more complex methods should also be developed to retrieve aerosol and cloud properties accounting for the cloud heterogeneities. Several theoretical or case studies have already been conducted. Some tends to mitigate cloud contamination for aerosol property retrieval (Davis et al., 2013; Stap et al., 2016b).

Since the retrieval errors due to cloud heterogeneity are large, the conclusion of the study should be that one should develop new retrieval algorithms, which somehow consider cloud heterogeneity. I miss this conclusion in the introduction and/or conclusion section.
Steps in this directions are presented in the following papers:
W. Martin, B. Cairns, G. Bal, Adjoint methods for adjusting three-dimensional atmosphere and surface properties to fit multi-angle/multi-pixel polarimetric measurements, J. Quant. Spectrosc. Radiat. Transfer 144 (2014) 68–85 doi:10.1016/j.jqsrt.2014.03.030
W. G. Martin, O. P. Hasekamp, A demonstration of adjoint methods for multidimensional remote sensing of the atmosphere and surface, J. Quant. Spectrosc. Radiat. Trans- fer 204 (Supplement C) (2018) 215 – 231 doi:10.1016/j.jqsrt.2017.09.031
A. Levis, A. Aides, Y. Y. Schechner, and A. B. Davis, Airborne Three-Dimensional Cloud Tomography. In Proceedings of the IEEE International Conference on Computer Vision 2015 (ICCV15), pp. 3379-3387 (201 5). Available online at: http://www.cvfoundation. org/openaccess/content_iccv_2015/html/Levis_Airborne_Three-Dimensional_Cloud_ICCV_2015_paper.html
Levis, Y. Y. Schechner, and A. B. Davis, Multiple-Scattering Microphysics Tomography. In Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR17). Available online at:
http://openaccess.thecvf.com/content_cvpr_2017/papers/Levis_Multiple, Scattering_Microphysics_Tomography_CVPR_2017_paper.pdf

We agree with the comment and add a paragraph in the conclusion:

Further that assessments of cloud heterogeneity uncertainties, more complex methods should also be developed to retrieve aerosol and cloud properties accounting for the cloud heterogeneities. Several theoretical or case studies have already been conducted. Some tends to mitigate cloud contamination for aerosol property retrieval (Davis et al., 2013; Stap et al., 2016b). Others aim to use 3D radiative transfer model to retrieve 3D cloud properties and hence account for some cloud heterogeneity effects. It requires then more complex inversion methods. Feasibility studies has been conducted using neural network method (Cornet et al., 2004, 2005), 3D tomography with a surrogate function (Levis et al., 2015, Levis et al. 2017) or adjoint method (Martin et al., 2014; Martin and Hasekamp, 2018). The latter two methods are very promising but have been developed in the framework of high resolution measurements (ten to hundred meters) involving no or small plane-parallel bias. They are so not directly applicable to POLDER/PARASOL measurements.

.

The core of the study, the 3D radiative transfer (RT) model 3DMCPOL, is not described (only reference Cornel et al. 2010 is given). There should be a brief description on which methodology is used to solve the vector radiative transfer equation and also on the accuracy. Also later, in the results section, it is not mentioned, how accurate the radiative transfer simulations are. Can we trust the RT results, has the model been validated? The first paragraph in section 2 provides a short description of the cloud model 3DCLOUD; I would expect a similar description for 3DMCPOL.

For the description of the 3DMCOL model, we add in section 2:
It is a forward Monte-Carlo model able to compute radiative reflected or transmitted Stokes vector as well as upwelling and downwelling fluxes in three-dimensional atmospheres. Initially, developed for solar radiation (Cornet et al., 2010), it was next extended to thermal radiation (Fauchez et al., 2014). To save time and for an accurate computation of reflectances, the local estimate method (Marshak and Davis, 2005) is used. Periodical boundary conditions at the horizontal domain limits are used. For highly peaked phase function, the potter truncation is implemented. Molecular scattering is computed according to the pressure profile. A heterogeneous surface can also be specified with Lambertian reflection, ocean or snow bidirectional function. The model participated and was improved during the Intercomparison of Polarized Radiative Transfer model (IPRT) on homogeneous cloud cases (Emde et al., 2015) and on 3D cloud cases (Emde et al., 2018).

Concerning the accuracy of the computations used is the paper, we add:
Simulations are run with a total of 10^7 photons and 10^9 photons for the homogeneous and heterogeneous clouds respectively. The Monte-Carlo uncertainties are estimated with the computation of standard deviation with 10 and 50 independent realizations of 10^6 and 20.10^6 photons for the homogeneous and heterogeneous cloud respectively. For the homogeneous case, the relative standard deviation is below 0.12% for the total reflectances and below 1.2% for the polarized reflectances. For the heterogeneous clouds, at 50m resolution, the mean relative standard deviation is below 1.3% for the total reflectances. For polarized reflectances at 50m, the mean relative standard deviation varies according to the angular geometry and is between 2% and 107% for very small reflectance values with an mean value of 23%. At 7km, as the reflectances are averaged, relative standard deviation values are much lower below 0.01% and 0.8% for total and polarized reflectances respectively.

p5, l4: "To remain consistent with assumptions made within POLDER operational algorithm, an oceanic surface with a wind speed of 7 m.s-1 is included for total reflectances while a black surface is included for polarized reflectances." -> This is an odd assumption. I think that this could introduce large errors, because the sun-glint is highly polarized. Why is the surface inconsistently included in the POLDER operational algorithm? Is there any document where this assumption is justified. Please explain/discuss this issue.

As already discussed in the reply to reviewer 1, the operational algorithm using polarized reflectances assumes a black surface because the multi-angularity of POLDER allow to not use the directions close to the sun glint where polarized reflectances can be high. In the other directions the polarized ocean surface reflection is almost null (black).

We add page 7 :
Indeed, for retrieval using polarized reflectances, the multi-angular ability of POLDER provides the advantage of not using the directions close to the sun-glint where polarized reflectances can be high.

p5, l17: "Note that in the three cases, the operational algorithm retrieves a cloud cover equal to one." -> can the operational algorithm retrieve cloud cover different from one? If yes, why does it not work for the fractional cloud?

The cloud cover is an output of the algorithm for the super pixel POLDER but you are right as the pixel level the value can only be zero or one. We removed the sentence

- p.6, l1: "That confirms that heterogeneity parameters can be at first order used to characterize plan-parallel bias" -> could the heterogeneity parameter be derived from observations?

The heterogeneity parameter cannot directly be obtained for one reflectance measurement but it may be estimated from higher spatial resolution measurements. This is the idea of the sentence wrote in the conclusion section:
"The Multi-viewing, Multi-Channel, Multi-Polarization Imaging mission (3MI) that will fly on METOP-A SG as part of EUMETSAT Polar System after 2021, will have a spatial resolution of 4 x 4 km. The plane-parallel bias is thus expected to be lower than for the POLDER instrument. In addition, as 3MI will be on the same platform as the Visible Infrared Imager (VII), a multispectral radiometer with a resolution of 500 m, the correction of the plane parallel biases may be possible while the multi-angular capability of 3MI would help to detect the illumination and shadowing effects."

p.6, l31: "Contrarily, using 1D cloud radiative model in the inversion and in the direct computation as it is done in the operational algorithm, is coherent and leads to a sound cloud albedo. The plane-parallel bias is indeed almost canceled." This sounds as if the operational algorithm would retrieve a good cloud albedo, but it does of course not. The reality always "uses" a 3D radiative transfer model, so retrieval algorithms based on 1D RT models are always inconsistent and yield wrong results.

Using an homogeneous cloud model for the cloud optical thickness retrieval from real or 3D reflectances and also for the computation of the cloud albedo almost cancel the plan-parallel bias effect. The residual error is due to the non-linearity degree of the reflectances/albedo as a function of the cloud optical thickness and to the 3D effects such as illumination, shadowing or even smoothing effects for high resolution. The reviewer 1 also found this paragraph unclear, we rephrased it hoping to be clearer:
The assessment of cloud heterogeneity effects on cloud albedo is realized by comparing the retrieved POLDER algorithm albedos with the ones directly computed with the 3DMCPOL radiative transfer model identified as the true one. Direct comparison of retrieved albedos values from homogeneous or from the heterogenous clouds as done for other parameters are not suitable for cloud albedo. Indeed, the plane-parallel bias leads to reflectances off of a heterogenous cloud lower than the reflectances off of an equivalent homogenous cloud with the same (mean) COT. The retrieved optical thickness is lower than the mean optical thickness of 10 (Figure 4). Using it to recompute the albedo in the POLDER algorithm leads to a too low value comparing to the albedo of the equivalent homogeneous cloud. Contrarily, using 1D cloud radiative model in the inversion and in the direct computation as it is done in the operational algorithm, is consistent and leads to a sound cloud albedo. The plane-parallel bias is indeed almost cancelled.

p7, l1: "Albedos are simulated simply by summing the proportion of the Monte-Carlo photons going up at the top of atmosphere." -> This is then not the cloud albedo but the total albedo,

since it includes also contributions from molecular scattering and surface reflection, right?

Good point, it is indeed a misnomer. The total albedo including molecular scattering, cloud scattering and surface reflection is indeed computed. It is done in the same way for the LUT used in the POLDER algorithm (Buriez et al., 2005). We replaced *cloud albedo* by the terms albedo of a cloudy scene or albedo only.

p8, Sec4.2: The effective variance retrieval uses the amplitude of the surnumerary bows. The aerosol above cloud retrieval (Sec 4.3) obtains information about AOT from the attenuation of the cloud bow. If effective variance and AOT above cloud both influence the amplitude of the cloudbow region, how does the retrieval distinguish between higher AOT and narrower size distribution? Does the amplitude also depend on cloud optical thickness?

The POLDER "operational algorithm" for aerosol above cloud retrieval uses a specific retrieval strategy. The cloud bow is indeed used for above cloud aerosol retrievals only in case of dust particles above clouds. The magnitude of the primary cloud bow primarily depends on the cloud droplet effective radius and this parameter must be also estimated. Collocated cloud properties from MODIS at high resolution (1 km × 1 km) are used to characterize and to select the cloudy scenes within a POLDER pixel (6 km × 6 km at nadir). The MODIS cloud products are notably used in the POLDER "operational algorithm" to estimate the droplets effective radius. The magnitude of the primary cloud bow is only weakly impacted by the choice of the droplets effective variance and this parameter is then fixed to 0.06 in the "operational algorithm".

We added this paragraph in the manuscript :
The magnitude of the primary cloud bow primarily depends on the cloud droplet effective radius and this parameter must be also estimated or included in the retrieval process. Collocated cloud properties from MODIS at high resolution (1 km × 1 km) are used to characterize and to select the cloudy scenes within a POLDER pixel (6 km × 6 km at nadir). The MODIS cloud products are notably used in the "operational algorithm" to estimate the droplets effective radius. The magnitude of the primary cloud bow is only weakly impacted by the choice of the droplets effective variance and this parameter is then fixed to 0.06 in the "operational algorithm".

and this information concerning the test realized :
Note, that for the synthetic retrievals discussed here below, we assumed that the operational algorithm knows the effective radius and effective variance of the cloud droplets.

For cloud optical thickness larger than 3, the amplitude of the cloud bow does not depend on the cloud optical thickness. We added this precision in the manuscript :
The retrievals are restricted to cloudy pixels associated with cloud optical thicknesses larger than 3.0, since the polarized radiance reflected by the cloud layer is then saturated and does not depend anymore on the cloud optical thickness.

Sec. 4.2: Is the optimal estimation method a good approach for Reff/Veff retrieval based on the polarization of the cloudbow region? You write that the radiance does not fit very well, so that the retrieval does not converge, although the retrieval of the size distribution parameters is very accurate. I would think that the retrieval should not minimize the fit to radiances but it should only fit the position of the cloudbow and its amplitude. This could be realized using an optimal estimation approach but may be a simple lookup-table method would also work well. Somehow the retrieval should provide a criterion, whether it provides good results or not, here the cost function is not a good number for the quality of the retrieval.

Beside the computation cost, the optimal estimation approach was chosen because of its flexibility. We want to keep this in order to have the freedom of adding new measurements or parameters in the state vector (like a second scattering layer above cloud for example).

We agree with the reviewer, that with a large sampling, a retrieval using only the position of the cloud bow and surnumerary bow would be much powerful than the absolute polarized radiance. However, one difficulty with POLDER/PARASOL measurements is that, because of the angular sampling, we never get the exact position of the maximum. A small error in the position of the maximum turns in a very large error in the effective radius. Therefore this « maximum position method » might give worst results than using the absolute polarized radiance.
The cost function is just an indicator of the goodness of the convergence within the errors provided by the measurements and forward model, and is also used as a criteria to stop the iteration process. Because the cost function is a sum of the square of standard normal variables, and because we have assumed that the conditional probability function of the measurements knowing the true state vector follows a normal distribution, the cost function follow a Chi-square law. We can therefore use this low together with a hypothesis testing to determine whether the weighed distance between the forward model and measurement is acceptable for a given confidence. This is just a statistical criteria which is working pretty well. A good indicator of the quality of the retrieval is always difficult to define, but the cost function at least can help when something went wrong in the retrieval, and especially when the forward model is not able, because it is too simple, to reproduce the measurements behavior (in the presence of highly heterogeneous cloud, or in the presence of an aerosol/cirrus layer above the liquid cloud).

We completed the sentence:
For all clouds, even if differences in polarized reflectances are large in amplitude, the retrieval algorithm still capture the general angular features of the three wavelengths, which results of small errors on the retrieved effective radius and effective variance.

And add concerning the cost function:
It means that the forward model (homogeneous model) used for the retrieval does not allow matching perfectly the heterogeneous cloud reflectances used as input.

- Sec. 4.2: "For the misrepresentation of 3D effects, we add 7.5% error in the cloudbow direction and 5% elsewhere." -> how are these errors estimated? Please justify.

These errors were estimated in previous work (Waquet et al., 2013) with the computation of 3-D and 1-D polarized radiances of a stratocumulus cloud close to the flat cloud presented here. Excepted for reflectances close to zero, relative errors were under 5-8%. We add the reference in the text as previous computations made in (Waquet et al., 2013)

- Table 3: I can not believe that for SZA=40_ the difference between true and retrieved AOT and Angstroem coefficient (here also SZA=20_) is exactly 0.0 (with 3 digits accuracy).
Please explain why it is exactly the same.

We checked the results and they are good. The rapid algorithm used for operational retrieval is based on precomputed tables. In the two cases, homogeneous and fractional cloud, the best model that minimized the cost function is the same so we obtain the same AOT. However, the cost function is more important for the heterogeneous cloud. We add the RMSE value between the input and the recalculated reflectances the table 3 and this sentence.
For SZA=40°, the best model that minimized the cost function is the same for the homogeneous and fractional cloud. Differences for the retrieved AOT are negligible, but we note that the RMSE between the input and recalculated reflectances is slightly larger for the fractional cloud than for the homogenous one.

Technical corrections:
Text has been modified according to the technical correction addressed by the reviewer that we would like to thank again. See the track changes file for the details..