Response to the reviews of manuscript amt-2015-338.

I am deeply grateful to the reviewers for reading this very long manuscript.

Its lengthiness and the explicit derivation of many scalar equations where matrix equations would be much more concise is the main objection of reviewers 1 and 2.

There are several reasons why I think this elaborateness is necessary.

Most of the lidar papers lack a thorough error analysis, especially regarding systematic errors. Why is that, and how can it be changed?

As can be seen from this manuscript, which deals only with polarisation related uncertainties, there are many different error sources, and it is laborious to consider them all together. I have not seen a complete error analysis of lidar measurements and their products published yet. If at all, uncertainties due to signal and background noise are processed, and in only few cases one or the other systematic uncertainty is included. The description how the error bars are achieved is often insufficient and important error sources are often neglected. But we urgently need verifiable error bars which reflect the true uncertainty without systematic biases, if we want our data to be assimilated in weather models and used in micro-physical retrieval algorithms, whose outcomes strongly depend on the uncertainty of the input data.

The persons which are able to prepare the error analysis of a lidar system are the ones which built it, because only these persons can be aware of all the error sources which have to be included. But these persons have enough troubles to deal with the instrumental issues and are often under pressure of time by the schedule of the project which funds the lidar or by the limited time span for their PhD study, which must present some new, exiting atmospheric measurements at the end. Usually, there is no time left for a thorough error analysis, which would cost maybe another year, and for which there would be only little appreciation value.

On the other hand, many error sources could already be eliminated in the design of the lidar systems, but concise matrix formulas, as proposed by the reviewers, don't reveal the influence of the individual elements and of the parameters of the matrices. If we want to estimate the importance of a certain parameter either for system design or error estimation, or to compare different system or calibration setups, it is necessary to expand the concise matrix equations as it is done in this manuscript, hopefully saving many colleagues a lot of time and mistakes.

There are several papers dealing with one or the other error source, and much could be learned from publications in the field of ellipsometry. But these publications are not only in another "language", but in many different "languages", with different assumptions, different definitions and conventions, different neglects also, which takes a lot of time to sort out what can be used. And then, finally, it takes again some time to write an error analysis code and to verify that it really works correctly.

To my opinion, concise matrix equations don't help those which actually have to run the lidar systems and do the error analysis.

I admit that such a lengthy manuscript as present here is a burden as well, but it tries to cover the different variants of the most common type of lidar systems with different calibration techniques, and it is also meant as a starting point for further developments in the direction of a complete error analysis and a related open source software. Therefore I want to present all definitions, derivations, and variants in one paper in such a way, that misunderstandings and mistakes are less probable. The manuscript is also meant as a reference for future papers, so that the derivations don't have to be repeated again and again with different definitions. The companion paper "*Assessment of lidar depolarization uncertainty by means of a polarimetric lidar simulator* " by J. A. Bravo-Aranda et al. (doi:10.5194/amt-2015-339) shows already an application.

Another critique is, that we should "let the computer do the hard part" and use software which can deal with matrix calculations. The original Monte-Carlo code in the above mentioned companion paper was written with such a software, and it took orders of magnitude longer than the scalar code we wrote for verification. And this is only the beginning, because we have to include more parameters for a complete error analysis and make it somehow operational. The error debugging of these two codes showed, that a detailed reference for the used equations is necessary.

Most of the above is already mentioned in the introduction of the manuscript.

Detailed answers to Anonymous Referee #2:

This analysis includes consideration of different polarization calibration schemes that, as far as I can tell, the author has previously published.

There is only one publication (Freudenthaler et al., 2009), which presents the basic idea of the Δ90-calibration in a very reduced form compared to this manuscript, and only for the rotation calibrator. Section 2 of this manuscript starts from the concept of that paper, and transfers the pure scalar concept there into the more general context of the Müller-Stokes formalism with adapted terminology, which serves as an introduction to the following sections.

It belabors a number of topics that have been thoroughly addressed in other published works while sometimes completely ignoring other aspects, creating the false impression of a comprehensive analysis.

The published works about these topics I am aware of are mentioned in the introduction. But none of those includes all the calibration techniques and all the error sources included in this manuscript. This manuscript describes several techniques in the same mathematical framework and compares the different advantages and disadvantages.

Already in the introduction I present the general setup of the sort of lidar systems which is covered by the presented model. It covers at least 80% of the lidar systems in EARLINET, and I guess this holds for other networks. I don't see the necessity to explicitly exclude everything which is not covered.

For all the meticulous steps in this section, there is absolutely no mention of oriented particles or multiple scattering. In my view, the reader may be given the false impression that this scattering matrix is somehow comprehensive, which it clearly is not, since it only covers single scattering by randomly oriented particles.

I am thankful for this hint. I include Section 13 (see below), which lists the assumptions and constraints of this model:

Assumptions and constraints of the model

1. The correction of the standard signals (Sect. 4) and of the calibration factor (Sects. 5 ff) are only applicable in scattering ranges without aerosol or with randomly oriented, non-spherical particles with rotation and reflection symmetry as described in Sect. 2.1, and not for clouds with oriented particles as in cirrus and rain clouds (Kaul et al. (2004); Hayman et al. (2014); Volkov et al. (2015)). However, the scattering volume for the calibration measurements can be chosen to avoid oriented particles in the calibration range, and then the calibration corrections in Sects. 5 to 10 can be applied for the retrieval of the calibration factor $\eta = (\eta_R\ T_R) / (\eta_T\ T_T)$, which itself is general for the considered types of lidar set-ups in Fig. 1.

2. We assume that the extinction in the range between the lidar and the scattering volume is polarisation independent and that signal contributions due to multiple scattering can be neglected.

3. We assume that the atmospheric depolarisation in the calibration range does not change between the two measurements of the Δ90-calibrations. This can be verified by comparison of standard measurements before and after and maybe even between the two calibration measurements.

4. Not considered are range dependent effects as the overlap function and the range dependent transmission and polarisation of interference filters and dichroic beam-splitters, which is caused by the range dependent incident angles on the optics.

5. We assume that the optical elements of the lidar do not depolarize. Such depolarization can be caused by variable retardation or diattenuation over the aperture of optical elements, for example due to crystalline (e.g. CaF_2 and MgF_2 lenses) or stress birefringence. The latter can be present in all optical elements if they are inappropriately restrained in their holders. Larger optics, as e.g. telescope windows, can exhibit inherent stress birefringence due to annealing and/or their own weight. Such optical elements can easily be visually inspected by means of crossed polarising sheet filters before and after the sample. Furthermore, non-parallel (converging or diverging) incident beams on optics with polarisation effects depending on the incidence angle will cause depolarisation. The manufacturers specification of dedicated polarisation optics should be sufficient to determine the maximum allowable divergence of the incident beam, but, for example, the coatings of 90° reflecting mirrors in Newtonian telescopes are usually not sufficiently specified to determine their polarisation effects. The depolarising effects of optics can additionally depend on the state of polarisation of the incident beam.

It is good there is a table of variable definitions, but maybe it would be better to break them up a bit based on where they are used in the paper.

Most of the listed definitions are used in many places in the manuscript. The idea of one comprehensive list is to make the definitions easy to find in one place.

Be very clear about the assumptions applied in this work (preferably itemize them or keep them in a table).

See above, new Sect. 13.

In Eq. (14) (and others) pick a matrix form and stick to it.

The various matrix forms are quoted on purpose in order to avoid mistakes, which would certainly arise because the definitions are different in different textbooks. It is especially important to emphasize that the diattenuation parameter is used in this manuscript and not the diattenuation as in most other literature and textbooks.

There really isn't a reason to evaluate to a scalar equation unless it produces a result that is simple, concise and reveals some previously not obvious fact. That is, evaluating the matrices should produce a reduction in complexity, not the other way around.

The fully evaluated scalar equations are necessary for the complete error calculation, as described above. Furthermore, I present for each calibration setup some equations (special cases) which are reduced in complexity and show the advantages of avoiding certain error sources already in the lidar design. This analysis of "previously not obvious facts" is continued in Sects. 11 and 12.

Present the overall Mueller equation describing I_s (Eq. (64)?), then give the evaluation in Eq. (68), but get rid of all those coefficients in front, which can just as easily fold into G and H and aren't important since absolute intensity measurements are almost never used in atmospheric lidar.

The "coefficients in front" and the absolute signal intensities are important for numerical error analysis when we include signal and background noise.

I don't want to fold the coefficients into G and H, because G and H are the parameters which describe the polarization and misalignment dependencies, while the coefficients in front are the unpolarised transmittances. G and H are simply 1 and ±1, respectively, for the ideal cases. Furthermore, also for the comparison of different lidar systems in intercomparison campaigns it makes sense to keep them separate.

The explicit definitions of G and H seem unnecessary.

With G and H the general Eqs. (62) and (65) can be formulated. This is the reduction in complexity as required by the reviewer. These general equations are a superset of several equations presented in other papers considering only a part of the error sources in this manuscript. They are very helpful for a general lidar data analysis software as described, e.g., in *Mattis et al., 2016, EARLINET Single Calculus Chain – technical Part 2: Calculation of optical products* (doi:10.5194/amt-2016-43).

Drop the <bra|ket> notation.

I include after Eq. (43):

This split-up of the equations in an analyser bra-vector and an input Stokes ket-vector is similar to the split-up in instrumental vectors of the transmitter and receiver in Kaul et al. (2004) and Volkov et al. (2015).

The <bra|ket> notation makes it easy to discern between the instrumental vectors of the transmitter and receiver. It allows for an elegant restructuring of the matrix equations as for example in Eqs. (133) and (177).

I don’t really understand all this interest in obtaining “correction equations” for scalar polarization variables. The approach used in Kaul Appl. Opt. 2004, Hayman, Opt. Express 2012, Volkov Appl. Opt. 2015 and countless other polarimetry papers avoids any need to belabor “corrections” and just retrieves the relevant scattering matrix terms based on the lidar’s operational parameters”errors” or whatever you want to call them.

Kaul (2004), Hayman (2012), and Volkov (2015) determine the gain ratio with a calibration in clean air ranges, where they assume molecular depolarization. This assumption can cause large errors. Furthermore, the mentioned papers don't describe any error calculation for systematic errors.

In this manuscript I describe various versions of the more robust Δ90-calibration, which must and can be corrected for systematic errors, if they are not already avoided in the lidar design. The necessary correction equations are described in Sects. 5 to 10. A numerical error calculation code can directly be written from the given equations.

In order to make the connections between the standard measurements, the model corrections and the additional calibration measurements more clear, I included in Section 5 the following figure and corresponding text:

Figure (4) shows the steps in which the measurements are corrected for systematic errors by means of the model. If all system parameters (Eqs. (56) and (57)) are known, the cross-talk parameters G_S and H_S can be calculated (see Eqs. (68) to (79)) and we only need to determine the calibration factor η by means of calibration measurements in step 2 and its correction for systematic errors (step 3) as explained in Sects. 6 to 10. Under certain conditions some instrumental parameters can be determined by means of additional calibration measurements (step 4) described under “special cases” in Sects. 6 to 10 and in Sects. 11 and 12.

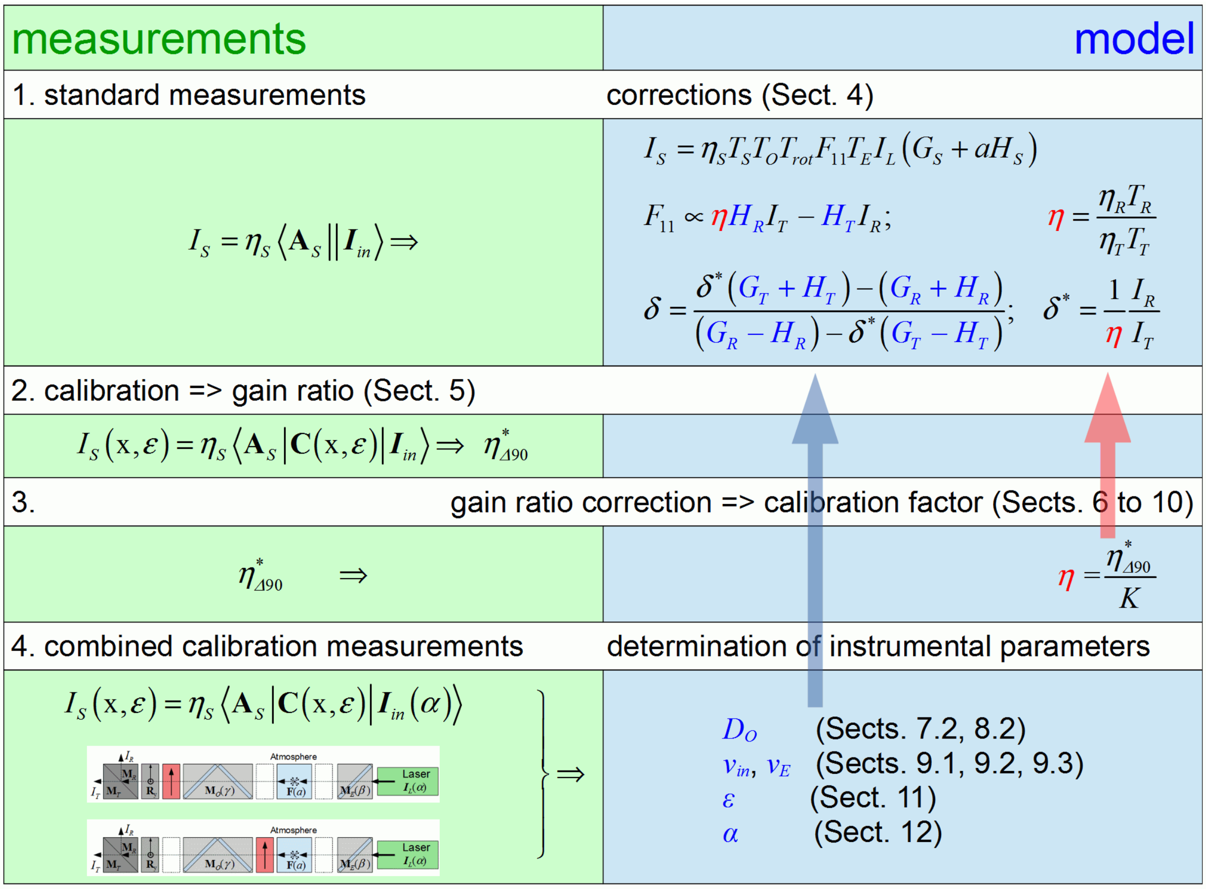

Figure 4: Four steps for calibrating and correcting the standard measurements for systematic errors by means of the model equations and additional calibration measurements.