



1 Information Content and Sensitivity of the $3\beta + 2\alpha$ Lidar 2 Measurement System for Aerosol Microphysical Retrievals

3 Sharon P. Burton¹, Eduard Chemyakin², Xu Liu¹, Kirk Knobelspiesse³, Snorre Stamnes¹, Patricia
4 Sawamura⁴, Richard H. Moore¹, Chris A. Hostetler¹, Richard A. Ferrare¹

5 ¹NASA Langley Research Center, Hampton, VA, USA

6 ²Science Systems and Applications, Inc., Hampton, VA, USA

7 ³NASA Goddard Space Flight Center, Greenbelt, MD, USA

8 ⁴Universities Space Research Association, Columbia, MD, USA

9 Email: Sharon.P.Burton@NASA.gov

10 Abstract

11 There is considerable interest in retrieving profiles of aerosol effective radius, total number concentration and complex
12 refractive index from lidar measurements of extinction and backscatter at several wavelengths. The combination of
13 three backscatter channels plus two extinction channels ($3\beta + 2\alpha$) is particularly important since it is believed to be
14 the minimum configuration necessary for the retrieval of aerosol microphysical properties and because the
15 technological readiness of lidar systems permits this configuration on both an airborne and future spaceborne
16 instrument. The 2nd-generation NASA Langley airborne High Spectral Resolution Lidar (HSRL-2) has been making
17 $3\beta + 2\alpha$ measurements since 2012. The planned NASA Aerosol-Clouds-Ecosystem (ACE) satellite mission also
18 recommends the $3\beta + 2\alpha$ combination.

19 Here we develop a deeper understanding of the information content and sensitivities of the $3\beta + 2\alpha$ system in terms
20 of aerosol microphysical parameters of interest. We use a retrieval-free methodology to determine the basic
21 sensitivities of the measurements independent of retrieval assumptions and constraints. We calculate information
22 content and uncertainty metrics using tools borrowed from the Optimal Estimation methodology based on Bayes'
23 theorem, using a simplified forward model look-up table, with no explicit inversion. The forward model is simplified
24 to represent spherical particles, monomodal log-normal size distributions, and wavelength-independent refractive
25 indices. Since we only use the forward model with no retrieval, the given simplified aerosol scenario is applicable as
26 a best case for all existing retrievals in the absence of additional constraints. Retrieval-dependent errors due to
27 mismatch between retrieval assumptions and true atmospheric aerosols are not included in this sensitivity study, nor
28 are retrieval errors that may be introduced in the inversion process.

29 The sensitivity metrics allow for identifying (1) information content of the measurements versus a priori information;
30 (2) error bars on the retrieved parameters; and (3) potential sources of cross-talk or “compensating” errors wherein
31 different retrieval parameters are not independently captured by the measurements. The results suggest that the $3\beta +$
32 2α measurement system is underdetermined with respect to the full suite of microphysical parameters considered in
33 this study, and that additional information is required, in the form of additional coincident measurements (e.g. sun-
34 photometer or polarimeter) or a priori retrieval constraints. A specific recommendation is given for addressing cross-
35 talk between effective radius and total number concentration.



1 1. Introduction

2 Aerosol effects on global and regional climate and human health depend on aerosol amount, vertical distribution, and
3 proximity to clouds, as well as the composition, size and absorption properties of the aerosol. The NASA Aerosol-
4 Cloud-Ecosystem (ACE) mission (http://dsm.gsfc.nasa.gov/ace/mission_details.html) recommended by NRC's 2007
5 Earth Decadal Study (National Research Council, 2007), is currently in pre-formulation stage, and aims to produce a
6 comprehensive set of vertically and horizontally resolved aerosol properties as a function of time and location. This
7 dataset will be used to constrain aerosol transport models and model estimates of globally averaged direct aerosol
8 radiative forcing, not just at the top of the atmosphere, but also near the surface and within the atmosphere. The mission
9 therefore addresses the quantification of (1) aerosol sources, sinks and transport, (2) direct aerosol forcing, and (3)
10 aerosol-cloud interactions (ACE Science Working Group, 2010).

11 To achieve these goals, ACE is planned to include a multi-wavelength High Spectral Resolution Lidar and a multi-
12 wavelength, multi-angle imaging polarimeter from which vertically resolved aerosol microphysical retrievals will be
13 made. While passive polarimeter measurements can provide accurate retrievals of column averaged microphysical
14 properties (Dubovik et al., 2011; Hasekamp et al., 2011), only lidar measurements can provide the vertical resolution
15 required. The combination of three backscatter and two extinction wavelengths ($3\beta + 2\alpha$) for the lidar is considered
16 to be the minimum number of channels required for an aerosol microphysical retrieval (Bockmann et al., 2005;
17 Veselovskii et al., 2002) based on a heritage of aerosol microphysical retrievals from ground-based Raman
18 measurements of varying wavelength combinations (e.g. Müller et al., 1999; Bockmann, 2001; Donovan and Carswell,
19 1997). Accordingly, the ACE plan calls for a High Spectral Resolution Lidar (HSRL) to measure the aerosol
20 backscatter coefficient at 355, 532, and 1064 nm and the aerosol extinction coefficient at 355 and 532 nm. This
21 combination is frequently referred to as " $3\beta + 2\alpha$ " lidar. The NASA Langley airborne HSRL-2 is one prototype for
22 the ACE lidar.

23 There exist various aerosol microphysics retrievals based on $3\beta + 2\alpha$ lidar measurements (e.g. Bockmann, 2001;
24 Veselovskii et al., 2002; Veselovskii et al., 2012; Chemyakin et al., 2014; Müller et al., 1999). In general, these
25 retrievals are performed for each vertically-resolved altitude level (grid-point) in the lidar profile, on a single set of
26 three backscatter and two extinction measurements at a time, with each altitude level being treated independently.

27 The Inversion with Regularization retrieval (Müller et al., 1999; Veselovskii et al., 2002) is the standard algorithm
28 used for $3\beta + 2\alpha$ retrievals. Mie theory kernels link the lidar optical measurements of aerosol backscatter and
29 extinction coefficients with aerosol size distributions, which are represented as a combination of five to eight triangular
30 basis functions. The size distribution is retrieved using inversion with regularization for a given complex refractive
31 index and set of minimum and maximum particle sizes (integration limits). The integration limits and complex
32 refractive index are then varied over a range of values that are typically found for aerosols, between 30 nm and 8 μm
33 for the integration limits, between 1.325 and 1.8 for the real part of the refractive index, and between 0 and 0.1 for the
34 imaginary refractive index. Specific solutions (sets of values for the size distribution parameterization and complex
35 refractive index) are selected based on limiting the amount of discrepancy between the measurements and the
36 backscatter and extinction coefficients reproduced from the Mie solutions. A few hundred individual solutions with
37 different integration limits and refractive indices are then averaged together to provide the mean value and error bars



1 for the final solution. The process of averaging multiple solutions together adds stability to the retrieval (Veselovskii
2 et al., 2002). The Inversion with Regularization retrieval was demonstrated with airborne HSRL-2 measurements by
3 Müller et al. (2014).

4 The Linear Estimation method (Veselovskii et al., 2012) solves for the particle size distribution represented as a linear
5 combination of the measurement kernels. Only the total integrated number concentration is retrieved rather than the
6 full size distribution. The refractive index is retrieved by iteration, solving the equation for an assumed refractive
7 index and minimizing the resulting systematic error. The systematic error to be minimized is estimated by using only
8 four of the measurements to attempt to reproduce the fifth and repeating for all five measurements. Like the Inversion
9 with Regularization technique, the final solution is an average of a family of individual solutions.

10 The Arrange and Average method (Chemyakin et al., 2014) is a simplified version of the $3\beta + 2\alpha$ retrieval which is
11 particularly helpful for experimental work in understanding retrieval behavior (Chemyakin et al., 2016). This
12 methodology makes use of a pre-computed look-up table (LUT), simplifying the exploration of the full space of
13 possible solutions. The look-up table used in this study and by Chemyakin et al. (2014) has only monomodal log-
14 normal size distributions. Since the complex refractive index is also included in the look-up table and therefore treated
15 identically to the size distribution parameters in this retrieval, all parameters are retrieved simultaneously and the
16 relationships between retrieval parameters are more transparent. Solutions are selected from the LUT that match the
17 optical measurements to within a small discrepancy.

18 While it has been demonstrated that $3\beta + 2\alpha$ lidar measurements can yield accurate retrievals of aerosol microphysical
19 parameters that agree with in situ measurements of effective radius and total integrated number concentration (Müller
20 et al., 2014), it is also understood that this retrieval is underdetermined (Veselovskii et al., 2002; Bockmann et al.,
21 2005; Pérez-Ramírez et al., 2013; Chemyakin et al., 2016). Therefore, characterization of the aerosol microphysical
22 state parameters requires additional information or constraints beyond the lidar measurements. In general, constraints
23 can take various forms, including smoothing, regularization, a priori state values, or limits on the ranges of values that
24 the retrieved parameters can take.

25 Previous studies of the $3\beta + 2\alpha$ lidar retrieval system also point out the difficulty of retrieving the complex refractive
26 index in particular (Veselovskii et al., 2012; Müller et al., 2014; Pérez-Ramírez et al., 2013). The $3\beta + 2\alpha$ retrievals
27 represent the relationship between the measured optical properties and the particle size distribution as Fredholm
28 integral equations of the first kind, with known limits of integration and known complex refractive index.
29 Consequently, the complex refractive index is generally assumed based on context, or else varied in a separate
30 minimization process (Müller et al., 1999; Veselovskii et al., 2012), which makes the retrieval performance and
31 sensitivities complicated to assess.

32 These challenges have been acknowledged and addressed in the existing retrievals, but there is still relatively little
33 published information about the true sensitivities of the $3\beta + 2\alpha$ lidar measurements with respect to aerosol
34 microphysical parameters of interest. We wish to rigorously and quantitatively deepen our understanding of the
35 information content of the retrieval system, by determining how much information in the retrieval stems from the lidar
36 measurements themselves and conversely which information is provided only by constraints or a priori information.
37 The results of this sensitivity study will also clarify how other measurements (e.g. polarimeter or sun-photometer



1 measurements) may significantly add to the information content. This study therefore supports ongoing work to
2 implement a full combined active+passive (lidar+polarimeter) vertically-resolved aerosol retrieval and to understand
3 the retrieval limitations in situations where only lidar data are available (i.e. night side of the orbit or gaps in broken
4 cloud systems). These studies will also help refine measurement requirements and determine retrieval uncertainties
5 for ACE or other future measurement systems.

6 An ideal framework for a study of retrieval sensitivity and information content is Optimal Estimation (OE). OE, based
7 on Bayesian statistics, is a formalized framework for combining measurements, measurement errors, external
8 information and constraints. Thoroughly described by Rodgers (2000), it provides a number of key tools for
9 characterizing the sensitivities and information content of a retrieval system. For example Knobelspiesse et al. (2012)
10 use the Shannon Information Content and the propagated retrieval errors to characterize the capabilities of multi-angle
11 multi-wavelength polarimeter for aerosol microphysics retrieval. Xu and Wang (2015) analyze the information
12 content of AERONET measurements with respect to aerosol microphysics retrievals using the propagated retrieval
13 errors and Degrees of Freedom (DOF) of the signal.

14 The diagnostics for sensitivities and information content in the OE framework do not depend on completing a retrieval.
15 Rather, they depend only on retrieval inputs: the forward model, measurement uncertainties, and the a priori
16 constraints. Therefore, although the lidar retrieval algorithms described above are not OE algorithms, these tools can
17 nevertheless be usefully applied to this problem to provide implementation-independent best-case sensitivity metrics.
18 Unlike a perturbation method, the strategy of performing the sensitivity study using only the forward-model allows
19 for mapping out the entire state space relatively quickly, without the need for time-consuming retrievals. In addition,
20 since the OE method is a matrix method, the measurement covariance matrix is handled as a single object, taking into
21 account measurement errors in all channels simultaneously, without requiring simplifying assumptions such as an
22 additive property (Pérez-Ramírez et al., 2013). Finally, the OE method provides a formalized means of representing
23 the retrieval constraints, a critical part of an underdetermined retrieval like this, but one which is not well represented
24 in a perturbation sensitivity study. In this study, we use a look-up table approach to simplify the forward model and
25 set the stage for a retrieval-independent study of sensitivity and information content of the $3\beta + 2\alpha$ lidar measurement
26 system with respect to a small set of aerosol microphysical parameters. While the simplifications necessarily ignore
27 some errors that would occur in a generic real-world aerosol situation, this strategy provides a transparent and rigorous
28 view of the basic sensitivities for this retrieval problem that is applicable to any retrieval with the same measurement
29 inputs, as long as the retrieval assumptions are no more restrictive than those consistent with the very simplified
30 aerosol under consideration. Note also that retrievals will also potentially include additional errors that are dependent
31 on the method used to converge to a solution, which, again, are not included in this assessment.

32 In Section 2 we describe the overall methodology for our sensitivity study and in Section 3 we describe the specific
33 cases used for illustration in this paper. In Section 4 we give a brief demonstration of the sensitivity of the $3\beta + 2\alpha$
34 lidar measurement system to the microphysical aerosol properties (the state parameters). Then in Sections 5 and 6 we
35 delve into specific metrics provided by the OE toolset: the degrees of freedom of the signal (Section 5) and the
36 propagated state errors (Section 6). In Section 7 we expand the discussion of the propagated state errors by discussing
37 the sensitivity to different levels of measurement uncertainty. Section 8 revisits the propagated state covariance matrix



1 with a new focus on the correlation terms. Section 9 discusses correlation in additional detail in terms of cross-talk
 2 between state parameters and gives a recommendation for resolving some of the ambiguity in $3\beta + 2\alpha$ retrievals.
 3 Section 10 provides a summary and outlook.

4 2. Methodology

5 With this study, we wish to develop a deeper understanding of the information content and sensitivities of the $3\beta +$
 6 2α measurement system in terms of aerosol microphysical parameters of interest, namely the complex refractive index,
 7 total number concentration, and a parameterization of the size distribution. The retrieval methodologies for this
 8 inversion system tend to be fairly complicated, particularly due to the difficulty in solving for the complex refractive
 9 index. For this study, we aim to determine the basic sensitivities common to all $3\beta + 2\alpha$ lidar retrievals using a
 10 methodology that is independent of any retrieval. We accomplish this by calculating the information content and
 11 uncertainty metrics using only a forward model, with no explicit inversion.

12 The measurements for these retrievals are bulk aerosol extinction and backscatter coefficients measured by an HSRL
 13 or Raman lidar system. They are related to the particle size distribution and complex refractive index of the volume
 14 of aerosols by this general relationship:

$$g_{i,\lambda} = \int_{r_{min}}^{r_{max}} K_i(m, r, \lambda) f(r) dr + \epsilon \quad (1)$$

15 where $g_{i,\lambda}$ represents a lidar measurement of either backscatter or extinction coefficient at wavelength λ . The function
 16 $f(r)$ represents the aerosol size distribution, which is a function of r , the particle size. K_i represents the extinction and
 17 backscatter measurement kernels, which are dependent on particle size, wavelength, and the complex refractive index,
 18 m . The measurements also include some measurement error, ϵ .

19 Eq. (1) is of the following general form:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \epsilon \quad (2)$$

20 in which \mathbf{F} , the forward function, relates the vector of state parameters, \mathbf{x} , to the vector of measurements \mathbf{y} . Comparing
 21 Eq. (2) with Eq. (1), the vector of measurements \mathbf{y} in Eq. (2) is comprised of $g_{i,\lambda}$, the five lidar measurements of
 22 backscatter and extinction. The state vector \mathbf{x} in Eq. (2) comprises the complex refractive index m and variables
 23 describing the size distribution $f(r)$.

24 If the forward model is linear or can be linearized, then Eq (2) can be written by the following matrix equation:

$$\mathbf{y} = \mathbf{J}\mathbf{x} + \epsilon \quad (3)$$

25 where \mathbf{J} , the Jacobian matrix, relates the state vector \mathbf{x} to the measurement vector \mathbf{y} .

26 Rodgers (2000) describes the generalized inverse problem, the optimal estimation methodology for solving it, and
 27 also a number of useful diagnostics for assessing the information content and retrieval errors. Although the existing
 28 lidar aerosol microphysical retrievals solve the generalized inverse problem in various ways not limited to optimal
 29 estimation, the metrics described by Rodgers (2000) are useful for the retrieval-free information assessment in this
 30 project. These include the scalar degrees of freedom (DOF) metric and the state error covariance matrix, propagated
 31 from the measurement errors. To calculate these metrics, it is necessary to have the weighting function matrix or



1 Jacobian matrix, \mathbf{J} , whose elements are the partial derivatives of the forward model elements with respect to the state
 2 vector elements.

$$J_{ij} = \frac{\partial F_i(\mathbf{x})}{\partial x_j} \quad (4)$$

3 To generate a Jacobian matrix for the purpose of the sensitivity study, we first simplify the problem by assuming
 4 single scattering processes from spherical particles, monomodal log-normal size distributions, and wavelength-
 5 independent refractive indices. The assumption of wavelength-independent refractive indices has been used in all 3β
 6 + 2α lidar aerosol microphysical inversions to date (Müller et al., 1999; Veselovskii et al., 2002; Bockmann et al.,
 7 2005; Chemyakin et al., 2014) and is necessitated in part by lack of knowledge of the wavelength dependence of the
 8 complex refractive index for real aerosols. However, some aerosol types may have significant spectral dependence
 9 of the complex refractive index which can impact microphysical retrievals (Veselovskii et al., 2016). The assumption
 10 of monomodal log-normal size distributions is used by the Arrange and Average algorithm (Chemyakin et al., 2014)
 11 but not the Inversion with Regularization algorithm (Müller et al., 1999), the Hybrid Regularization method
 12 (Bockmann et al., 2005), or the Linear Estimation method (Veselovskii et al., 2012). The retrievals which do not
 13 make this assumption can retrieve more general size distribution shapes of which the monomodal log-normal can be
 14 seen as a special case. Similarly, the assumption of spherical particles is generally found in these retrievals, due to
 15 limitations in the accuracy of non-spherical models for lidar measurements, but some retrieval studies (Veselovskii et
 16 al., 2010; Veselovskii et al., 2016) have allowed limited retrievals for non-spherical particles with more generalized
 17 assumptions about shape. Our forward model adopts the most restrictive assumptions used by any of these retrievals,
 18 and so has the fewest unknown state parameters. Therefore, it is useful for determining an upper limit on the
 19 measurement sensitivity applicable to all of these retrievals. That is, we are characterizing the retrieval of aerosols that
 20 conform perfectly to the most restrictive forward model assumptions. The same set of measurements would have less
 21 information content with respect to a forward model with more unknown state parameters. Additionally, mismatch
 22 between retrieval assumptions and true atmospheric aerosols will also generate errors which are not assessed by this
 23 analysis and which will be retrieval-dependent. Sensitivity studies to assess the measurement content with respect to
 24 more complex aerosol scenarios (specifically bimodal size distributions) are part of our ongoing work.
 25 Consistent with the assumption of spherical particles and single scattering processes, we use Mie kernels, which are
 26 calculated with code from Bohren and Huffman (1983). The size distributions are represented as monomodal log-
 27 normal size distributions characterized by the total number concentration, N ; median radius, r_{med} ; and geometric
 28 standard deviation, s . The mode width is the natural logarithm of s .

$$f(r) = \frac{N}{r\sqrt{2\pi}\ln s} \exp\left[-\frac{\left(\ln\frac{r}{r_{med}}\right)^2}{2(\ln s)^2}\right] \quad (5)$$

29 In all, five state parameters are used in this study: the median radius and geometric standard deviation of the
 30 monomodal log-normal size distribution, the total number concentration, and the complex refractive index (real and
 31 imaginary parts). From these, the extinction and 180° backscatter are calculated from Eq. (1) at the wavelengths



1 measured by the $3\beta + 2\alpha$ lidar system, which are 355 and 532 nm for extinction and 355, 532, and 1064 nm for
2 backscatter. The integrals are performed for values of r from 1 nm to 50 μm . The state parameters and the output
3 extinction and backscatter values are saved in the form of a LUT for a wide range of state variable values meant to
4 conservatively include realistic aerosol states. The original LUT was developed by Chemyakin et al. (2014), who
5 describe it in more detail. The version of the LUT used for this study includes median radii from 17 to 605 nm;
6 geometric standard deviations from 1.425 to 2.625; real refractive indices from 1.37 to 1.75; and imaginary refractive
7 indices including 0 plus increments from 0.00025 to 0.10175. For the purpose of this study, we have also included
8 total number concentration as a fifth dimension. The range of total number concentration values in the modified LUT
9 is 1-40,000 cm^{-3} .

10 For our purposes, the Jacobian matrix is calculated from the LUT using finite differences, using the increments of the
11 LUT itself. The use of finite differences amounts to an assumption that the increments are small enough that the
12 derivatives are locally linear. Testing with both smaller and larger increments confirms that the derivatives are
13 insensitive to the size of the increments from about one tenth the size of the increments used to at least about 5 times
14 the size used. However, the derivatives and associated retrieval sensitivities are not constant across the entire state
15 space. Therefore, the Jacobians and the metrics describing information content and error propagation have been
16 calculated for several specific realistic cases and also over multiple continuous slices of the hypercube defined by the
17 five state variables, to develop a sense of how these metrics vary over the state space.

18 Although the published aerosol microphysical retrievals referenced in the introduction solve the inverse problem in
19 various ways, the LUT can be thought of as a generalized realization of the forward function, given the simplifications
20 described above. Since the calculation of the sensitivity and error metrics (Rodgers, 2000) depends on the forward
21 function but not on any explicit retrieval, the LUT can be used to assess the $3\beta + 2\alpha$ measurement sensitivities with
22 respect to aerosol microphysical retrievals, independent of any particular retrieval strategy, not just the Arrange and
23 Average retrieval for which the LUT was developed.

24 Besides the Jacobian matrix, the sensitivity calculations also require the measurement covariance matrix, which
25 depends on the observation system. We use a simple but realistic matrix to describe the measurement errors for this
26 study, modeling the uncertainties as constant, normally distributed relative values with standard deviation of nominally
27 20% for the extinction coefficients and 5% for the backscatter coefficients, and with no correlations between the
28 uncertainties in each channel. Zero or near-zero correlation for the uncertainties between channels is realistic for lidar,
29 for which uncertainties are primarily from random processes (e.g. shot noise) and channel-specific systematic sources
30 (e.g. filter transmittance). The uncertainty levels used in this study are chosen as realistic targets for a space-based
31 lidar system, based on existing HSRL-2 technology (Hair et al., 2008; Burton et al., 2015). Later in this study (Section
32 7), we explore a few other benchmark values of measurement uncertainties. In reality, uncertainty will not be constant
33 for all aerosol scenarios, but for the purpose of this study, a few benchmark values are sufficient to explore the
34 sensitivities.

35 The third input needed for these calculations is the a priori covariance matrix. This matrix represents the uncertainty
36 of the prior knowledge of the state. The diagonal terms represent the variance, and are chosen so that the standard
37 deviation is represented as one half of the full range in the look-up table for each state variable. The off-diagonal terms



1 represent the correlation or covariance between state variables; we assume zero correlation in the a priori. These large
 2 prior variances and zero correlations are an intentionally conservative choice. For an actual retrieval, prior information
 3 about aerosol type and real aerosol variability would typically be used to decrease these prior variance terms, which
 4 can certainly decrease the uncertainty in the final result. Likewise, if it were known a priori that the state variables
 5 were correlated, this could also be used to decrease the uncertainty in the final result. However, since our aim is
 6 primarily to assess the information content of the measurements themselves, we use conservative prior variance and
 7 covariance values for the sensitivity study. We recognize that the state variables are not normally distributed in reality,
 8 although the OE formalism makes the assumption that they are (and that the measurement errors likewise are normally
 9 distributed). A more advanced strategy would be to use the Markov Chain Monte Carlo method (Posselt and Mace,
 10 2014) which allows for generalized error distributions. However, for this initial study, we use the more straightforward
 11 OE method and partially compensate by choosing conservatively large prior variances values.

12 3. Case Definitions

13 In describing the calculation of the metrics, we will illustrate the procedures and interpretation using five particular
 14 sets of values in the state space, which we collectively call “the reference cases.” The values of the state variables for
 15 the five reference cases are given in Table 1, as well as values of effective radius, effective variance, and single
 16 scattering albedo (SSA) which are calculated from the state variables. For a log-normal distribution, the effective
 17 radius and effective variance can be expressed as analytical functions of r_{med} and s :

$$18 \quad r_{eff} = r_{med} \exp\left(\frac{5 \ln^2 s}{2}\right) \quad (6)$$

$$19 \quad v_{eff} = \exp(\ln^2 s) - 1 \quad (7)$$

20 The SSA is calculated from the state variables using Mie theory.

21 The first of the reference cases has been constructed to approximately reflect a real measurement scenario
 22 encountered during the DOE TCAP (Two-Column Aerosol Project) field mission by HSRL-2 (Berg et al., 2015;
 23 Müller et al., 2014); the parameters model a plume of urban outflow. The complex refractive index for this constructed
 24 case is $1.47 - 0.00325i$, corresponding to a very weakly absorbing aerosol with SSA value of 0.98 at 532 nm. The
 25 aerosol is composed of accumulation mode particles; the constructed monomodal size distribution has effective radius
 26 of $0.17 \mu\text{m}$ and effective variance of 0.16, or median radius $0.12 \mu\text{m}$ and geometric standard deviation of 1.48. The
 total number concentration is moderate, with a value of 1100 cm^{-3} .

27 For the other four reference cases, we vary the state variables in sets. Cases 2 and 3 have the same complex refractive
 28 index as Case 1, but different size distributions. For Case 2, the effective radius and effective variance are somewhat
 29 larger at $0.24 \mu\text{m}$ and 0.23, respectively (median radius = $0.15 \mu\text{m}$ and geometric standard deviation = 1.58). Like
 30 Case 1, this size distribution is considered fine mode. For Case 3, the effective radius and effective variance are 1.60
 31 μm and 1.27, respectively (median radius is $0.20 \mu\text{m}$ and geometric standard deviation is 2.48). The total number
 32 concentration is 50 cm^{-3} . The total number concentration is much lower than Cases 1 and 2, but the larger particles are
 33 more scattering and therefore the signal levels are comparable. For comparison, the 532 nm extinction value is 0.092



1 km⁻¹ for Case 1 and 0.084 km⁻¹ for Case 3. This larger size distribution approximately reflects a coarse-mode marine
2 aerosol, although the complex refractive index is not necessarily appropriate for marine aerosol. Since we will be
3 using these cases to understand the dependencies of the retrieval sensitivities on the state space, we choose to vary the
4 state variables relating to the size distribution separately from those relating to the complex refractive index.
5 Case 4 has a size distribution equal to Case 1, but larger real and imaginary refractive index values of 1.61 and 0.03,
6 respectively. For this size distribution, this complex refractive index corresponds to a 532 nm SSA value of 0.89.
7 This can be thought of as similar to a biomass burning plume.
8 Case 5 is similar to Case 4 in everything except total number concentration. Now the total number concentration has
9 been increased dramatically to 20,000 cm⁻³, approximating as a very intense smoke plume.
10 These 5 cases will be used for illustrating the results of the sensitivity analysis, starting in Section 5.

11 4. Dependence of Lidar Intensive Variables on Aerosol Microphysical Parameters

12 First, to build an intuition of the information content encoded within the $3\beta + 2\alpha$ dataset, we briefly examine the
13 dependence of some of the lidar intensive variables on the effective radius, Eq. (6), and complex refractive index. We
14 use Mie modeling of spherical particles and use the simplified assumption of a monomodal log-normal size
15 distribution (as discussed in Section 2) for this exercise.

16 Recall that aerosol intensive variables are those that do not scale with the amount of aerosol loading. Of the five state
17 variables, total number concentration is an extensive variable while the other four (real and imaginary parts of the
18 refractive index, median radius, and geometric standard deviation) are intensive variables. Aerosol extinction and
19 backscatter coefficients, the direct measurements of a lidar using the HSRL or Raman techniques, are extensive
20 variables; ratios of these basic measurements are intensive variables. Burton et al. (2012) show that intensive variables
21 such as the lidar ratio (extinction-to-backscatter ratio at a given wavelength) and backscatter color ratio (i.e. ratio of
22 backscatter at two different wavelengths) encode information about the type of aerosol present in broad categories,
23 i.e. marine vs. smoke vs. urban pollution. It is also known that the extinction Ångström exponent is sensitive to the
24 particle size distribution (e.g. Schuster et al., 2006; Ångström, 1929; Kaufman et al., 1994).

25 Figure 1 (left panel) illustrates the monotonic dependence of extinction Ångström exponent (355/532 nm) on the
26 effective radius, for monomodal log-normal size distributions. The sensitivity of this parameter to either the real or
27 imaginary part of the refractive index is smaller, as demonstrated by shallower slopes in the middle and right panels
28 of Figure 1. On the other hand, Figure 2 illustrates the dependencies of the lidar ratio (at 532 nm and 355 nm), which
29 is the ratio of the extinction to backscatter and is also the inverse of the product of the aerosol 180° phase function
30 and the single scattering albedo. The dependence on effective radius is non-monotonic, and there is a complicated
31 dependence on the real refractive index. For effective radius larger than about 0.1 μm, there is significant sensitivity
32 to the real refractive index. There is a monotonic relationship with the imaginary refractive index, with greater
33 sensitivity (steeper slopes) for the 355 nm lidar ratio compared to the 532 nm lidar ratio, reflecting the relationship
34 between lidar ratio and absorption. The lidar ratio increases as the imaginary part of the refractive index increases and
35 (for large enough sizes) as the real part of the refractive index decreases. In Figure 3 the backscatter color ratio



1 (532/1064 nm) is shown to vary in a complicated way with the real and imaginary refractive indices and with the
 2 effective radius, with differently shaped curves compared to the dependence of the extinction Ångström exponent and
 3 lidar ratios. Total number concentration is not reflected in any of the intensive parameters of course, but by definition
 4 the extensive parameters (backscatter and extinction) are linearly related to N .
 5 While this simple sensitivity check illustrates that changes in the aerosol microphysical parameters are reflected in the
 6 measurements, it is not sufficient to determine if the measurements are enough to retrieve all five state parameters.
 7 For that, we must turn to more quantitative tools.

8 5. Degrees of freedom and averaging kernel matrix

9 The retrieval problem as specified above consists of five direct aerosol measurements (two extinction and three
 10 backscatter measurements) from a lidar system at a single level in the atmosphere and five state vector elements (three
 11 describing the number and size distribution and two to specify the complex refractive index). We would like to know
 12 if the five measurements are sufficient to determine the five unknowns, in other words, to determine if the inverse
 13 system is fully determined, overdetermined, or underdetermined and by how much. Rodgers (2000) describes a useful
 14 metric to quantify the number of pieces of independent information in the measurement, the degrees of freedom for
 15 the signal, d_s . It is defined as the trace of the matrix $\tilde{\mathbf{J}}$, which is known as either the pre-whitening matrix (Rodgers,
 16 2000) or the error-normalized Jacobian matrix (Xu and Wang, 2015). This matrix is defined in terms of the Jacobian
 17 matrix, \mathbf{J} , the measurement error covariance matrix, \mathbf{S}_s , and the a priori covariance matrix, \mathbf{S}_a .

$$\tilde{\mathbf{J}} = \mathbf{S}_s^{-\frac{1}{2}} \mathbf{J} \mathbf{S}_a^{\frac{1}{2}} \quad (8)$$

18 Since the error-normalized Jacobian matrix is weighted by the prior covariance in the numerator and the measurement
 19 error in the denominator, elements greater than unity indicate where variability in the true state exceeds measurement
 20 noise. The trace of the matrix, d_s , therefore indicates the number of independent pieces of information about the state
 21 contained in the measurements. For a fully determined retrieval system, the degrees of freedom would be equal to the
 22 number of state parameters.

23 For the first reference case, Case 1 (see Table 1 for description), the signal DOF, d_s , is determined using this method
 24 to be 4.5. The implication of this calculation is that of the five pieces of information required to specify the state, 4.5
 25 of them are provided by the measurement signal.

26 The quoted d_s is only applicable to one particular value of the state vector. In general, the information content is
 27 regime-dependent (dependent on the state). For the other cases in Table 1, d_s is 4.6 for the slightly larger fine mode
 28 case, 3.9 for the coarse mode case, 4.2 for the absorbing case, and 3.8 for the case with large total number
 29 concentration. Figure 4 provides a more detailed look at the regime dependence for two orthogonal “slices” through
 30 the 5-dimensional state space, illustrating that values of approximately 4 are typical of most of the space, except for
 31 the smallest particle radii, where the signal DOF decreases closer to 3.

32 Signal values for the degrees of freedom less than 5 mean that some of the information in the five retrieved parameters
 33 is not provided directly by the measurements and will be “filled in” by a priori information or other constraints in a
 34 retrieval. A value of d_s less than five is not surprising, because it is already well understood that this problem is



1 underdetermined (Veselovskii et al., 2002; Bockmann et al., 2005; Pérez-Ramírez et al., 2013). In general for this
 2 system, we find that approximately 4 independent pieces of information are provided by the measurements, with slight
 3 regime dependence.

4 **6. Propagated state uncertainties**

5 While the signal DOF is useful for providing an overall estimate of the under-determinedness of the system as a single
 6 number, the a posteriori (i.e. propagated) state error covariance matrix provides more detail about the sensitivity of
 7 the system to each state variable. The state error covariance matrix, $\hat{\mathbf{S}}$, is propagated from the measurement error
 8 covariance matrix \mathbf{S}_ε and a priori covariance matrix \mathbf{S}_a using the Jacobian matrix, \mathbf{J} , by

$$\hat{\mathbf{S}} = (\mathbf{J}^T \mathbf{S}_\varepsilon^{-1} \mathbf{J} + \mathbf{S}_a^{-1})^{-1} \quad (9)$$

9 Table 2 shows $\hat{\mathbf{S}}$ for the first reference case as an example. The diagonal elements in the covariance matrix are the
 10 variance terms, and their square roots are the standard deviations. These standard deviations, which we will also call
 11 the propagated uncertainties, are shown in Table 3 for the five reference cases. Table 3 also shows the prior uncertainty
 12 from the a priori covariance matrix. Comparing the propagated uncertainty with the prior uncertainty shows how the
 13 measurements constrain the retrieval beyond the prior knowledge. For the size distribution parameters, the assigned
 14 prior standard deviations are 0.3 for the median radius, 0.6 for the geometric standard deviation and 20,000 for the
 15 total number concentration. In each of the reference cases, the propagated uncertainty values from Table 3 for these
 16 three variables represent a significant reduction in the standard deviation by 40-87% for the median radius, 17-84%
 17 for the geometric standard deviation, and 31-99% for the total number concentration. The measurements also reduce
 18 the prior standard deviation of the RRI significantly, by a factor of 26-79% from the prior standard deviation of 0.19.
 19 For IRI, there is a reduction of 52-90% from the prior standard deviation of 0.05. So, the measurements constrain
 20 knowledge of all of the state variables beyond the prior knowledge.

21 Since the prior covariance matrix was defined rather conservatively in this study, the reduction from the prior
 22 uncertainty may be less useful than comparing to uncertainty values defined in terms of a desired goal. Part of the
 23 motivation of this study is to determine the extent to which a $3\beta + 2\alpha$ lidar system can meet the requirements outlined
 24 in the ACE satellite white paper (http://acemission.gsfc.nasa.gov/Study_Report_2010.html, accessed 22 Oct 2015).
 25 These draft ACE requirements, shown in Table 3, in some cases specify retrieval precisions defined with respect to a
 26 vertically resolved profile with resolution of 1.5 km in the free troposphere and 500 m in the boundary layer. These
 27 include the total number concentration with a retrieval precision (one standard deviation) to within 100% (relative)
 28 and SSA to within 0.02 (absolute). Other ACE draft requirements are specified for column-equivalent values. These
 29 include RRI to within 0.02 (absolute), effective radius to within 10% (relative) and effective variance to within 50%
 30 (relative). The ACE satellite is planned to include both a multi-wavelength lidar and multi-wavelength, multi-angle
 31 polarimeter. The requirements reflect the expectation that both instruments will be used in a combined retrieval, but
 32 this measurement configuration is out of the scope of the current sensitivity study.

33 Some of the ACE requirements, are stated in terms of the effective radius, effective variance and SSA, quantities that
 34 are not part of the nominal set of state variables described above; however, they are directly related to the state



1 variables and can be derived from them. In general, if a secondary variable, z , can be expressed as some function of
 2 the state variables, \mathbf{x} ,

$$z = F(\mathbf{x}) \quad (10)$$

3 then the random error of the secondary variable can be calculated using the state error covariance matrix $\hat{\mathbf{S}}$ and the
 4 partial derivatives of the secondary variable with respect to the state variables.

$$\sigma_z = \sqrt{\sum_i \sum_j \hat{S}_{ij} \frac{\partial z}{\partial x_i} \frac{\partial z}{\partial x_j}} \quad (11)$$

5 For our purpose, the variable z can represent either the effective radius, effective variance, or SSA.

6 The effective radius and effective variance can be calculated for a monomodal log-normal size distribution using Eqs.
 7 (6) and (7). The functional dependence of single scattering albedo, which is the ratio of the scattering efficiency to the
 8 total extinction efficiency, can be obtained using Mie theory. Then the propagated uncertainties for these quantities
 9 can be obtained using Eq. (11) with partial derivatives that are calculated either analytically or using finite differencing
 10 on the output of the Mie code from Bohren and Huffman (1983). The propagated uncertainties for the effective radius,
 11 effective variance, and SSA are also shown in Table 3.

12 As with the signal degrees of freedom, the propagated errors for the state vector elements and for effective radius,
 13 effective variance and SSA are regime dependent, varying over different parts of the state space. In the Supplemental
 14 Material are figures like Figure 4 that show two orthogonal slices through the 5-dimensional state space for the five
 15 state variables and also the derived variables effective radius, effective variance, and SSA. These illustrate the ease
 16 with which the sensitivity metrics can be calculated for the whole state space, but since some of the states represented
 17 in these slices may not be particularly realistic, it can be hard to interpret the results. Therefore, the five reference
 18 cases in Table 1 were designed to provide a focus for understanding the regime dependence more easily.

19 Recall that the differences between Case 1, 2 and 3 are related to the size distribution. The size distribution for Case
 20 3 is a coarse mode with a larger particle size, larger geometric standard deviation and smaller total number
 21 concentration than Cases 1 and 2. Compared to Case 1 or 2, Case 3 has larger propagated relative uncertainty of the
 22 effective radius, 50% uncertainty compared to 23% and 29%, and also for total number concentration, 122% compared
 23 to 98 and 103%, mostly due to the increase in the geometric standard deviation. For the most part, we found increasing
 24 relative uncertainties for the size distribution parameters for increasing geometric standard deviations (with some
 25 exceptions, as can be seen in the Supplemental figures). However, compared to Case 1, Case 3 has smaller
 26 uncertainties on the complex refractive index and SSA, although the complex refractive index did not change between
 27 cases.

28 Case 4 has the same size distribution as Case 1, but the complex refractive index corresponds to a more absorbing
 29 aerosol. There are only minor differences in the size distribution uncertainties, but the uncertainties on the complex
 30 refractive index and SSA increase, suggesting less sensitivity in the retrieval to the complex refractive index of
 31 absorbing aerosols.

32 Case 5 is identical to Case 4 except that it has a very large total number concentration. Although such a large total
 33 number concentration in a real world measurement scenario would mean greatly increased signal-to-noise ratio (SNR),
 34 recall that the measurement errors for this study are defined as constant percentages, so the SNR effect is not reflected



1 in this study. Instead, total number concentration behaves essentially as a scaling variable in the retrieval, and
2 therefore most of the propagated uncertainties are very similar for Case 5 compared to Case 4. The exception is the
3 uncertainty in the total number concentration itself, which decreases from 94% to 68%.

4 Comparing the propagated uncertainties to the ACE requirements, note that ACE calls for an uncertainty on the
5 column total number concentration of 100%. The uncertainties in Table 3 show that the $3\beta + 2\alpha$ retrieval already
6 meets this requirement even for vertically resolved measurement levels in the absorbing aerosol cases and meets it or
7 is very close to meeting it in the non-absorbing fine mode cases. The coarse mode case has the largest total number
8 concentration uncertainty, 122%, but is still reasonably close to the column uncertainty target. Note that a requirement
9 on the column uncertainty is less restrictive than a requirement on vertically resolved measurement levels. This study
10 focuses only on the sensitivities of single-level retrievals, but full profile retrievals are also possible (Kolgotin and
11 Müller, 2008) which optimize the use of simultaneous information content from multiple related vertical levels. If an
12 aerosol layer extends across multiple measurement levels, then a profile retrieval which combines measurements from
13 multiple levels within the column would include proportionately more measurement information content (since the
14 noise in the measurements is mostly uncorrelated and the aerosol properties are correlated), and so the uncertainty
15 would be reduced, compared to a single-level retrieval. In the future, we will perform sensitivity studies for such a
16 retrieval system.

17 The uncertainties on the vertically-resolved effective variance are 36% and 41% for the absorbing cases, which already
18 meets the proposed ACE column requirement of 50%. The non-absorbing fine mode and coarse mode cases have
19 effective variance uncertainties of 61% to 68%, not very much larger than the ACE column requirement.

20 The requirement of 10% column uncertainty for the effective radius is not met for any of the five illustrated cases on
21 a vertically resolved basis; the propagated uncertainties are two to three times larger for the three fine mode cases, and
22 five times larger for the coarse mode case. A factor of two or three may be recoverable by a profile retrieval which
23 uses multiple vertically resolved measurement levels simultaneously, assuming the aerosol properties are correlated
24 across several levels.

25 The propagated uncertainty on the real refractive index is two to seven times the proposed ACE column requirement,
26 in this case smallest for the coarse mode case and worst for the two absorbing fine mode cases.

27 The proposed ACE requirement for SSA is 0.02 on a vertically resolved basis. The propagated uncertainties for all
28 four cases are 3 to 5 times larger than this proposed requirement, which may be sufficient for distinguishing extreme
29 cases such as intense biomass burning plumes, and also may be reducible to some extent by a profile retrieval.

30 **7. Performance assessment for varying measurement errors**

31 It should perhaps be mentioned that the ACE requirements are not necessarily finalized and the values quoted here
32 are draft requirements. Similarly, the instrument performance used for the results described above is only approximate
33 based on a best-guess estimate of realistic targets for a space-based lidar system, based on the technology used for the
34 airborne HSRL-2. Since the motivation for this study is to determine what retrieval performance is possible from a
35 lidar-only microphysical retrieval, it is useful to briefly explore the retrieval uncertainties as a function of instrument
36 performance. Table 4 accordingly shows the propagated uncertainties, using reference case #1, for three different



1 instrument configurations with different measurement uncertainties for backscatter and extinction. The first
2 measurement configuration assumes that the uncertainties are larger than previously described, 10% for aerosol
3 backscatter and 30% for aerosol extinction. The second of the three configurations in Table 4 is a repetition from
4 Table 3, with uncertainties of 5% and 20% on aerosol backscatter and extinction, respectively. The third theoretical
5 instrument configuration is more ambitious, with assumed uncertainty 5% on aerosol backscatter and 10% on aerosol
6 extinction. Comparing the first and second scenarios, if the measurement uncertainties are allowed to increase as
7 described, the retrieval uncertainties increase by a factor of 20-50%. Comparing the second and third scenarios, if
8 instead the extinction measurement uncertainty is decreased by half, then the retrieval uncertainties all decrease by
9 approximately 30-40%. In the third scenario, the draft ACE requirement for vertically resolved total number
10 concentration is met; the requirement for column effective variance is met even on a vertically resolved basis; and the
11 vertically resolved effective radius uncertainty is less than twice the column requirement. However, the real refractive
12 index and SSA uncertainties are still large compared to the ACE draft requirements. This level of precision and
13 accuracy may be difficult to achieve with a satellite lidar.

14 Recall that the proposed ACE system consists of both a multi-wavelength High Spectral Resolution Lidar and a multi-
15 wavelength multi-angle polarimeter. The current sensitivity study addresses only the lidar. A combined retrieval with
16 both lidar and polarimeter will certainly have higher information content particularly pertaining to aerosol absorption,
17 and a better chance of meeting all of the draft ACE requirements. To quantitatively assess the information content of
18 this more complicated system, a full column retrieval using a combined lidar-plus-polarimeter forward model would
19 be required, which is out of the scope of the current paper.

20 Based on the current study, it seems likely that a $3\beta + 2\alpha$ lidar-only system with measurement errors similar to those
21 studied here will have trouble retrieving single scattering albedo to the target level of uncertainty, and that additional
22 information content must be provided, such as from coincident passive (sun-photometer or polarimeter) measurements
23 at more wavelengths, or if additional measurements are not available, then from a priori constraints.

24 **8. Correlation matrix**

25 Besides the diagonal variance elements, the state error covariance matrix includes off-diagonal terms that describe the
26 interaction between pairs of state variables in the retrieval. Prior similar sensitivity studies for other systems do not
27 explicitly address the off-diagonal terms of the propagated matrix (Xu and Wang, 2015; Knobelspiesse et al., 2012),
28 but these terms give critical information about retrieval performance. To illustrate, Table 5 and Table 6 give the state
29 error correlation matrix for Case 1 and Case 2, respectively. These can be easily converted from the state error
30 covariance matrices, like the one given in Table 2 for Case 1. The correlation matrices show that there is some
31 correlation between all pairs of variables, with the highest correlations between the real and imaginary parts of the
32 complex refractive index and between the total number concentration and median radius. The correlations have a
33 complicated regime dependence, illustrated in Figure 5, S9 and S10. Although Cases 1 and 2 vary only a small amount
34 in median radius and geometric standard deviation, there is a significant increase for Case 2 in the magnitude of the
35 correlations between size distribution variables. For Case 2, the correlation is -0.97 between the median radius and
36 total number concentration and also between the median radius and geometric standard deviation.



1 High magnitude correlations between the retrieved variables indicate the potential for cross-talk between these
2 parameters. Cross-talk can cause additional error in the retrieved parameters that is not reflected in the variance terms,
3 due to non-unique solutions which have compensating errors. In an ideal case with no cross-talk, the forward model
4 evaluated at the true state would produce output equal to the measurements (ignoring measurement error), while the
5 forward model evaluated using an incorrect state vector should produce output that does not agree with measurements.
6 However, in the case of compensating errors or cross-talk, an incorrect solution may also reproduce the measurements
7 if, for example, an error in the median radius that tends to produce larger backscatter and extinction values is
8 compensated by an error in the total number concentration that tends to produce smaller values. Such compensating
9 errors make it impossible for the measurements to distinguish between the true state and the erroneous state.
10 The cause of the cross-talk can broadly be described as a lack of sensitivity in the measurements. The cross-talk
11 between total number concentration and median radius occurs because particles significantly smaller than the shortest
12 measured wavelength (355 nm) contribute little to observed optical properties. Therefore, the measurements can be
13 insensitive to the difference between large numbers of very small particles and smaller numbers of larger particles.
14 This problem, and a partial remediation, are examined in more detail in Section 9. The cross-talk between the real
15 and imaginary index of refraction is related to a relative lack of sensitivity to absorption in the lidar measurements.
16 Probably the best remedy for this latter problem is to incorporate additional information content into the retrieval,
17 preferably in the form of additional coincident measurements, as from a polarimeter on the same platform.

18 **9. Cross-talk between size parameter and total number concentration**

19 Taking measurement error into account, there are always multiple solutions that reproduce the measurements to within
20 the measurement error. This is not a concern if the solutions are clustered around the true solution, but can be a
21 significant issue in the case of cross-talk or compensating errors as discussed above. Figure 6 and Figure 7 show
22 histograms of the number of solution states in the gridded LUT that reproduce the backscatter and extinction values
23 of Case 1 to within the prescribed error bars (5% for backscatter coefficient and 20% for extinction coefficient). Figure
24 6 shows the total number concentration and Figure 7 shows the median radius, respectively. Note that although the
25 peaks of the histograms do not exactly match the specified values for Case 1 (indicated by dashed lines), the solutions
26 are clustered around those values.

27 In contrast, Figure 8 and Figure 9 illustrate the set of solutions from the gridded LUT that match Case 2 within the
28 measurement error bars (shown in red). Figure 8 shows that this set of solutions covers an enormous range in total
29 number concentration. The range of total number concentration for these solutions is much larger than indicated by
30 the propagated standard deviation shown in Table 6. Cases 1 and 2 have similar propagated standard deviations; the
31 problem with Case 2 is only evident in the near-unity correlation value between total number concentration and median
32 radius, shown in Table 6. The very high correlation indicates that the solutions with very large total number
33 concentrations are those solutions that also have very small median radii. Very small particles contribute little to
34 extinction or backscatter at the lidar wavelengths. So, large numbers of very small particles can be included in the
35 retrieved solution without significantly affecting the agreement with the measurements; therefore, the measurements
36 by themselves are not sufficient to determine if these very small particles are actually present.



1 This situation emphasizes the value of examining the cross-terms of the propagated error matrix. The regime
2 dependence of this situation is complex and the problem can be detected only by studying the correlation matrix or
3 else by examining the distribution of solutions for a given retrieval.

4 A resolution of the cross-talk can be achieved by adding an additional constraint on either the total number
5 concentration or the radius using a priori information. For example, it is probably unrealistic to allow total number
6 concentrations up to $40,000 \text{ cm}^{-3}$. However, it is not clear how one would determine a realistic upper bound on the
7 total number concentration. We argue that a better solution is to constrain the radius of the particles. After all, the
8 limitation of the lidar measurements is a lack of sensitivity to particles much smaller than the smallest lidar
9 wavelength; it is not a limitation on the sensitivity to large total number concentrations. To approximate such a
10 constraint, we repeated the retrieval for Cases 1 and 2 using a limited version of the LUT, where some solutions are
11 disallowed depending on the size of particles in the size distribution. The blue histograms in Figure 6 – Figure 9
12 illustrate the solutions for which 80% or more of the particles in the size distribution are larger than 50 nm radius. As
13 seen in the blue histograms, limiting the retrieval to larger particles improves the cross-talk problem for Case 2, and
14 the solutions are now much better constrained around the truth solution.

15 Fifty nm radius is proposed for the cutoff value based on the sensitivity of the lidar measurements and on the naturally
16 occurring lower bound of the atmospheric aerosol accumulation mode. Typically aerosol size distributions are
17 described in terms of three or four size modes, depending on if one is examining the number- or mass-based size
18 distribution (Seinfeld and Pandis, 2006). Most particles, on a number basis, exist in the ultrafine diameter size range
19 of a few nanometers up to a few hundred nanometers, with two distinct modes: the nucleation mode ($D < 10 \text{ nm}$) and
20 the Aitken mode ($10 \text{ nm} < D < 300 \text{ nm}$). Nucleation mode particles are fresh aerosols created via gas-to-particle
21 nucleation, while the Aitken mode encompasses directly emitted particles and particles that have grown via
22 coagulation or gas-to-particle condensation. Meanwhile, on a mass basis, the aerosol size distribution is dominated by
23 two larger modes (with the ultrafine particles contributing almost negligible mass): the accumulation mode (100 nm
24 $< D < 2.5 \text{ }\mu\text{m}$; consisting of direct particle emissions, coagulation of smaller particles, and gas-to-particle condensation
25 of sulfates, nitrates, and organics) and the coarse mode ($2.5 \text{ }\mu\text{m} < D < 50 \text{ }\mu\text{m}$; consisting of particles formed via
26 mechanical processes such as wind-blown dust or sea salt). Atmospheric photo-oxidation and cloud processing can
27 also affect these modes and cause both the number and mass size distributions to shift toward larger sizes, as is often
28 seen for cloud processed marine aerosol (Hoppel et al., 1986).

29 Clearly, if an Aitken or nucleation mode with large number concentration does exist, limiting the size range of the
30 retrieval introduces the possibility of bias in total number concentration. Yet it is important to realize that even if it is
31 known from external sources (such as in situ measurements) that an observation is occurring in a region of significant
32 new particle production, lowering the cutoff radius will not resolve the systematic error in the retrieval, since the
33 measurements cannot distinguish between large numbers of very small particles and smaller numbers of larger
34 particles. Therefore, we think it is sensible to limit the particle size in the retrievals to reflect the measurement
35 sensitivity to larger sized particles. This strategy also has the benefits of making the constraint explicit and leading to
36 a clear and understandable interpretation of the results. In this case, the retrieval should not be described as a retrieval



1 of total aerosol number concentration, but rather as a retrieval of accumulation mode and coarse mode aerosols, more
2 accurately reflecting the retrieval sensitivities.

3 This strategy has heritage in existing retrievals. In Inversion with Regularization (Müller et al., 1999), the under-
4 determination of the retrieval is addressed by putting strong constraints on the window of particle sizes that are
5 considered, effectively limiting the minimum particle radius to 50 nm (Veselovskii et al., 2002). However, in that
6 retrieval the limit varies from case to case and even from one solution to another within the set of solutions that are
7 averaged for a particular retrieval, with the minimum radius being anything between 50 nm and 500 nm. Since we
8 argue that the need for a minimum particle radius cutoff is related to the limited sensitivity of the measurements to
9 very small particles, we believe that a single cutoff would be more consistent with our understanding of the retrieval
10 sensitivities. In any case, it is important to recognize that the size cutoff amounts to prior information supplementing
11 the information content of the measurements; explicitly describing the prior information is essential to understanding
12 and evaluating retrieval systems and their products.

13 To investigate the potential for bias associated with the particle size cutoff, it is useful to examine how much of an
14 effect the Aitken mode would have on the backscatter and extinction measurements. For this exercise, we start with
15 a retrieval case similar to an actual measurement from the NASA Langley HSRL-2 on 17 July 2012, from the Two-
16 Column Aerosol Project, described by Müller et al. (2014), and then add on a simulated mode with particle radius of
17 15 nm (diameter = 30 nm) and varying number concentration. For the purpose of this exercise, we limit the simulated
18 Aitken mode to a narrow mode width ($s = 1.48$). For the complex refractive index of the simulated Aitken mode, we
19 used values given by Costabile et al. (2013). Figure 10 shows the result of this numerical experiment, and demonstrates
20 that even for $40,000 \text{ cm}^{-3}$ of simulated Aitken particles, the maximum effect on the measurements is less than 2% for
21 the 355 nm backscatter measurement, and less for other wavelengths and for the extinction (i.e. compared to the actual
22 backscatter and extinction measurements for the TCAP case). Two percent is not a significant effect on the
23 measurements. Given that it is significantly smaller than the assumed measurement errors for this sensitivity study, it
24 is fair to say that the measurements are not sensitive to this mode. (For airborne HSRL-2 measurements with 5-minute
25 averaging such as were used by Müller et al. (2014), the effect on backscatter is about the same size as the random
26 errors and significantly smaller than the extinction random errors.) For number concentrations larger than $40,000$
27 cm^{-3} , the effect of the simulated Aitken mode is of course larger, due to the linear dependence of backscatter and
28 extinction coefficients on total number concentration. For these particles, it would require a number concentration of
29 approximately 10^6 cm^{-3} to have a significant impact on the measurements. Examples of measured Aitken and
30 nucleation mode number concentrations include values of about $13,000 \text{ cm}^{-3}$ for each of the two modes from a case
31 of new particle formation in an urban environment described by Cheung et al. (2013) and a maximum of $50,000 \text{ cm}^{-3}$
32 of particles with radii less than 5 nm for a case of new particle formation in an agricultural region (Mozurkewich et
33 al., 2004). For this latter case, since the particles are much smaller, the effect on the backscatter and extinction is
34 smaller than the $40,000 \text{ cm}^{-3}$ of 15 nm radius particles simulated above, so it seems reasonable to suggest that number
35 concentrations of particles in this size range would rarely be large enough to significantly affect the lidar
36 measurements.



1 As the particle radius gets larger, the sensitivity of the measurements to these aerosols increases. Figure 11 shows the
2 effect as a fraction of backscatter and extinction (again using the measurements from the TCAP case on 17 July 2012
3 as a reference) of 1000 cm^{-3} of particles of varying median radius. At about 50 nm median radius, the approximate
4 boundary between the Aitken and accumulation modes, the effect is a few percent to 10% of the backscatter and
5 extinction, which is on the order of the measurement uncertainty. For larger particles in the accumulation mode, the
6 effect is a significant portion of the measurements, reflecting that the measurements have good sensitivity to the
7 accumulation mode. This suggests that a 50 nm radius is a reasonable cutoff to use in retrievals, representing the
8 approximate boundary where the measurements have reasonable sensitivity. Of course, the true sensitivity of the
9 measurements depends on the number concentration, but since N is unknown, a constant cutoff is a good strategy.

10 It is worth pointing out that although it is true that lidar measurements lack sensitivity to particles much smaller than
11 the smallest wavelength, they do not lack sensitivity to particles much larger than the longest wavelength, as is
12 sometimes stated. For instance, it is not true that “pollens cannot be observed with lidar systems” (Bockmann et al.,
13 2005). See Figure 12 for an illustration of lidar measurements simulated by Mie modeling for very large particles. At
14 these large particle sizes, a forward model for the lidar based only on the single scattering Mie calculations is no longer
15 applicable, but this simple illustration serves to show that the backscatter and extinction coefficients are much larger,
16 not smaller, than the benchmark observations of the lidar. The scattering efficiency of large particles is significant
17 even at wavelengths much smaller than the particle size and so the effect of laser light scattering from large particles
18 is easily seen using lidar. However, since the particle size dependence of the lidar measurements is not monotonic at
19 large particle sizes and the single scattering forward model is no longer applicable, microphysical retrievals of particle
20 properties are challenged at large particle sizes. See Gasteiger and Freudenthaler (2014) for a further discussion of
21 retrieval of large particle size from multi-wavelength lidar.

22 **10. Summary and Discussion**

23 There is considerable interest in retrievals of aerosol size distribution parameters and absorption properties using
24 multi-wavelength High Spectral Resolution Lidar or Raman lidar. While there have been successful $3\beta + 2\alpha$ retrievals
25 of some particle properties (Müller et al., 2014; Veselovskii et al., 2016), there is also well-justified concern that these
26 retrievals are somewhat underdetermined. In this study we have taken a rigorous look at the information content of
27 single-height-level $3\beta + 2\alpha$ lidar measurements with respect to the microphysical parameters of interest, using
28 implementation-independent tools from the field of optimal estimation, which allows for combining measurements,
29 measurement errors, and constraints within a single coherent framework. By avoiding a retrieval and using the forward
30 model only (along with reasonable measurement uncertainties and a conservative a priori covariance matrix) we isolate
31 the sensitivities of the measurements themselves for a best case aerosol scenario, a monomodal log-normal distribution
32 of spherical particles with spectrally independent complex refractive index. Retrieval-dependent uncertainties related
33 to retrieval methodology or mismatch between the assumptions and the real-world aerosols are not included. On the
34 other hand, actual retrievals generally benefit from using various constraints and a priori information. A priori
35 knowledge is intentionally minimized in this study to focus on the measurement sensitivities, but in general it will
36 improve retrieval performance from this basic level.



1 We find that the five $3\beta + 2\alpha$ lidar measurements provide approximately four independent pieces of information to
2 describe the aerosol microphysical state space, with only slight regime dependence. Using reasonable lidar
3 measurement uncertainties, the retrieval uncertainties are closest to the proposed ACE satellite precision requirements
4 for the size distribution parameters, particularly the total number concentration, and worst for the complex refractive
5 index, and provide a reduction of the uncertainty from the conservative a priori values for all five variables. We find
6 that the total number concentration and particle median radius can be affected by cross-talk which increases the true
7 uncertainty beyond the propagated standard deviation, for some parts of the state space, related to limited sensitivity
8 of the lidar measurements to particle radii smaller than about 50 nm. We recommend limiting the radii in the retrieval
9 to a range where the measurements have greater sensitivity, to address the high correlation between total number
10 concentration and the particle median radius.

11 In general, information about the state vector that is not provided by the measurements comes from assumptions,
12 constraints, or other a priori information. Smoothing and regularization are examples of retrieval constraints, as is the
13 idea of limiting the minimum particle radius. Retrieval constraints and assumptions can also be hidden or difficult to
14 characterize. For specific retrieval methodologies, we would like to emphasize the importance of explicitly describing
15 any prior information and constraints that affect retrieval results.

16 In this sensitivity study, only very conservative constraints were used in order to pinpoint the sensitivity of the
17 measurements. To achieve better performance with a retrieval, three strategies can be adopted either singly or in
18 combination: 1. Add a priori information that constrains the retrieval using known information about the observed
19 aerosol. 2. Reduce the measurement uncertainties. 3. Add additional measurements to the system.

20 One method to assign a priori covariance information is to use aerosol classification from the lidar intensive parameters
21 (Burton et al., 2012) to infer what type of aerosol is present and then assign prior variances for the state parameters
22 that are specific to that aerosol type. It has been demonstrated that the lidar intensive parameters from an HSRL lidar
23 have sufficient information content to categorize aerosol into broad categories. Assigning a priori values based on
24 these categories additionally requires representative information about the microphysical properties of aerosols in each
25 category from in situ measurements or from modeling.

26 Reducing the measurement uncertainty involves either designing the observing system to stricter requirements (to the
27 extent practical) or reworking the retrieval problem to make more optimal use of the measurement information. For
28 example, a simultaneous profile retrieval that uses the $3\beta + 2\alpha$ lidar information from the whole column with
29 appropriate constraints on the correlations between levels is likely to have somewhat improved information content
30 compared to the baseline uncertainties for the level-by-level retrieval system discussed in this work.

31 Finally, measurement information content can be increased by adding more measurements to the system, for example
32 by combining coincident lidar plus polarimeter measurements from the same platform. This combination is expected
33 to add significantly more information content and reduce the need for constraints or a priori information

34 Research is ongoing into each of the three retrieval strategies described above, aerosol-type-specific prior covariance
35 matrices, profile retrievals, and combined lidar plus polarimeter retrievals. Additional sensitivity studies for these
36 scenarios will be performed in the future.



1 **Acknowledgements**

- 2 Funding for this research came from the NASA Aerosol/Clouds/Ecosystems project, NASA Earth Science Division's
- 3 Remote Sensing Theory program, and NASA Radiation Science Program.



1 **Tables**

2 **Table 1. State variables and selected derived variables for five constructed reference cases.**

	Case 1 (urban)	Case 2 (larger particles)	Case 3 (coarse mode)	Case 4 (absorbing)	Case 5 (large number conc.)
Total number conc. (cm ⁻³)	1101	1101	50	1101	20001
Median radius (µm)	0.12	0.15	0.20	0.12	0.12
Geom. Std. dev.	1.48	1.58	2.48	1.48	1.48
Effective radius (µm)	0.17	0.24	1.60	0.17	0.17
Effective variance	0.16	0.23	1.27	0.16	0.16
Real refractive index	1.47	1.47	1.47	1.61	1.61
Imaginary refractive index	0.003	0.003	0.003	0.03	0.03
Single Scat Albedo (532 nm)	0.98	0.98	0.91	0.89	0.89
Lidar ratio (532 nm) (sr)	72	62	27	81	81

3

4 **Table 2. Propagated state error covariance matrix for the first reference case, assuming measurement errors of 5% for backscatter and 20% for extinction and a priori covariance as described in Section 6 and Table 3.**

5

	Median radius	Geometric standard deviation	Total number concentration	RRI	IRI
Median radius	0.0027	-0.0083	-55.2	-0.0036	-0.0004
Geometric standard deviation	-0.0083	0.032	134	0.014	0.0024
Total number concentration	-55.2	134	1.30e6	43.0	2.29
RRI	-0.0036	0.014	43.0	0.010	0.0018
IRI	-0.0004	0.0024	2.29	0.0018	0.0003

6



1 Table 3. Propagated uncertainties (standard deviations) for state variables and selected additional variables derived from the state variables, shown for the reference cases
 2 described in Table 1. The uncertainties are shown as absolute value for all variables with relative uncertainty in parenthesis for the size distribution variables. The
 3 propagated uncertainties, Eq. (9), depend on assumed measurement errors of 5% for backscatter and 20% for extinction and depend on a priori covariance as described
 4 in the text. The assumed a priori uncertainty and the requirements described in the ACE white paper (also one standard deviation) are listed for comparison.

Retrieval state variables	ACE requirement	Prior uncertainty	Case 1:		Case 2:		Case 3:		Case 4:		Case 5:	
			propagated uncertainty	propagated uncertainty	propagated uncertainty	propagated uncertainty	propagated uncertainty	propagated uncertainty	propagated uncertainty	propagated uncertainty		
Median radius	--	0.30 μm	0.05 μm (46%)	0.07 μm (47%)	0.18 μm (88%)	0.04 μm (41%)	0.04 μm (31%)					
Geometric standard dev.	--	0.6	0.18 (12%)	0.20 (13%)	0.50 (20%)	0.11 (7%)	0.10 (6%)					
Total number concentration	100% (vertically resolved)	20000 cm^{-3}	1142 cm^{-3} (103%)	1074 cm^{-3} (98%)	61 cm^{-3} (122%)	1035 cm^{-3} (94%)	13712 cm^{-3} (68%)					
Real Refractive Index	0.02 (column)	0.19	0.10	0.06	0.04	0.14	0.13					
Imaginary Ref. Index	--	0.050	0.018	0.018	0.005	0.024	0.024					
Derived variables												
Effective radius	10% (column)		0.05 μm (29%)	0.06 μm (23%)	0.80 μm (50%)	0.05 μm (31%)	0.04 μm (23%)					
Effective variance	50% (column)		0.11 (68%)	0.14 (61%)	0.83 (65%)	0.07 (41%)	0.06 (36%)					
SSA (532 nm)	0.02 (vertically resolved)		0.10	0.10	0.06	0.08	0.08					



1 **Table 4. Propagated uncertainties for Case 1, expressed as absolute values and percentage (for the size distribution**
 2 **parameters) for three different theoretical instrument configurations with different backscatter and extinction**
 3 **uncertainties. The last column repeats the draft requirements from the ACE white paper as in Table 3 for reference.**

Retrieved variables	10% / 30%	5% / 20%	5% / 10%	ACE draft requirement
Total number	1677 cm ⁻³	1142 cm ⁻³	700 cm ⁻³	N/A
Concentration	152.4%	103.7%	63.6%	100%
Effective Radius	0.07 μm	0.05 μm	0.03 μm	N/A
	41.0%	28.6%	17.2%	10% (column)
Effective Variance	0.15	0.11	0.08	N/A
	93.1%	67.7%	49.7%	50% (column)
Real Refractive Index	0.12	0.10	0.07	0.02 (column)
Single Scattering Albedo (532 nm)	0.12	0.10	0.07	0.02

4

5



1 **Table 5. State correlation matrix derived from the covariance matrix shown in Table 2, showing the correlations between**
 2 **retrieved variables for Case 1, assuming measurement errors of 5% for backscatter and 20% for extinction and a priori**
 3 **uncertainties from Table 3.**

	Median radius	Geometric standard deviation	Total number concentration	RRI	IRI
Median radius	1	-0.88	-0.92	-0.67	-0.47
Geometric standard deviation	-0.88	1	0.65	0.78	0.72
Total number concentration	-0.92	0.65	1	0.37	0.11
RRI	-0.67	0.78	0.37	1	0.93
IRI	-0.47	0.72	0.11	0.93	1

4

5

6

7 **Table 6. Correlation matrix of the retrieved variables for Case 2, assuming the same measurement errors of 5% for**
 8 **backscatter and 20% for extinction and a priori uncertainties listed in Table 3.**

	Median radius	Geometric standard deviation	Total number concentration	RRI	IRI
Median radius	1	-0.97	-0.97	0.23	0.51
Geometric standard deviation	-0.97	1	0.88	-0.14	-0.42
Total number concentration	-0.97	0.88	1	-0.38	-0.62
RRI	0.23	-0.14	-0.38	1	0.95
IRI	0.52	-0.42	-0.62	0.95	1

9



1 Figures

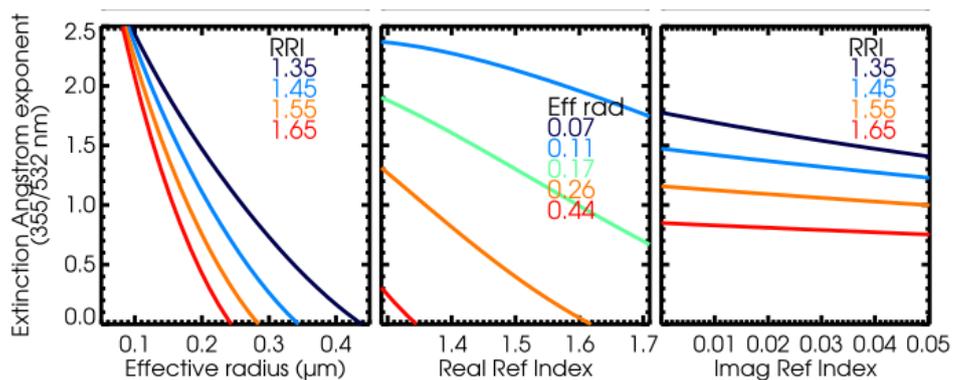


Figure 1. The sensitivity of the extinction Ångström exponent to the effective radius and complex refractive index is shown. The left panel shows the dependence of extinction Ångström exponent (y-axis) on effective radius (x-axis) for four values of the real refractive index (colors); the imaginary refractive index is held fixed at 0.003 and the geometric standard deviation is held fixed at 1.48 (values from Case 1 in Table 1). The middle panel shows the dependence on real refractive index along the x-axis, parameterized by effective radius. The right panel shows the dependence on the imaginary refractive index (x-axis) for four values of the real refractive index; in this case, median radius is held fixed at 0.12 μm and geometric standard deviation is held fixed at 1.48 (values also from Case 1 in Table 1).

2

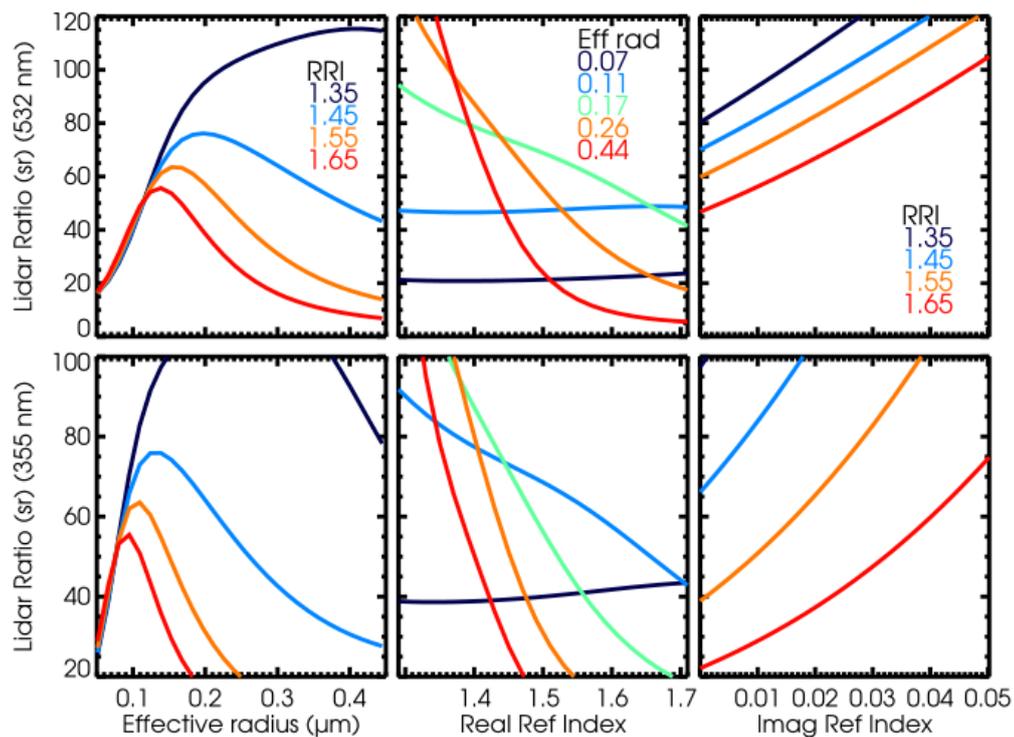


Figure 2. Like Figure 1 but for 532 nm lidar ratio (top row) and 355 nm lidar ratio (bottom row).

1

2

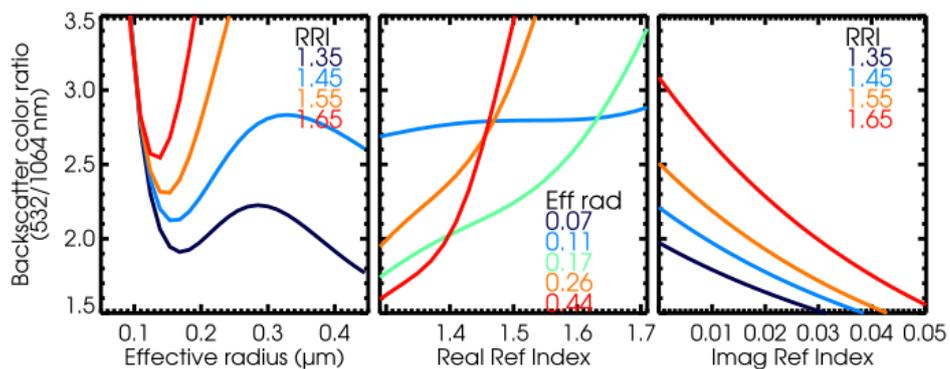


Figure 3. Like Figure 1 but for the backscatter color ratio (which is the ratio of the aerosol backscatter coefficient at 532 nm divided by the aerosol backscatter coefficient at 1064 nm).

1
2

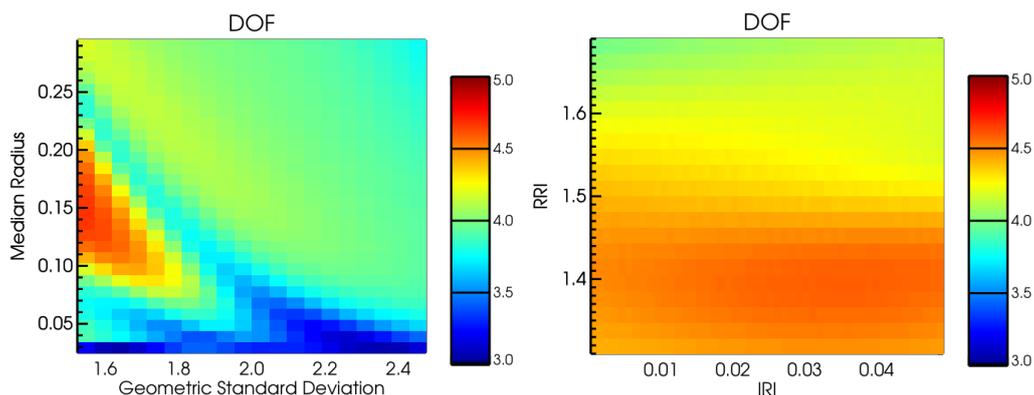


Figure 4. The degrees of freedom (DOF) of the signal, d_s , is shown color-coded, as orthogonal 2-D slices through the 5 variable state space. The left graph shows the dependence on median radius and geometric standard deviation, with the complex refractive index held fixed as $1.47-0.00325i$ and the total number concentration held fixed at 1001 cm^{-3} . The right graph shows the dependence on the complex refractive index (RRI = real refractive index and IRI = imaginary refractive index) with the total number concentration held fixed at 1001 cm^{-3} , the median radius = $0.115 \mu\text{m}$, and the geometric standard deviation = 1.475 . Dependence on total number concentration is very slight and is not illustrated here.

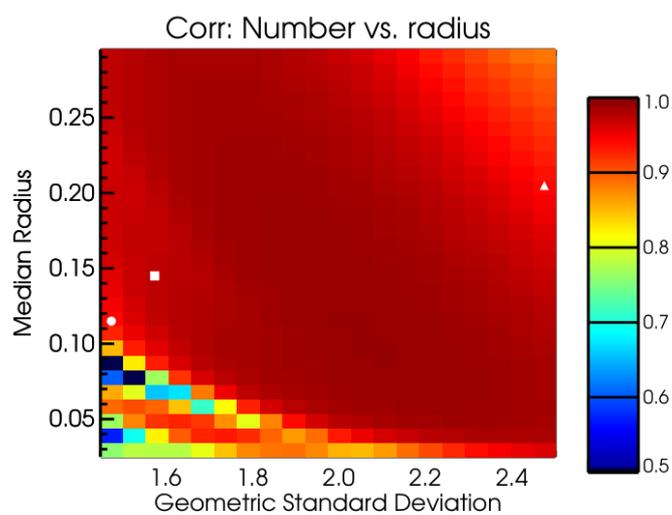


Figure 5. The a posteriori correlation between retrieved total number concentration and median radius is here shown as a 2-D slice through the five-variable state space. The complex refractive index is held fixed at $1.470-0.00325i$, the total number concentration is held fixed at 1001 cm^{-3} , and the dependence on median radius and geometric standard deviation is depicted. Symbols show the values of median radius and geometric standard deviation for Cases 1 (circle), 2 (square) and 3 (triangle), which also have the same complex refractive index as the illustrated slice.

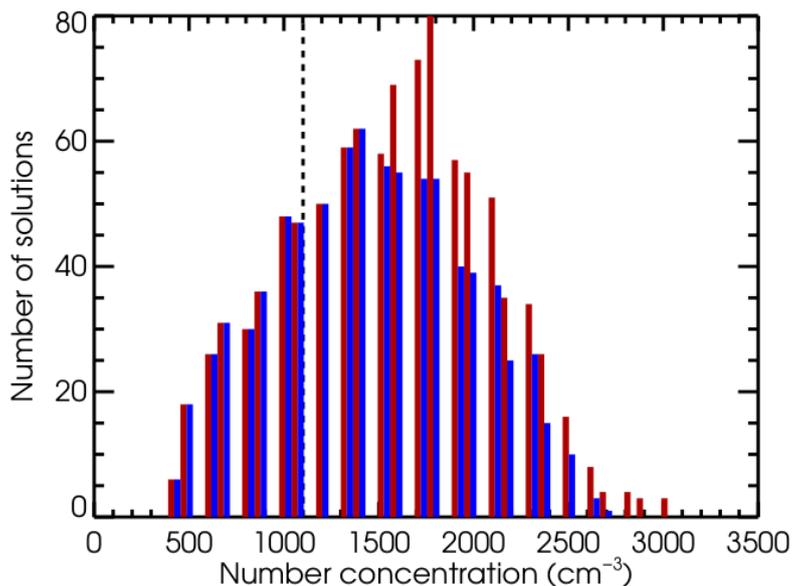


Figure 6. Histograms showing the total number concentration value for all solutions in the gridded LUT (i.e. without interpolation) that match the backscatter and extinction coefficients of Case 1 within measurement errors of 5% for backscatter and 20% for extinction. The total number concentration value for Case 1 is marked with a dashed line. Red histogram bars show solutions from the full gridded LUT. Blue histogram bars show solutions from the modified LUT which excludes size distributions that have a significant contribution from particles of smaller than 50 nm radius.

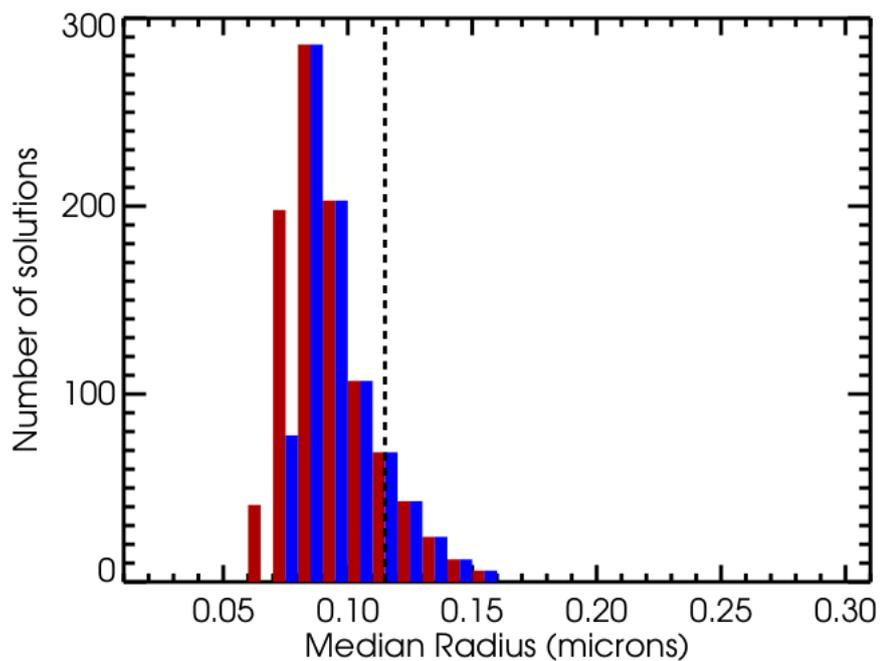


Figure 7. Histograms showing the median radius for all solutions in the gridded LUT that match the backscatter and extinction coefficients of Case 1 within measurement errors of 5% for backscatter and 20% for extinction. The dashed line indicates the median radius for Case 1. Red and blue are as in Figure 6.

1

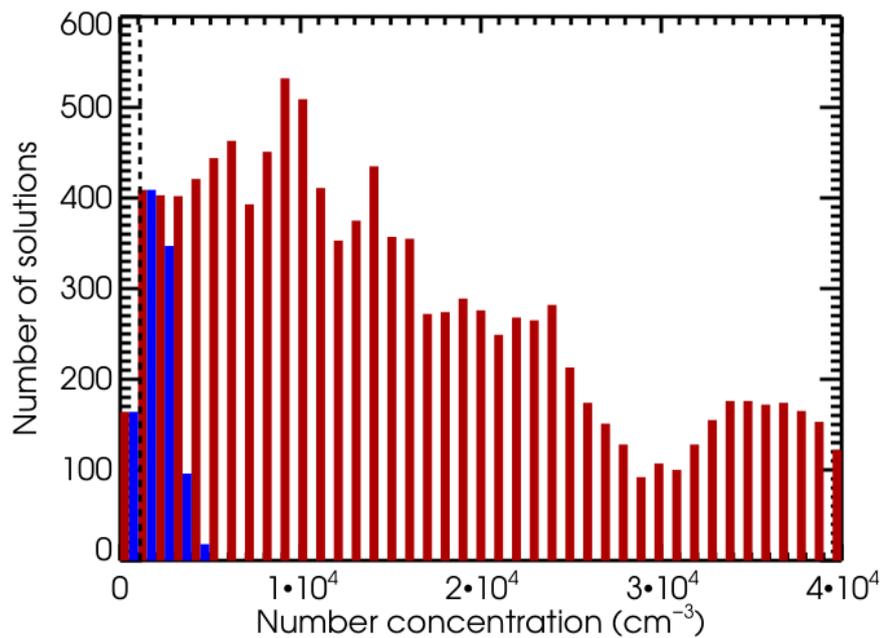


Figure 8. Like Figure 6 but for Case 2.

1

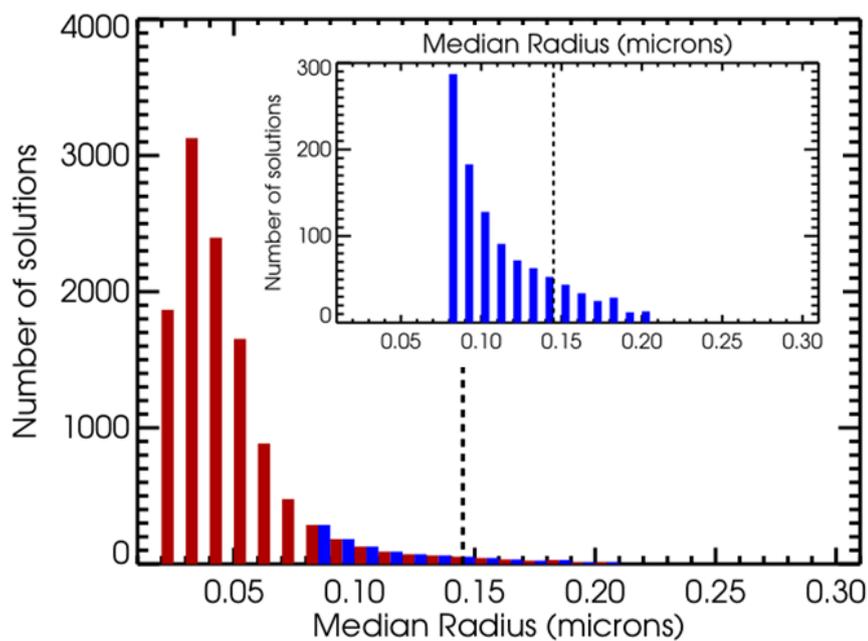


Figure 9. Like Figure 7 but for Case 2. The inset box shows the blue histograms (reduced solution set) with an expanded y-axis scale for better readability.

1

2

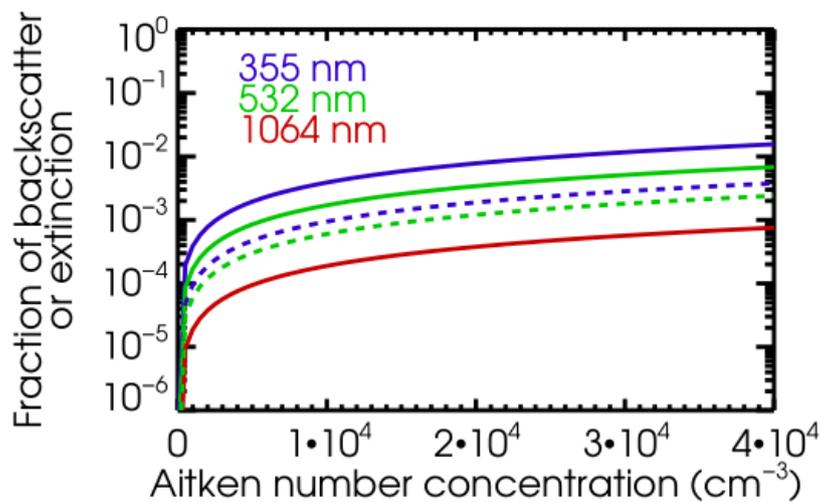


Figure 10. The effect on the backscatter coefficient (solid lines) and extinction coefficient (dashed lines) of a narrow Aitken mode with median radius = 15 nm, $s = 1.48$, and varying number concentration, expressed as a fraction backscatter and extinction measured in a typical observation by HSRL-2 on July 17, 2012 during the TCAP campaign.

1
2

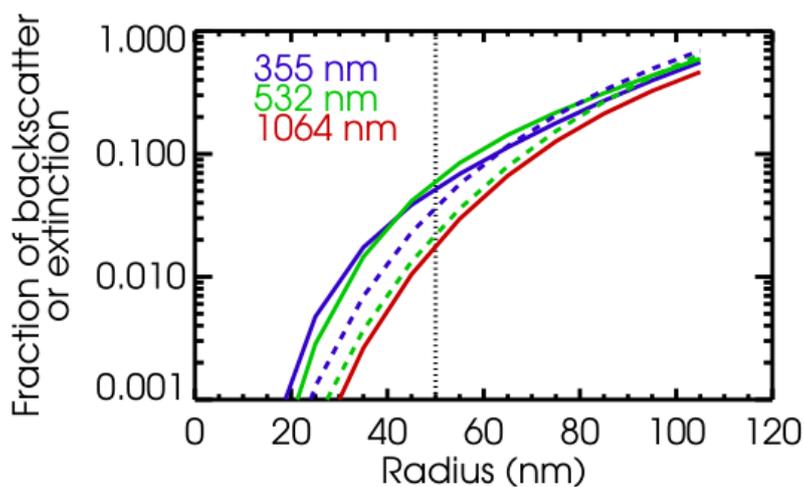


Figure 11. The effect on the backscatter coefficient (solid lines) and extinction coefficient (dashed lines) of 1000 cm^{-3} particles in a narrow mode ($s=1.48$) with varying median radius, expressed as a fraction of the backscatter and extinction measured in a typical observation by HSRL-2 on July 17, 2012 during the TCAP campaign. The dotted line indicates the 50 nm particle radius cutoff discussed in the text.

1

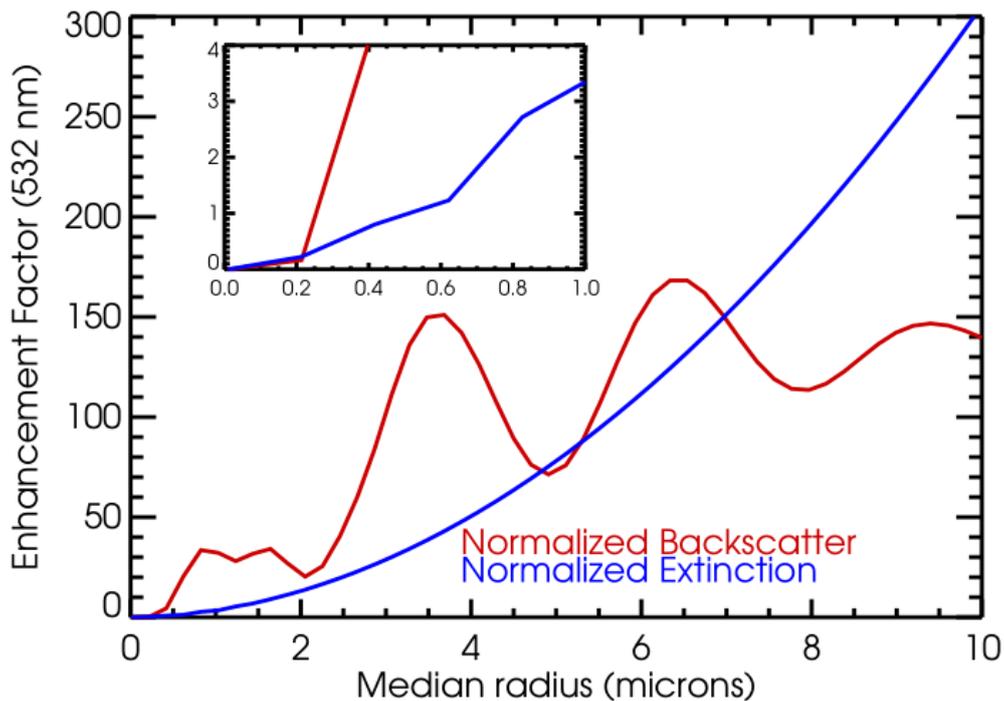


Figure 12. The effect on the backscatter (red line) and extinction (blue line) coefficients at 532 nm of a narrow mode (geometric standard deviation 1.1) of large particles (radius varies along x-axis) with number concentration 36 cm^{-3} and complex refractive index $1.57 - i0.0037$, using the same single scattering Mie modeling as before. The backscatter and extinction coefficients of the modeled coarse mode are in this case much larger than the benchmark values measured by HSRL-2 on July 17, 2012 during the TCAP campaign. The y-axis is expressed as the coarse mode backscatter or extinction divided by the benchmark measurement.

1
2



1 References

- 2 ACE Science Working Group: Aerosol, Cloud and Ecosystems (ACE) Proposed Satellite Mission, NASA, Study
3 2010.
- 4 Ångström, A.: On the Atmospheric Transmission of Sun Radiation and on Dust in the Air, *Geografiska Annaler*, 11,
5 156-166, 10.2307/519399, 1929.
- 6 Berg, L. K., Fast, J. D., Barnard, J. C., Burton, S. P., Cairns, B., Chand, D., Comstock, J. M., Dunagan, S., Ferrare, R.
7 A., Flynn, C. J., Hair, J. W., Hostetler, C. A., Hubbe, J., Jefferson, A., Johnson, R., Kassianov, E. I., Kluzek,
8 C. D., Kollias, P., Lamer, K., Lantz, K., Mei, F., Miller, M. A., Michalsky, J., Ortega, I., Pekour, M., Rogers,
9 R. R., Russell, P. B., Redemann, J., Sedlacek, A. J., Segal-Rosenheimer, M., Schmid, B., Shilling, J. E.,
10 Shinozuka, Y., Springston, S. R., Tomlinson, J. M., Tyrrell, M., Wilson, J. M., Volkamer, R., Zelenyuk, A.,
11 and Berkowitz, C. M.: The Two-Column Aerosol Project: Phase I—Overview and impact of elevated aerosol
12 layers on aerosol optical depth, *Journal of Geophysical Research: Atmospheres*, 336–361,
13 10.1002/2015JD023848, 2015.
- 14 Bockmann, C.: Hybrid regularization method for the ill-posed inversion of multiwavelength lidar data in the retrieval
15 of aerosol size distributions, *Appl Optics*, 40, 1329-1342, 2001.
- 16 Bockmann, C., Miranova, C., Muller, D., Scheidenbach, L., and Nessler, R.: Microphysical aerosol parameters from
17 multiwavelength lidar, *J. Opt. Soc. Am-A*, 518-528, 10.1364/JOSAA.22.000518, 2005.
- 18 Bohren, C. F., and Huffman, D. R.: *Absorption and Scattering of Light by Small Particles*, edited, p. 530, John Wiley,
19 Hoboken, NJ, 1983.
- 20 Burton, S. P., Ferrare, R. A., Hostetler, C. A., Hair, J. W., Rogers, R. R., Obland, M. D., Butler, C. F., Cook, A. L.,
21 Harper, D. B., and Froyd, K. D.: Aerosol Classification of Airborne High Spectral Resolution Lidar
22 Measurements – Methodology and Examples, *Atmospheric Measurement Techniques*, 5, 73-98,
23 10.5194/amt-5-73-2012, 2012.
- 24 Burton, S. P., Hair, J. W., Kahnert, M., Ferrare, R. A., Hostetler, C. A., Cook, A. L., Harper, D. B., Berkoff, T. A.,
25 Seaman, S. T., Collins, J. E., Fenn, M. A., and Rogers, R. R.: Observations of the spectral dependence of
26 linear particle depolarization ratio of aerosols using NASA Langley airborne High Spectral Resolution Lidar,
27 *Atmos. Chem. Phys.*, 15, 13453-13473, 10.5194/acp-15-13453-2015, 2015.
- 28 Chemyakin, E., Müller, D., Burton, S., Kolgotin, A., Hostetler, C., and Ferrare, R.: Arrange and average algorithm for
29 the retrieval of aerosol parameters from multiwavelength high-spectral-resolution lidar/Raman lidar data,
30 *Appl Optics*, 53, 7252-7266, 10.1364/AO.53.007252, 2014.
- 31 Chemyakin, E., Burton, S., Kolgotin, A., Müller, D., Hostetler, C., and Ferrare, R.: Retrieval of aerosol parameters
32 from multiwavelength lidar: investigation of the underlying inverse mathematical problem, *Appl Optics*, 55,
33 2188-2202, 10.1364/AO.55.002188, 2016.
- 34 Cheung, H. C., Chou, C. C. K., Huang, W. R., and Tsai, C. Y.: Characterization of ultrafine particle number
35 concentration and new particle formation in an urban environment of Taipei, Taiwan, *Atmos. Chem. Phys.*,
36 13, 8935-8946, 10.5194/acp-13-8935-2013, 2013.



- 1 Costabile, F., Barnaba, F., Angelini, F., and Gobbi, G. P.: Identification of key aerosol populations through their size
2 and composition resolved spectral scattering and absorption, *Atmos. Chem. Phys.*, 13, 2455-2470,
3 10.5194/acp-13-2455-2013, 2013.
- 4 Donovan, D. P., and Carswell, A. I.: Principal component analysis applied to multiwavelength lidar aerosol backscatter
5 and extinction measurements, *Appl Opt*, 36, 9406-9424, 10.1364/AO.36.009406, 1997.
- 6 Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J. L., Ducos, F., Sinyuk, A., and Lopatin, A.:
7 Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-
8 angle polarimetric satellite observations, *Atmos. Meas. Tech.*, 4, 975-1018, 10.5194/amt-4-975-2011, 2011.
- 9 Gasteiger, J., and Freudenthaler, V.: Benefit of depolarization ratio at $\lambda = 1064$ nm for the retrieval of the aerosol
10 microphysics from lidar measurements, *Atmos. Meas. Tech.*, 7, 3773-3781, 10.5194/amt-7-3773-2014, 2014.
- 11 Hair, J. W., Hostetler, C. A., Cook, A. L., Harper, D. B., Ferrare, R. A., Mack, T. L., Welch, W., Izquierdo, L. R., and
12 Hovis, F. E.: Airborne High Spectral Resolution Lidar for profiling aerosol optical properties, *Appl Optics*,
13 47, 6734-6752, 10.1364/AO.47.006734, 2008.
- 14 Hasekamp, O. P., Litvinov, P., and Butz, A.: Aerosol properties over the ocean from PARASOL multiangle
15 photopolarimetric measurements, *Journal of Geophysical Research: Atmospheres*, 116, D14204,
16 10.1029/2010JD015469, 2011.
- 17 Hoppel, W., Frick, G., and Larson, R.: Effect of nonprecipitating clouds on the aerosol size distribution in the marine
18 boundary layer, *Geophys Res Lett*, 13, 125-128, 1986.
- 19 Kaufman, Y. J., Gitelson, A., Karnieli, A., Ganor, E., Fraser, R. S., Nakajima, T., Mattoo, S., and Holben, B. N.: Size
20 distribution and scattering phase function of aerosol particles retrieved from sky brightness measurements,
21 *Journal of Geophysical Research: Atmospheres*, 99, 10341-10356, 10.1029/94JD00229, 1994.
- 22 Knobelspiesse, K., Cairns, B., Mishchenko, M., Chowdhary, J., Tsigaridis, K., van Diedenhoven, B., Martin, W.,
23 Ottaviani, M., and Alexandrov, M.: Analysis of fine-mode aerosol retrieval capabilities by different passive
24 remote sensing instrument designs, *Optics Express*, 20, 21457-21484, 10.1364/OE.20.021457, 2012.
- 25 Kolgotin, A., and Müller, D.: Theory of inversion with two-dimensional regularization: profiles of microphysical
26 particle properties derived from multiwavelength lidar measurements, *Appl Optics*, 47, 4472-4490,
27 10.1364/AO.47.004472, 2008.
- 28 Mozurkewich, M., Chan, T. W., Aklilu, Y. A., and Verheggen, B.: Aerosol particle size distributions in the lower
29 Fraser Valley: evidence for particle nucleation and growth, *Atmos. Chem. Phys.*, 4, 1047-1062, 10.5194/acp-
30 4-1047-2004, 2004.
- 31 Müller, D., Wandinger, U., and Ansmann, A.: Microphysical particle parameters from extinction and backscatter lidar
32 data by inversion with regularization: theory, *Appl Optics*, 38, 2346-2357, 10.1364/AO.38.002346, 1999.
- 33 Müller, D., Hostetler, C. A., Ferrare, R. A., Burton, S. P., Chemyakin, E., Kolgotin, A., Hair, J. W., Cook, A. L.,
34 Harper, D. B., Rogers, R. R., Hare, R. W., Cleckner, C. S., Obland, M. D., Tomlinson, J., Berg, L. K., and
35 Schmid, B.: Airborne Multiwavelength High Spectral Resolution Lidar (HSRL-2) observations during TCAP
36 2012: vertical profiles of optical and microphysical properties of a smoke/urban haze plume over the
37 northeastern coast of the US, *Atmos. Meas. Tech.*, 7, 3487-3496, 10.5194/amt-7-3487-2014, 2014.



- 1 National Research Council: Earth Science and Applications from Space: National Imperatives for the Next Decade
2 and Beyond, The National Academies Press, Washington, D.C., 400 pp., 2007.
- 3 Pérez-Ramírez, D., Whiteman, D. N., Veselovskii, I., Kolgotin, A., Korenskiy, M., and Alados-Arboledas, L.: Effects
4 of systematic and random errors on the retrieval of particle microphysical properties from multiwavelength
5 lidar measurements using inversion with regularization, *Atmos. Meas. Tech.*, 6, 3039-3054, 10.5194/amt-6-
6 3039-2013, 2013.
- 7 Posselt, D. J., and Mace, G. G.: MCMC-Based Assessment of the Error Characteristics of a Surface-Based Combined
8 Radar–Passive Microwave Cloud Property Retrieval, *J Appl Meteorol Clim*, 53, 2034-2057, 10.1175/JAMC-
9 D-13-0237.1, 2014.
- 10 Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, Series on Atmospheric, Oceanic
11 and Planetary Physics, 2, edited by: Taylor, F. W., World Scientific, New Jersey, 2000.
- 12 Schuster, G. L., Dubovik, O., and Holben, B. N.: Angstrom exponent and bimodal aerosol size distributions, *J.*
13 *Geophys. Res.*, 111, D07207, 10.1029/2005jd006328, 2006.
- 14 Seinfeld, J., and Pandis, S.: Atmospheric Chemistry and Physics, A Wiley-Inter Science Publication, John Wiley &
15 Sons Inc, New York, 2006.
- 16 Veselovskii, I., Kolgotin, A., Griaznov, V., Müller, D., Wandinger, U., and Whiteman, D. N.: Inversion with
17 regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding, *Appl*
18 *Optics*, 41, 3685-3699, 10.1364/AO.41.003685, 2002.
- 19 Veselovskii, I., Dubovik, O., Kolgotin, A., Lapyonok, T., Di Girolamo, P., Summa, D., Whiteman, D. N., Mishchenko,
20 M., and Tanré, D.: Application of randomly oriented spheroids for retrieval of dust particle parameters from
21 multiwavelength lidar measurements, *Journal of Geophysical Research: Atmospheres*, 115, D21203,
22 10.1029/2010JD014139, 2010.
- 23 Veselovskii, I., Dubovik, O., Kolgotin, A., Korenskiy, M., Whiteman, D. N., Allakhverdiev, K., and Huseyinoglu, F.:
24 Linear estimation of particle bulk parameters from multi-wavelength lidar measurements, *Atmos. Meas.*
25 *Tech.*, 5, 1135-1145, 10.5194/amt-5-1135-2012, 2012.
- 26 Veselovskii, I., Goloub, P., Podvin, T., Bovchaliuk, V., Derimian, Y., Augustin, P., Fourmentin, M., Tanre, D.,
27 Korenskiy, M., Whiteman, D. N., Diallo, A., Ndiaye, T., Kolgotin, A., and Dubovik, O.: Retrieval of optical
28 and physical properties of African dust from multiwavelength Raman lidar measurements during the
29 SHADOW campaign in Senegal, *Atmos. Chem. Phys.*, 16, 7013-7028, 10.5194/acp-16-7013-2016, 2016.
- 30 Xu, X., and Wang, J.: Retrieval of aerosol microphysical properties from AERONET photopolarimetric
31 measurements: 1. Information content analysis, *Journal of Geophysical Research: Atmospheres*, 120, 7059-
32 7078, 10.1002/2015JD023108, 2015.

33

34