METHODS OF ASSESSING UNCERTAINTY IN AEROSOL PROPERTIES AS INTERPRETED FROM LIDAR MEASUREMENTS

by

BENJAMIN R. HERMAN

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2008

UMI Number: 3296967

UMI Microform 3296967

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the
Graduate Faculty in Engineering in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Barry Gross

Date — Chair of Examining Committee

Mumtaz K. Kassir

Date — Executive Officer

Fred Moshary

Samir Ahmed

Yonghua Wu

Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

METHODS OF ASSESSING UNCERTAINTY IN AEROSOL PROPERTIES AS INTERPRETED FROM LIDAR MEASUREMENTS

by

Benjamin R. Herman

Advisers: Professors Barry Gross and Fred Moshary

This dissertation is an investigation into how measurements of backscatter and extinction from lidar can be interpreted to gain information about other aerosol properties by determining uncertainties of aerosol size distributions and refractive indexes. A derivation of aerosol optical properties from particle size distributions and their resulting effect on lidar signals is shown. Procedures for retrieving optical properties from lidar signals are shown and an analysis of the dependency of optical coefficient retrieval error on true value is performed. Methods of retrieving aerosol size distributions including linear basis representation and parameterization are discussed. A graphical uncertainty analysis is discussed and applied to analyzing the effects on uncertainty when extinction measurements are included to a three-wavelength backscatter measurement set in an assumed index of refraction model. The problems associated with using a mono-modal model in the presence of a bimodal aerosol distribution are analyzed.

The Bayesian inverse method is introduced and a reverse Monte Carlo method based on conditional probability is used to illustrate discrepancies from a conventional forward Monte Carlo method. An analytic formula for the color ratio retrieval error probability density function (PDF) is derived for computing the posterior PDF. Discrepancies between posterior cumulative distribution functions (CDFs) and estimated

CDFs from forward Monte Carlo methods using Bayesian and non-Bayesian retrieval methods are demonstrated and quantified for an array of scenarios. General differences in marginal and measurement-conditional error PDFs are formulated. Posterior PDFs are used to derive credible sets of aerosol distribution parameters including complex index of refraction from synthetic optical coefficient measurements.

A plan for using posterior PDFs for uncertainty analysis in high dimensional aerosol models based on the Metropolis-Hastings Markov chain Monte Carlo algorithm is presented. Preliminary results of its application in the parameterized model are shown and compared with analytic posterior CDFs.

Acknowledgements

There are so many people who have made this dissertation possible. I thank Barry Gross, Fred Moshary, and Sam Ahmed for all their encouragement, support, and guidance over the years and for providing a laboratory for lidar research. I thank Brian Cairns for suggesting a reference which greatly helped in revising section 3 and for pointing out the depolarization effect in molecular scattering which I had originally neglected. I thank Yonghua Wu for his stimulating conversations which have been helpful to me in the processing of lidar data discussed in section 2. I thank Tom Legbandt for helping me in the design and construction of the amplifier for the infrared detectors, without which the images in figures 2.1 and 2.2 would not be as rich in detail. I thank NASA-COSI for my financial support. It has been a privilege to have had the opportunity to do this work, and for that I am grateful to them. I also thank NOAA-CREST for their enrichment of the scientific culture at CUNY.

No acknowledgements section would be complete without a sappy sentimental thank you to that someone who you'd never expect, so I thank Sally Drob, my 6^{th} and 8^{th} grade math teacher at I.S. 70 of New York City. Her enthusiasm and teaching was inspirational, and I have vivid memories of my first senses of gratification in mathematics which have continued to this day. She truly deserves these thanks.

I thank my mom, Marian Woertz Brawer, and dad, Ed Herman for raising me in an environment that values intellectual pursuits.

Finally I thank my wife, Aimee Cegelka. I cannot overstate the help she has given at the eleventh hours of completing this document. More importantly I thank her for all the love and support she has given me over the past four and a half years since we've met. She has continually been patient with me while I have been working towards this dissertation. No one has ever made me feel as appreciated as much as she.

Table of Contents

List of Tables

List of Figures

1. Introduction

The advent of Q-switched pulsed lasers has led to their application in sensing of atmospheric gases and aerosols with lidar. The use of optical harmonic generation has led to the capability of measuring optical properties of aerosols at multiple wavelengths using a single laser. Multi-wavelength aerosol lidar systems typically use a pulsed Nd:YAG laser producing light at a wavelength of 1064nm and its 2^{nd} and 3^{rd} harmonics at 532 and 355nm, and aerosol backscatter coefficients at these wavelengths can be retrieved through appropriate signal processing of the elastic scattering signal. Furthermore, by detecting the wavelength-shifted light that gets returned to the lidar by Raman-scattering off of the nitrogen in the atmosphere and using knowledge of the atmospheric molecular density, the signal can by analyzed to retrieve aerosol extinction profiles by observing its attenuation.

Networks of multi-wavelength lidars are present in North America (REALM), Europe (EARLINET), and Asia (NIES lidar network). The most comprehensive of these is EARLINET which includes 4 lidars, each with three elastic backscatter channels and at least one Raman channel. Two of these lidars also include a second Raman channel, and one of these lidars also utilizes multiple lasers providing additional elastic backscatter channels at 800, 710, and 400nm. In the REALM network one lidar uses the high spectral resolution technique to distinguish molecular scattering from aerosol scattering by making use of Doppler broadening due to thermal molecular motion to retrieve aerosol extinction rather than using Raman scattered optical returns. The optical remote sensing laboratory at the City College of New York has built and operates a multi-

wavelength lidar system in the REALM network that is currently (January 2008) set up to retrieve backscatter at the three Nd:YAG wavelengths and extinction at 355nm.

This dissertation deals with the problem of retrieving particle size distributions and indexes of refraction of aerosols from multi-wavelength lidar measurements. In section 2 the optical properties of aerosols will be explained along with the retrieval of these properties from lidar measurements and the errors associated with these retrievals. Section 3 gives a brief account of past research into the retrieval of aerosol particle size distributions from optical coefficients that can be retrieved by lidar. The theory behind the past research is explained and used to demonstrate its shortcomings. Section 4 discusses a graphical technique of plotting confidence bound contours that can give simple estimates of uncertainty in model parameters of a unimodal, fixed index of refraction aerosol. The technique uses ratios of optical coefficients so that size parameters can be retrieved without having to consider the number density. It is improved by using error bounds from a probability density function for ratio errors that is derived from Gaussian distributed optical coefficient fractional errors. The graphical uncertainty estimate is used to assess how well uncertainty can be improved by including extinction coefficients to a set of backscatter coefficients for particle size distribution retrieval.

In section 5 the Bayesian approach to uncertainty assessment is introduced. A Monte Carlo method directly based on conditional probability is used to demonstrate that there are significant differences between a Bayesian approach and a simple forward Monte Carlo approach. These differences are explored in depth in section 6 where cumulative distribution functions (CDFs) estimated from forward Monte Carlo sample

ensembles are used for comparison against Bayesian formulated CDFs. A CDF deviation parameter is imaged over a domain of size parameters showing regions where differences are more pronounced. To understand these differences it is shown through probability density function (PDF) analysis that the measurement error PDF takes a different form when the information present in the measurements is considered and any of a number of common non-ideal conditions are met. Two of these conditions which were existent in the study presented in section 6 are: 1) non-linear relationships between retrieval parameters and the underlying optical coefficient values, and 2) measurement error PDFs that are dependent on the retrieval parameters.

Unfortunately this shows a fault with the graphical uncertainty estimate, since the error bounds used in it are valid only for the uninformed error PDF. This problem is circumvented by using the Bayesian posterior PDF to derive credible sets in the aerosol size parameter domain. A procedure for deriving credible sets of both size parameters and complex indexes of refraction is described, and the results of its application to synthetic measurements are shown for a variety of measurement configurations.

This dissertation is concluded by presenting a plan of how to implement Bayesian uncertainty assessments with higher dimensional aerosol models such as a linearized size distribution representation or a multimodal aerosol model. This plan makes use of a Markov chain Monte Carlo algorithm to generate random samples of aerosol parameter retrievals in a manner consistent with the Bayesian posterior PDF. The methodology in the plan is applied to the retrieval of monomodal aerosol model parameters, and CDFs estimated from the random sample ensembles are compared with Bayesian CDFs.

2. Optical Coefficients and Their Retrieval with Lidar

2.1 Optical properties of aerosols

To illustrate optical properties of aerosol as a whole, definitions are presented of the terminology that describes how individual particles behave when exposed to incident light. When a particle is exposed to a plane wave with intensity, I_{inc}, a quantity of power, $W_{\text{sca}} = C_{\text{sca}} I_{\text{inc}}$ is scattered and a quantity, $W_{\text{abs}} = C_{\text{abs}} I_{\text{inc}}$, is absorbed, where C_{sca} , C_{abs}, and their sum $C_{\text{ext}} = C_{\text{sca}} + C_{\text{abs}}$ are known as the scattering and absorption cross-sections. Considering an aerosol made up of spherical particles with different radii r_i, when light is incident on a volume with dimensions $\mathrm{d}A$ in area perpendicular to the incident light and $\mathrm{d}z$ in the direction parallel to the incident light, the parcel of the aerosol removes an amount of power, $\sum_i C_{\text{ext}}(r_i) I_{\text{inc}}$.

If the parcel of aerosol is viewed as having N particles whose sizes are realizations of a random variable, then the extinguished power has an expected value of $N\mathrm{E}\{C_{\text{ext}}\} I_{\text{inc}}$. This then results in the differential equation,

$$\frac{\mathrm{d}I}{\mathrm{d}z} = -\frac{N\mathrm{E}\{C_{\text{ext}}\}}{\mathrm{d}z\mathrm{d}A} I \,, \tag{2.1}$$

whose solution is $I = I(z_0)\exp(-\int_{z_0}^{z} \alpha(z')\mathrm{d}z')$, where

$$\alpha = \int_0^\infty \frac{N}{\mathrm{d}z\mathrm{d}A} f_r(r) C_{\text{ext}}(r, m)\mathrm{d}r \,, \tag{2.2}$$

is known as the extinction coefficient, $f_r(r)$ is the probability density function for particle size, and *m* represents the complex index of refraction which is modeled as being constant for all wavelengths and particles in the distribution. This leads to the definition of the particle size distribution,

$$n(r) = \rho_v f_r(r), \tag{2.3}$$

where

$$\rho_v = \frac{N}{\mathrm{d}z\mathrm{d}A}, \tag{2.4}$$

is the particle density per volume.

A similar derivation can be done for scattered power where an area *A* at a distance of *R* at angle θ is

$$\sum_i \frac{C_{\mathrm{sca}}(r_i)P(\theta)}{R^2} I_{\mathrm{inc}}, \tag{2.5}$$

where $P(\theta)$ is the scattering phase function in sr^{-1}, normalized to integrate to 1. This leads to the definition of the backscatter coefficient,

$$\beta = \int_0^\infty C_{\mathrm{sca}}(r)P(\pi)n(r)\mathrm{d}r. \tag{2.6}$$

This leads to the lidar equation,

$$Y(R) = \frac{EA}{R^2}\beta(R)\exp(-2\int_0^R \alpha(R')\mathrm{d}R'), \tag{2.7}$$

where $Y(R)$ is the energy received per unit range. In relating this to an electric signal in time R is replaced by $\frac{1}{2}ct$ and $Y(R)$ is multiplied by $\frac{1}{2}cG$ to convert to energy per unit time and a detector response that converts power into volts or current.

There is a more complicated process going on in lidar operation. Since the laser beam is relatively narrow both in area and direction, it is assumed that the scattered power does not enter substantially into the receiver's field of view and hence Eq. 2.7 is known as the single scattering equation. This assumption becomes less appropriate for clouds and denser aerosols. In fact these effects are actually exploited by some lidar systems that take measurements at multiple fields of view (Hutt, Bissonnette, and Durand 1994; Bissonette, Roy, and Roy 2005); however, this topic is beyond the scope of this dissertation.

2.2 Processing techniques

To reiterate, the equation for the single scattering elastic lidar signal is

$$P(R) = \frac{ECO(R)}{R^2}\left(\beta_a(R) + \beta_m(R)\right)T_m(R)T_a(R), \tag{2.8}$$

where

$$T_m(R) = \exp(-2\int_0^R \alpha_m(R')\mathrm{d}R') \text{, and } T_a(R) = \exp(-2\int_0^R \alpha_a(R')\mathrm{d}R') \tag{2.9}$$

are the molecular and aerosol 2-way transmittances from the receiver to range R, E is the laser pulse energy, C is the calibration constant, $O(R)$ is the overlap function, $\beta_a(R)$ and $\beta_m(R)$ are the aerosol and molecular backscatter profiles respectively, and $\alpha_a(R)$ and

$\alpha_m(R)$ are the extinction profiles. It is convenient to express extinction as $\alpha = S\beta$, where the extinction-to-backscatter ratio *S* is an intensive parameter specific to a given atmospheric constituent. It is a common assumption in lidar research that the molecular S-ratio is 8π/3 (e.g., Böckmann et al. 2004). However, even for non-absorbing molecules there is a discrepancy in this value due to their anisotropic nature which results in non-zero off-diagonal terms in the molecular polarizability matrix which varies among different orientations. Hence, the induced molecular dipole moments are not entirely parallel to the incident light, and the scattering phase function is deviated with respect to the parallel dipole case (Miles, Lempert, and Forkey 2001). An improved formula for molecular S-ratio is

$$S_m = \frac{8\pi}{3}\left(1 + \frac{\gamma}{1+\gamma}\right), \tag{2.10}$$

which is a direct result from Eq. 12 in (Bucholtz 1995) which gives values of γ as specified in (Bates 1984) for dry air which result in γ/(1+γ) = 1.51%, 1.42%, and 1.37% for λ = 355, 532, and 1000 nm respectively (values were given only for λ between 200 and 1000nm).

The lidar equation has a solution for $\beta_a(R)$ expressed in a way equivalent to Sasano, Browell, and Ismail (1985) as

$$\beta_a(R) = \frac{X(R)\,\mathrm{tc}(R, R_0)}{ECT_m(R_0)T_a(R_0) - 2\int_{R_0}^{R} S_a(R')X(R')\,\mathrm{tc}(R', R_0)\,\mathrm{d}R'} - \beta_m(R) \tag{2.11a}$$

$$\mathrm{tc}(R,R_0)=\exp\left\{-2\int_{R_0}^{R}\left[S_a(R')-S_m\right]\beta_m(R')\mathrm{d}R'\right\}, \tag{2.11b}$$

and

$$X(R)=R^2P(R)/O(R). \tag{2.11c}$$

The molecular backscatter profile can be determined from the single molecule backscatter cross-section and radiosonde measured profiles of temperature and pressure. For the aerosol backscatter presented in this disertation as retrieved from lidar signals taken with the CCNY lidar, the molecular backscatter profile is derived from radiosonde data taken at Brookhaven National Laboratory, 91 km east of our lidar site with molecular cross sections as given by Bucholtz (1995). A value for $ECT_m(R_0)T_a(R_0)$ can be estimated from the lidar signal coming from an aerosol scarce region in the atmosphere with R_0 within that region (Fernald, Herman, and Reagan 1972). There the lidar signal is approximately

$$P(R)=\frac{EC}{R^2}\beta_m(R)T_m(R_0)T_a(R_0)\exp\left(-2\int_{R_0}^{R}\alpha_m(R')\mathrm{d}R\right), \tag{2.12}$$

and therefore

$$\frac{R^2P(R)}{\beta_m(R)}\exp\left(2\int_{R_0}^{R}\alpha_m(R')\mathrm{d}R\right)=ECT_m(R_0)T_a(R_0), \tag{2.13}$$

and the left hand side of Eq. 2.13 is averaged in the aerosol scarce region to get the constant. This procedure works best for the 355 and 532nm backscatter signals, but is often not viable for the 1064nm signal due to the decrease in molecular scattering

according to the λ^{-4} dependency. The 1064nm signal is calibrated using assumed cloud spectral properties and their backscatter measured by the other channels. This is performed in a systematic fashion where a user views an image of the range corrected 1064nm signal and draws the regions where there is cloud and where there is scarce aerosol with a graphic user interface. Low molecular backscatter makes the 1064nm signal the best choice for initial imaging. The cloud region is initially drawn larger than the cloud seen in the image and the final cloud region is determined as beginning when retrieved backscatter from the other channels rises above a threshold over increasing range. The upper cloud calibration boundary is determined by when the range integrated backscatter rises above another threshold or a maximum range span is reached. The fact that the inverse formula is decreasing with increasing $ECT_m(R_0)T_a(R_0)$ is used so that the inverse formula can be repeatedly applied to the 1064nm signal using different values for $ECT_m(R_0)T_a(R_0)$ in a divide-and-conquer algorithm until the regional integrated backscatter matches the assumed value within some tolerance.

An example of the results of lidar data processing are shown in RGB image representations of backscatter coefficients derived from the CCNY lidar in Figures 2.1 and 2.2. These figures are shown as examples of how variations in color ratios (defined as the ratio of two backscatter coefficients at different wavelengths) can quickly be seen in this representation. Readers who are not viewing this in color may skip ahead to subsection 2.3 without loss of continuity.

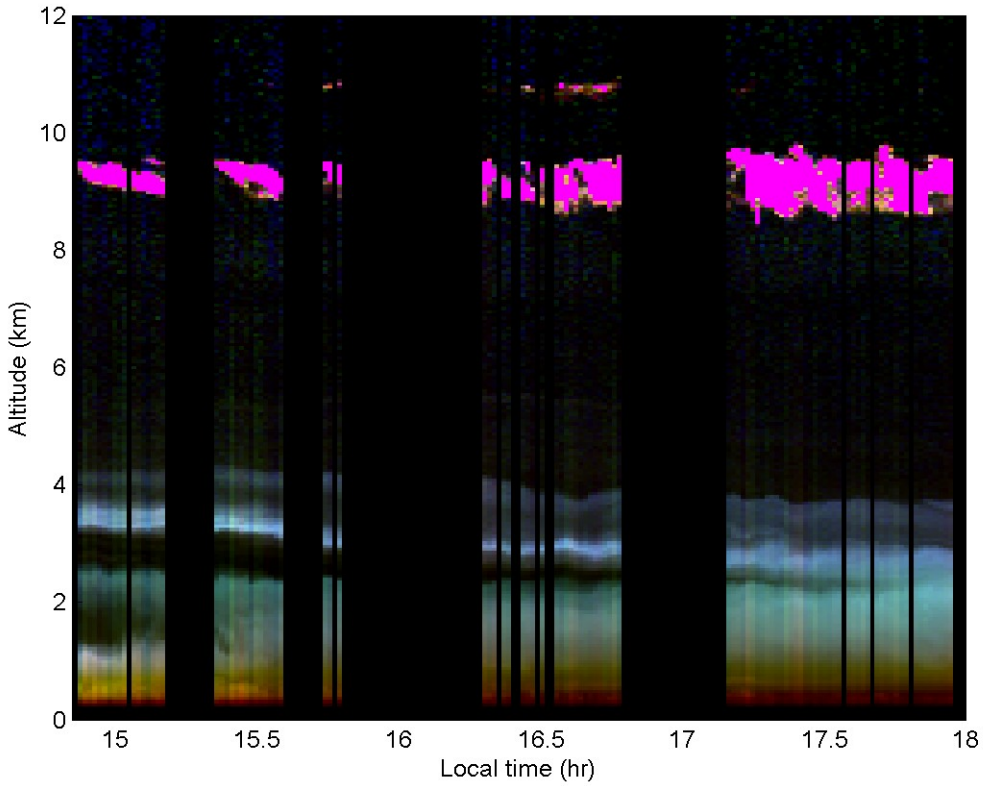

Figure 2.1. RGB image representation of multi-wavelength backscatter coefficients retrieved from lidar on 11-Apr-2006. Unequal weights are used to facilitate viewing variations in color ratios. Red, green, and blue pixel intensities of weighted backscatter at the three wavelengths are computed as,

$$(r, g, b) = (\beta_{1064} / 1.41 , \beta_{532} / 2.00 , \beta_{355} / 2.45) .$$

Pixel saturation is represented as magenta.

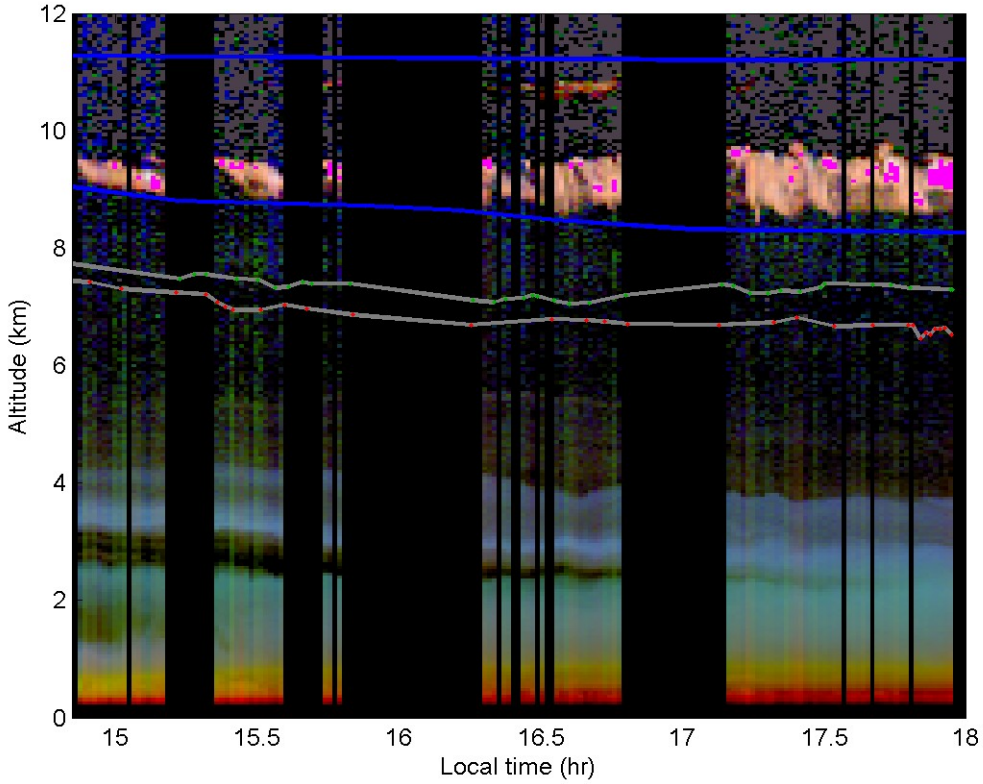

Figure 2.2. Log-scaled RGB image representation of backscatter with user input bounds overlaid. To simultaneously view constituents with different orders of magnitude, this log-scaled image is also shown according to,

$$(R,G,B) = c\left[\log(r+g+b) - L\right]\frac{(r,g,b)}{r+g+b},$$

thereby maintaining the same hue and saturation as in the linearly scaled image; the constants c and L are chosen heuristicly to see the most detail. The altitudes between the gray plots were used for molecular calibration for the 532 and 355 nm signals. Infrared signals were calibrated from cirrus clouds when available. The user input bounds indicated by the blue plots mark the altitudes at which a cirrus cloud S-ratio is used in Eq's 2.11(a, b, and c) instead of an aerosol S-ratio. The cirrus clouds appear orange rather than white due to the unequal weighting of backscatter at different wavelengths. Color ratio variations are much more clearly seen here than in the linear scaled image in regions where backscatter coefficients are smaller.

2.3 Retrieval Error Analysis

A Monte Carlo analysis was performed to see how deviations of the solution are dependent on the true value of aerosol backscatter. The purpose here is not to advocate a comprehensive model of error in optical coefficients retrieved from lidar measurements, but only to demonstrate that using an additive error model independent of the true aerosol distribution is generally not appropriate. The Monte Carlo analysis consisted of trials where lidar signals were generated from fixed aerosol and molecular backscatter profiles and a random S_a which was constant for all ranges. In each trial the signal was inverted using Eq's 2.11(a, b, & c) with a fixed S_a at the mean of S_a's random generator and inverted with a random generated deviation between the assumed and true overlap profile. The molecular backscatter profile was modeled from radiosonde data taken at Brookhaven National Laboratory (station identifier OKX). The aerosol backscatter profile was modeled as a boxcar function,

$$\beta_a(R) = \begin{cases} \beta_0, & 2\,\mathrm{km} \le R \le 2\,\mathrm{km} + \Delta R \\ 0, & \text{otherwise} \end{cases}, \tag{2.14}$$

and the analysis was performed for 50 values of β_0 spaced uniformly between 4×10^{-4}–2×10^{-2} sr^{-1} km^{-1}. This was repeated for ΔR = 0.1, 0.2, and 0.5 km.

S_a was generated by using a random number generator to generate log-normal spherical aerosol distributions with an index of refraction of 1.5 – 0i and with the median radius and GSD being uniformly distributed on [0.01um, 0.9um] and [1.1,2.4] respectively. Lookup tables were then used to interpolate the S_a value. Random

deviation in the assumed from true overlap function was generated by modeling their ratio with a piecewise linear function,

$$\frac{O_{\text{assumed}}(R)}{O_{\text{true}}(R)} = \begin{cases} 1+\varepsilon_{\text{OL}}\left(R-3\,\text{km}\right), & R < 3\,\text{km} \\ 1 \quad , & 3\,\text{km} \le R \end{cases}, \tag{2.15}$$

where ε_{OL} was distributed uniformly over $[-\Delta\varepsilon_{\text{OL}},\Delta\varepsilon_{\text{OL}}]$ repeated for $\Delta\varepsilon_{\text{OL}} = 0.05$, 0.1, and 0.15 km^{-1}. For each β_0, 10 000 solutions were obtained from simulated signals and the retrieved backscatter was averaged over the [2km,2km+ΔR] window. The root mean square deviation between retrieved and true backscatter, $\Delta\beta$, was calculated from the samples of averaged backscatter and the results are shown plotted versus β_0 in Figure 2.3 for the different setups, showing that errors are significantly dependent on the true values.

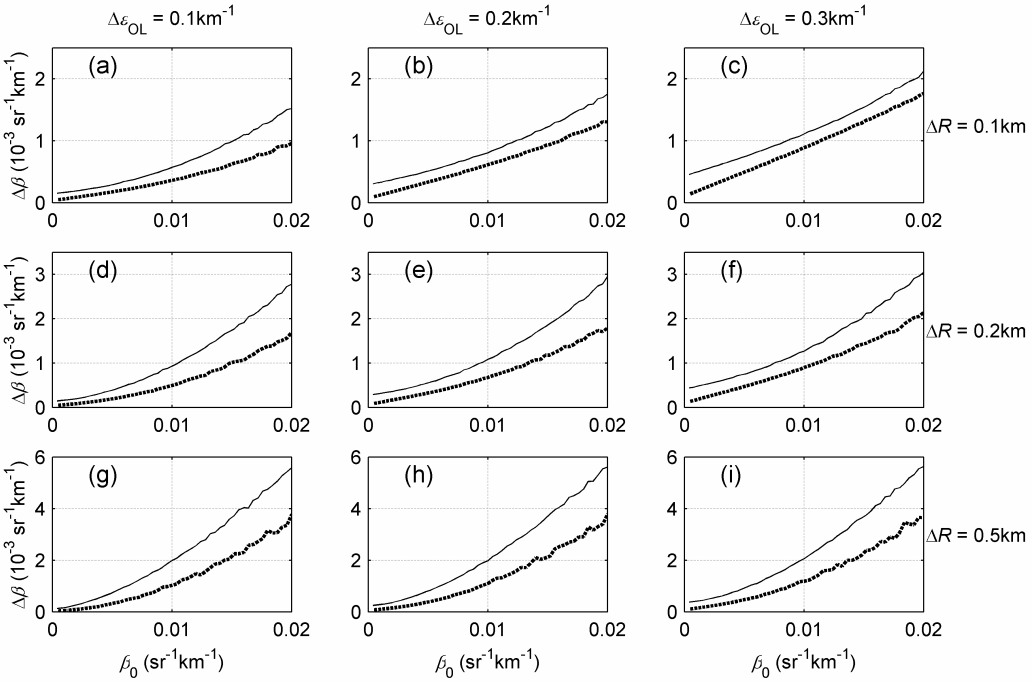

Figure 2.3. RMS backscatter retrieval deviation. Columns of plots have identical $\Delta\varepsilon_{\mathrm{OL}}$s as specified above; Rows of plots have identical ΔRs as specified to the right.

Raman lidar makes use of the inelastic Raman scattering from molecular species where the light that is scattered back is at a wavelength of

$$\lambda_D = \left(\lambda_T{}^{-1} - \lambda_{\text{Raman}}{}^{-1}\right)^{-1}, \tag{2.16}$$

where λ_T is the transmitted wavelength and λ_{Raman} is one of the Raman scattering wavelengths specific to a molecular constituent (Ansmann et al. 1992). This is useful for the detection of trace gases, but can also be used for the retrieval of aerosol extinction by detecting the signal from scattering off of N_2 molecules since the molecular density of nitrogen can be determined from radiosonde profiles.

The expression for the detected Raman signal is

$$P(R) = \frac{EC}{R^2}\beta_{\text{Ram}}(R)T'_m(R)\exp(-\int_0^R \alpha_a(R',\lambda_T) + \alpha_a(R',\lambda_D)\mathrm{d}R'), \tag{2.17}$$

where $\alpha_a(R,\lambda)$ is the aerosol extinction coefficient profile at any given wavelength, $\beta_{\text{Ram}}(R)$ is the Raman backscattering coefficient, and

$$T'_m(R) = \exp(-\int_0^R \alpha_m(R',\lambda_T) + \alpha_m(R',\lambda_D)\mathrm{d}R'), \tag{2.18}$$

is the 2-way molecular transmittance from the receiver to range *R*. It can easily be seen that an inverse formula for combined extinction can be obtained by

$$\alpha_a(R,\lambda_D) + \alpha_a(R,\lambda_D) = \frac{-\frac{\mathrm{d}Y(R)}{\mathrm{d}R}}{Y(R)}, \tag{2.19}$$

where

$$Y(R) = \frac{R^2 P(R)}{\beta_{\mathrm{Ram}}(R) T_m'(R)} . \tag{2.20}$$

Combined extinction at both wavelengths can be used just as readily as extinction at one wavelength for size distribution retrieval since the Fredholm equations are linear in the kernel functions. Thus the kernel for combined extinction is the sum of the kernels for the individual extinctions at λ_T and λ_D. This is obviously preferable to using an assumed Angstrom coefficient to estimate a single wavelength extinction.

Since Raman lidar signals are much weaker than elastic signals and the inverse formula requires estimating a derivative, signal noise is the dominant driver of retrieval error. To investigate the dependence of the statistical properties of retrieval error on true values of extinction, a similar Monte Carlo analysis of retrieval error in Raman-lidar retrieved aerosol extinction was performed. The aerosol extinction profiles used were piecewise-linear with combined extinction equal to 0.97 km^{-1} for R < 1.5km and zero aerosol above 2km with extinction at 1.75km being variable as shown in Figure 2.4a. Gaussian white noise was added to the synthetic signals at critical levels, (i.e., levels where noise is low enough that a meaningful retrieval is possible but retrieval error is substantial). Two noise levels were tried which were 3.2% and 8.0% of the signal at 1.5km. The derivative was estimated by a sliding linear regression with a span of 41 samples, corresponding to range span of 0.15km. Plots of estimated root-mean-square error of range averaged extinction within four different range windows are shown in Figure 2.4b (3.2% noise) and 2.4c (8.0% noise) with window regions indexed as shown in Figure 2.4a. These plots demonstrate that these errors generally have a weaker

dependency on true values than was found for backscatter error. It is suspected that the stronger dependency for window 4 is due to its adjacency to the aerosol free zone.

Therefore, it has been demonstrated that optical coefficient retrieval error is not independent of coefficient true values. Backscatter error dependency results from incorrect S-ratio assumptions and also has the potential to be affected by speckle noise. Raman retrievals of extinction show a weaker dependency. As will be illustrated in sections 5 and 6, an analysis of error dependency is needed for constructing forward models for Bayesian uncertainty assessment.

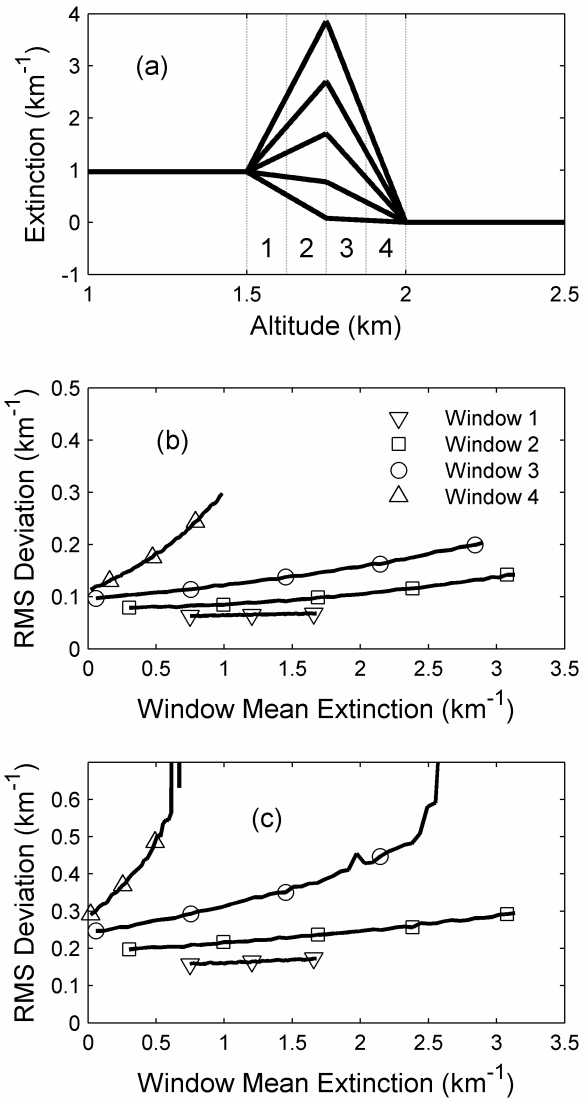

Figure 2.4. Synthetic aerosol extinction profile and averaging windows(a), RMS retrieval deviation for 3.2% (b), and 8.0% noise (c) relative to 1.5km signal.

3. Background on Particle Size Distribution Retrieval

The problem under consideration is to determine what can be inferred about an aerosol's particle size distribution (PSD) from lidar measurements of extinction and backscatter from the equations,

$$y_i = \int_0^\infty K_i(r,m)n(r)dr \quad , \tag{3.1}$$

where K_i are the kernel functions as defined by Eq's 2.3, 2.4, and either 2.2 for extinction or 2.6 for backscatter. To pose the lidar PSD retrieval problem in a way so that it can feasibly be solved by computation, it is necessary to describe the aerosol distribution in terms of a finite set of parameters. One approach is to express the size distribution function as a linear combination of basis functions so that

$$n(r) = \sum_j x_j \bar{n}_j(r) \quad , \tag{3.2}$$

where $\bar{n}_j(r)$ is the *j*th basis function. Eq. 3.1 becomes a matrix equation,

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad , \tag{3.3}$$

with the components of the kernel matrix **A** being

$$A_{i,j} = \int_0^\infty K_i(r,m)\bar{n}_j(r)\mathrm{d}r \quad . \tag{3.4}$$

Most linear inverse methods applied to the lidar PSD retrieval problem employ B-spline basis functions of some order as described by Böckmann (2001). One exception to this was the work done by Donovan and Carswell (1997) who effectively used the kernel functions as the basis with the purpose of retrieving aerosol macroscopic properties rather than the size distribution itself. The B-spline basis functions are non-zero only for particle radii within contiguous segments of a partition that, together with the spline order, defines the basis functions. First order (rectangular) B-spline functions are piecewise constant, 2nd order (triangular) are piecewise linear and continuous over the support of the basis (defined as the region in which there are basis functions with non-zero contributions), 3rd order are piecewise quadratic, continuous, and have continuous 1st derivatives, etc.

The main problem in retrieving particle size distributions from lidar measurements is that there are typically only a small number of wavelengths at which measurements are taken, and therefore the inverse problem is under-determined for general aerosol models. Even the most sophisticated of systems can measure only 6 backscatter and 2 extinction coefficients (Althausen 2000). One of the earlier studies on this topic was done by Müller and Quenzel (1985) who used a randomized minimization search technique to invert synthetic measurements of backscatter and extinction at wavelengths of 0.347, 0.530, 0.694, and 1.06μm. This iterative technique was formulated by Heinzenberg et al. (1981) and is equivalently performed as follows. One begins with an initial solution $\mathbf{x}_0 = \mathbf{0}$. For each iteration, a random component, *r*, of $\mathbf{x}_n$ is chosen and one considers solutions of the form

$$\mathbf{x}_{n+1} = a\mathbf{e}_r + b\mathbf{x}_n, \tag{3.5}$$

where $\mathbf{e}_r$ is a vector whose elements are zeros except for the r^{th} element, which is one. The reproduction of measurements from the tentative solution is then

$$\mathbf{y}_{n+1} = \mathbf{A}\left(a\mathbf{e}_r + b\mathbf{x}_n\right), \tag{3.6}$$

and a and b are chosen so that

$$D_{\text{opt}} = \sum_i \left(\frac{y_{n+1,i} - y_i}{y_i} \right)^2 \tag{3.7}$$

is minimal. The solution $\mathbf{x}_{n+1}$ is accepted if none of its elements are negative. Otherwise it is rejected and another component is randomly chosen. This is repeated until either D_{opt} falls below some predefined threshold or a maximum number of iterations is reached. This method is a rather elegant approach to the problem since: 1) both over- and under-determinacy are dealt with, and 2) the positivity constraint is always satisfied. Over-determinacy is dealt with, since an optimal solution is obtained. Under-determinacy is dealt with, since the output of this method is non-deterministic. Hence the algorithm may be applied several times, and the average of solutions can be taken to give a single solution.

The results of a search on the ISI Web of Knowledge showed that there had been few publications on lidar PSD retrieval since the Müller and Quenzel (1985) study until a dramatic increase in the early 1990s. Typical approaches taken for size distribution retrieval from multi-wavelength lidar measurements during this period consisted of finding the parameters of log-normal particle size distribution model parameters that

optimally match the estimates of backscatter (e.g. Larsen et al. 1994), or of backscatter and extinction (e.g. Stein et al. 1994, and Del Guasta et al. 1994).

The year 1999 marked the beginning of a research thrust as seen in the publications of Müller et al. (1999a, 1999b, 2001), Böckmann (2001), and Veselovskii et al. (2002, 2004), who all used linear particle size distribution models with B-spline basis functions of one form or another, and included the complex index of refraction as an unknown parameter. A regularized solution (to be explained later in this section) was utilized in most of this body of research, with the one exception being the work of Böckmann (2001) who used a truncated singular value (or singular vector) decomposition (SVD) solution. The truncated SVD solution seems to be a poor choice of approach to the problem and it is necessary to give an explanation of SVD to understand why.

The singular value decomposition has been described by Hansen (1998), Rodgers (2000), and Kaipio and Somersalo (2004) whose formulations will be summarized. The SVD concisely describes the action of a matrix on a vector by expressing the matrix in the form given by Rodgers (2000) as

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T, \tag{3.8}$$

where for **A** being an $m \times n$ matrix with rank *p*, the salient matrix properties are: 1) **U** is an $m \times p$ matrix with its columns being a set of orthonormal vectors which span the range of **A**, 2) **V** is an $n \times p$ matrix with its columns also being a set of orthonormal vectors which span the row space of **A**, and 3) $\mathbf{\Lambda}$ is a $p \times p$ diagonal matrix, $\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p)$, with ordered positive elements so that $\lambda_n \geq \lambda_{n+1} > 0$. The diagonal components of $\mathbf{\Lambda}$ are known as the singular values. The specific form the

decomposition is not universal. Hansen (1998) specifies $\mathbf{\Lambda}$ and $\mathbf{V}$ to be $n \times n$ with the last $n - p$ values on the diagonal of $\mathbf{\Lambda}$ being zero, and the $n - p$ rightmost vectors in $\mathbf{V}$ spanning the null-space of $\mathbf{A}$. Another form specified by Kaipio and Somersalo (2004) has $\mathbf{U}$ as $m \times m$, $\mathbf{V}$ as $n \times n$, and $\mathbf{\Lambda}$ being $m \times n$ consisting of a diagonal matrix concatenated with a zero matrix. The form as given by Rodgers (2000) will primarily be used since it will allow some equations to be written more clearly. From Eq. 3.8 it can be seen that

$$\mathbf{Ax} = \sum_{i=1}^{p} \mathbf{u}_i \lambda_i (\mathbf{v}_i \bullet \mathbf{x}) \quad , \tag{3.9}$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the columns of $\mathbf{U}$ and $\mathbf{V}$. The multiplication $\mathbf{V}^T\mathbf{x}$ decomposes the component of $\mathbf{x}$ orthogonal to the null space of $\mathbf{A}$ into its representation in terms of the columns of $\mathbf{V}$. When $\mathbf{y} = \mathbf{Ax}$, then $\mathbf{\Lambda V}^T\mathbf{x}$ is the representation of $\mathbf{y}$ in terms of the columns of $\mathbf{U}$.

The singular value decomposition offers a way of forcing a solution to an ill-posed linear inverse problem by forming the pseudo-inverse, which is expressed for $\mathbf{A}$ as

$$\mathbf{A}^+ = \mathbf{V\Lambda}^{-1}\mathbf{U}^T \quad , \tag{3.10}$$

with $\mathbf{U}$, $\mathbf{\Lambda}$, and $\mathbf{V}$ being the SVD of $\mathbf{A}$.

A solution,

$$\mathbf{x} = \mathbf{A}^+\mathbf{y} = \sum_{i=1}^{p} \frac{\mathbf{v}_i (\mathbf{u}_i \bullet \mathbf{y})}{\lambda_i} \quad , \tag{3.11}$$

deals with over-determinacy in the sense that the product $\mathbf{U}^T\mathbf{y}$ has no contribution from the component of **y** orthogonal to the range of **A**. It is easy to show that

$$\mathbf{A}\mathbf{A}^+ = \mathbf{U}\mathbf{U}^T, \tag{3.12}$$

thus $\mathbf{x} = \mathbf{A}^+\mathbf{y}$ is an exact solution to

$$\mathbf{U}\mathbf{U}^T\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{3.13}$$

$\mathbf{U}\mathbf{U}^T\mathbf{y}$ is the projection of **y** onto the range of **A,** therefore **x** minimizes the Euclidean distance, $|\mathbf{y} - \mathbf{A}\mathbf{x}|$. Under-determinacy is dealt with by simply choosing the solution that is constrained to the row-space of **A**, hence $\mathbf{x} = \mathbf{A}^+\mathbf{y}$ is also the solution to Eq. 3.13 that minimizes $|\mathbf{x}|$. If the matrix **A** has full rank, meaning $p = \min(m,n)$, then either $\mathbf{U}\mathbf{U}^T$ or $\mathbf{V}\mathbf{V}^T$ is the identity matrix, and the pseudo-inverse can then be expressed as either

$$\mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T \tag{3.14a}$$

for over-determined systems, or

$$\mathbf{A}^+ = \mathbf{A}^T\left(\mathbf{A}\mathbf{A}^T\right)^{-1} \tag{3.14b}$$

for under-determined systems, thus avoiding the need for actually performing the SVD. It is possible that the problem is simultaneously over-determined and under-determined if both the rows and columns are not respectively linearly independent, in which case, $p < \min(m,n)$ and **A** is said to be rank deficient. However, common sense among those

familiar with machine computations indicates that this would rarely be the case exactly (an exception would be a situation where there were multiple measurements of the same coefficient). The fact that there is almost always an inaccuracy in the machine representation of a matrix due to truncation errors requires the concept of *numerical ε-rank*. Hansen (1998) gives a definition of it as

$$p_\varepsilon(\mathbf{A},\varepsilon) \equiv \min_{\|\mathbf{E}\|_2 \le \varepsilon} \operatorname{rank}(\mathbf{A}+\mathbf{E}), \tag{3.15}$$

where $\|\mathbf{E}\|_2$ represents the matrix 2-norm of **E** which is defined as

$$\|\mathbf{E}\|_2 \equiv \max_{|\mathbf{x}|=1} |\mathbf{E}\mathbf{x}|. \tag{3.16}$$

In words, if one considers all matrixes that are close to **A** within a tolerance specified by ε, then the numerical ε-rank is the lowest rank among all of those matrixes. The numerical ε-rank is related to the singular value decomposition by the property that

$$\lambda_{p_\varepsilon} > \varepsilon \ge \lambda_{p_\varepsilon+1}, \tag{3.17}$$

hence the meaningfulness of the singular values that are less than the precision to which a matrix is known is questionable. This is a serious problem since, as can be seen from Eq. 3.11, it is the components of **y** associated with the smallest singular values that have the most influence on the magnitude of the solution vector. The truncated SVD (TSVD) solution deals with this problem by simply omitting the contributions to the solution from the singular values below ε in Eq. 3.11.

In relation to the lidar PSD retrieval problem, one thing that appears to be an issue is that solutions are confined to the row-space of the kernel matrix, **A**. The low number of types of optical coefficients available means that the variety of attainable solutions may be overly constrained. These constraints are based on the physics governing the measurement of the coefficients without consideration of more relevant general aerosol qualities. The fact that even further constraints are imposed in the TSVD solution only magnifies this unsuitability. The use of singular value truncation hardly seems warranted in the context of its intended use, as can be seen in Figure 3.1 from histograms of condition numbers (defined as the ratio of highest to lowest singular values) taken on an ensemble of kernel matrixes. These matrixes were computed from Eq. 3.4 with triangular basises among different support intervals and different complex indexes of refraction for a variety of coefficient configurations similar to the ensemble used by Müller et al. (1999a), with the exception that no imaginary parts of the index of refraction were above 0.1. The ensemble corresponding to Figure 3.1 covers the values for:

particle radius support lower limits of 10, 50, 100, 150, and 200nm;

particle radius support upper limits of 1, 2, 3, 4, 6, 8, and 10μm;

index of refraction real parts of 1.33 and 1.35-1.8 in steps of 0.05; and

index of refraction imaginary parts of [0 2.5 7.5 20 40 60 80, and 100]x10^{-3}.

Although the relationship of the 2-norm defined in Eq. 3.16 to a norm that is suitable for describing matrix precision is a complicated matter, these histograms are at least useful for classifying clear-cut cases based on order of magnitude. Double floating point numeric representation has a fractional precision of less than 10^{-15}, thus in referring to Figure 3.1, it seems that the only kernel matrixes that could conceivably be ε-rank

deficient is a small subset of the matrixes corresponding to the 6-backscatter/2-extinction measurement configuration. This subset of 597 ensemble members (defined as having a condition number greater than 10^{11}) corresponds almost exclusively to ensemble members where the particle radius support lower limit is 10nm; only 5 members in the subset have other values. This last result is not surprising since for particles much smaller than the wavelength of incident light, the spectral properties of backscatter are identical (specifically a dependence $\propto \lambda^{-4}$) other than differing multiplying constants for all small particle sizes (Bohren and Huffman 1983).

Even though it has been shown that a great majority of the kernel matrixes are not ε–rank deficient, Figure 3.1 does demonstrate that error magnification will be a problem in addition to the problem of having solutions constrained to the row-space. A more reasonable approach to deal with these issues is to use Twomey-Tikhonov regularization. This method finds a solution that minimizes a penalty function that both formulates how well the coefficients produced from the solution match the measurements, and penalizes solutions that are considered less likely without consideration of measurements (Rodgers 2000). The penalty function in Twomey-Tikhonov regularization is expressed as

$$\gamma(\mathbf{Ax}-\mathbf{y})^T(\mathbf{Ax}-\mathbf{y})+\mathbf{x}^T\mathbf{Hx}, \tag{3.18}$$

with $\mathbf{H}$ being a symmetric matrix, and the minimizing solution is

$$\mathbf{x}_{\text{reg}} = \mathbf{A}^{+\text{reg}}\mathbf{y} = (\mathbf{A}^T\mathbf{A}+\gamma^{-1}\mathbf{H})^{-1}\mathbf{A}^T\mathbf{y}. \tag{3.19}$$

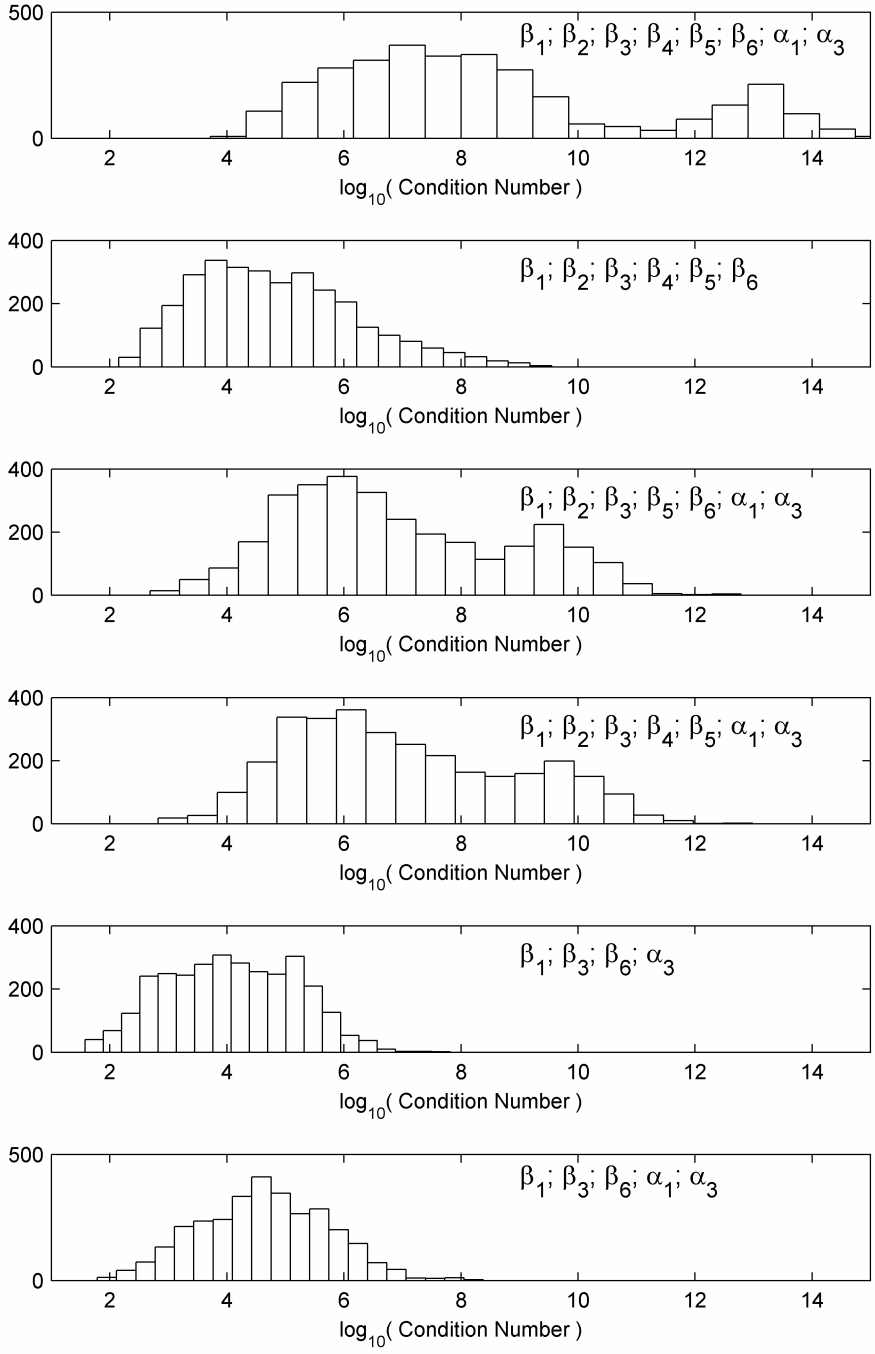

Figure 3.1. Condition number histograms. Indexes in the coefficient subscripts correspond to wavelengths as: 1 - 355; 2 - 400; 3 - 532; 4 - 710; 5 - 800; 6 - 1064 (all in nm).

When **H** is the identity matrix the second penalty is simply proportional to $|\mathbf{x}|^2$ and the solution is similar to the TSVD solution in the sense that the TSDV solution causes a reduction in solution magnitude. An analysis using SVD can be done to gain insight into this similarity (Rodgers 2000). This analysis requires using the form of the SVD as given by Hansen (1998) or Kaipio and Somersalo (2004). First it is noted that the lidar PSD retrieval problem is under-determined and the matrix **U** is the same and unitary in any SVD form. The full form of $\mathbf{\Lambda}$ and **V** matrixes are written as $\mathbf{\Lambda}_f$ and $\mathbf{V}_f$, where $\mathbf{\Lambda}_f = [\mathbf{\Lambda}\ \mathbf{0}]$ and $\mathbf{V}_f = [\mathbf{V}\ \mathbf{V}_\varnothing]$, where $\mathbf{V}_\varnothing$ represents an arbitrary orthonormal basis for the null-space of **A**. To reiterate, Eq. 3.8 is valid for either form of the SVD and in this analysis **H** is taken to be the identity matrix. The regularized solution in Eq. 3.19 is then

$$\mathbf{x}_{\text{reg}} = (\mathbf{V}_f \mathbf{\Lambda}_f^T \mathbf{U}^T \mathbf{U} \mathbf{\Lambda}_f \mathbf{V}_f^T + \gamma^{-1} \mathbf{I}_n)^{-1} \mathbf{V} \mathbf{\Lambda} \mathbf{U}^T \mathbf{y} \quad . \tag{3.20}$$

Noting that since the full form SVD results in the equality, $\mathbf{V}_f \mathbf{V}_f^T = \mathbf{I}_n$, Eq. 3.20 can be written as

$$\begin{aligned} \mathbf{x}_{\text{reg}} &= \mathbf{V}_f \left\{ \begin{bmatrix} \mathbf{\Lambda}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \gamma^{-1} \mathbf{I}_n \right\}^{-1} \mathbf{V}_f^T \mathbf{V} \mathbf{\Lambda}\, \mathbf{U}^T \mathbf{y} \\ &= \mathbf{V}_f \left\{ \begin{bmatrix} \mathbf{\Lambda}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \gamma^{-1} \mathbf{I}_n \right\}^{-1} \begin{bmatrix} \mathbf{\Lambda} \\ \mathbf{0} \end{bmatrix} \mathbf{U}^T \mathbf{y} \end{aligned} \quad . \tag{3.21}$$

Contemplation of Eq. 3.21 indicates that the regularized solution is analogous to the pseudo-inverse solution in Eq. 3.11 with λ_i' replacing λ_i by

$$\lambda_i' = \lambda_i + \lambda_0 \frac{\lambda_0}{\lambda_i}, \tag{3.22a}$$

where

$$\lambda_0 = \gamma^{-1/2}. \tag{3.22b}$$

Hence a lower limit of $2\lambda_0$ is placed on the effective singular values, and the regularized solution with $\mathbf{H} = \mathbf{I}_n$ is similar to the truncated SVD solution yet can be done with simpler matrix operations.

As is more apparent in Eq. 3.21, the regularized solution in this form still has the problem that the solution is constrained to the row-space of **A**. Müller et al. (1999a), and Veselovskii et al. (2002 and 2004) had applied the regularized solution to the lidar PSD retrieval problem with an **H** that favors smoother solutions in a manner analogous to penalizing a function for the magnitude of its integrated squared second derivative, specifically,

$$\mathbf{H} = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ -2 & 5 & -4 & 1 & 0 & 0 & 0 & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 & -4 & 6 & -4 & 1 \\ 0 & 0 & 0 & 0 & 1 & -4 & 5 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix}. \tag{3.23}$$

A computation of $\mathbf{V}_{\varnothing}^{T}\mathbf{A}^{+\mathrm{reg}}$ with $\mathbf{A}^{+\mathrm{reg}}$ derived using this **H** matrix was performed on one of the matrixes from the ensemble that was used to examine condition numbers

described above; it clearly showed that such smoothing regularized solutions can have components in the null space.

Both Müller et al. (1999a) and Veselovskii et al. (2002) used a method that repeatedly computed regularized solutions using the **H** matrix in Eq. 3.23 over multiple basises with different supports. The solution ensemble would then be refined by rejecting solutions using complicated criteria; the method of Müller et al. (1999a) rejected solutions that were either strictly positive or had negative portions that did not extend to either of the upper or lower limits of the support of the basis. The negative weights of the remaining ensemble members were set to zero, and these members were averaged to get a single output solution. These techniques were so complicated that essentially black-boxes were constructed. This makes assessing uncertainty difficult. One would only be able to perform a forward Monte Carlo method which is not appropriate for inferring aerosol size distributions from measurements, as will be shown in sections 5 and 6. When forward Monte Carlo analyses were done, only a handful of trials were taken (Veselovskii 2002).

In this study the problem of assessing uncertainty will be approached by representing the PSD in a simpler log-normal model, thereby making more advanced uncertainty computations easier to perform. Two approaches to assessing uncertainty will be presented: an *ad hoc* graphical approach applicable to models with a single index of refraction and a more rigorous Bayesian approach which will additionally be applied to assess uncertainty of index of refraction.

Prior research in retrieving aerosol distributions from lidar measurements does suggest that there is sufficient information in the measurements to retrieve index of

refraction. Böckmann (2001) was able to get solutions that included a limited range of indexes of refraction with the full Nd:YAG Raman lidar coefficient set (backscatter at 1064, 532, and 355nm and extinction at 532, and 355nm) with additional backscatter coefficients at 400, 710, and 800nm that can be retrieved with the 6-wavelength lidar at the Institute for Tropospheric Research in Leipzig, Germany (Althausen 2000). Müller et al. (2001) and Veselovskii et al. (2002) were able to do the same without the additional backscatter coefficients. Many of the qualitative results concerning the viability of different sensor channel configurations based on numerical tests will be confirmed with the Bayesian analysis in section 6.

4. Graphical Uncertainty Assessment

4.1 Introduction

A graphical approach can be used to estimate the PSD parameter values as well as the uncertainty in the retrieval. Rather than applying numerical algorithms, this approach utilizes the search of pre-generated look up tables (LUTs) that map optical coefficient ratios to a discretized PSD parameter space. The solution space is graphically visualized as the intersection of contours in the parameter space as was done by Post (1996). The framework for this paper uses this graphical methodology that is based on a lognormal monomodal PSD and an assumed index of refraction. Post (1996) tacitly assumed that the error bounds would directly follow from the optical coefficient uncertainties, but no procedure was outlined. As noted earlier, the optical coefficients used are not the extinction and backscatter coefficients but the ratio of these (independent) quantities. Such optical coefficient ratios are not Gaussian distributed nor independent; this can have a significant effect on the total cumulative distribution function and therefore, the uncertainty bounds for a given confidence level need to be redefined. This extension of earlier work is necessary to ensure that comparisons made between different optical coefficient sets are based on the same confidence level. Only then can the graphical technique be meaningfully applied to investigate improvements in the retrieval of PSD parameters with increased number of optical coefficients. Unfortunately direct comparisons of the results found by Post (1996) and Dellago and Horvath (1993) cannot be performed since no working definition of the confidence bounds was given. In this work, the potential benefits from the addition of nitrogen Raman extinction channels to a typical three wavelength Nd:YAG laser based lidar system are examined. The results

obtained are consistent with studies by Böckmann (2001) using regularization techniques reviewed earlier, but are obtained in a more direct manner.

In addition, the graphical method was previously applied only to retrieve the mean radius and GSD parameters with little attention to obtaining estimates of the total number desnity (Post 1996, Delago, and Horvath 1993). This section outlines a procedure for retrieval of both the number density and its uncertainty following the determination of PSD parameters. Finally, while the graphical technique is based on a single mode PSD assumption, it is also applied to underlying bimodal distributions to see if quick insight can be gained into the nature of the distribution.

In section 4.2, the basics of the graphical inversion are reviewed and methodology to determining confidence bounds is outlined. Further, a procedure to retrieve the number density and its uncertainty is introduced. In section 4.3, the method to study the retrieval capabilities of a typical 3 wavelength backscatter lidar using a Nd:YAG laser transmitter (355, 532, and 1064nm) for an a-priori known index of refraction is applied, and the results are compared to those obtained from adding additional extinction coefficients from Raman channels. The possibility of using the graphical method to distinguish monomodal from bimodal distributions is also explored. Section 4.4 provides a summary of the results.

4.2 Graphical Retrieval of PSD Parameters

Rather than applying numerical search algorithms that may have difficulties in converging, the graphical method does not need to find local minima. Therefore, the technique can be used to both retrieve PSD parameters as well as provide an estimate of retrieval errors without the need to perform numerical minimization. In this sub-section,

the general graphical approach to determining parameter uncertainty estimates is described.

To examine how a set of multi-wavelength optical coefficients can be used to determine underlying distribution parameters, and some of the difficulties involved, the graphical approach as presented by Post (1996) and applied to three-wavelength lidar is illustrated.

An optical coefficient set comprised of extinction ratios can be given as $\left\{R_1 = \frac{\alpha_{532}}{\alpha_{355}}, R_2 = \frac{\alpha_{1064}}{\alpha_{355}}, R_3 = \frac{\alpha_{1064}}{\alpha_{532}}\right\}$. Similar ratio sets can be considered for the backscatter β or lidar ratio ($S = \alpha/\beta$) parameters. The main steps in the graphical approach are as follows:

1. For a given point in the parameter space $\mathbf{p}(\bar{r}_0, \sigma_0)$ where the zero subscript denotes the parameters used in generating the optical coefficients, a forward calculation can generate the corresponding calculated optical coefficient ratios $\mathrm{R_k}(\mathbf{p})$. The parameter space is discretized into cells $\mathbf{p}(\bar{r}_i, \sigma_j) \equiv \mathbf{p}_{ij}$, and the corresponding values of the optical coefficient (ratios) are evaluated and stored in 3 matrices, $\mathbf{R}_k^{i,j}$ $k = 1:3$ (for a three wavelength lidar), which will then serve as look-up tables (LUTs).
2. Given ratios of initially assumed optical coefficients (for the purpose of this paper, or calculated coefficients from lidar data), and the associated ratio uncertainty bounds (ΔR_k), the pre-constructed database (LUTs) is searched to determine the cells (i, j) which satisfy,

$$R_k - \Delta R_k \le \mathbf{R}_k^{i,j} \le R_k + \Delta R_k \quad k = 1:3 \,, \tag{4.1}$$

This can be thought of as finding the intersection of the 6 curves in the parameter space **p**.

In this way, a visual representation of the solution space can be constructed and different optical measurement schemes compared. This procedure can be generalized to *n* distinct optical coefficients by forming the set of resulting $n(n-1)/2$ ratios,

$$R_{k\{i,j\}} = \frac{\chi_j}{\chi_i} \quad j \ge i \quad i = 1:n \,, \tag{4.2}$$

In this work, the $n(n-1)/2$ pre-calculated ratio matrices (LUTs) $\mathbf{R}_k^{i,j}$ were generated on the following grids: $\bar{r}$ was divided into small size (from 0.002 to 0.2 μm in 0.002 steps) and large size (from 0.2 to 0.9 μm in 0.01 steps) ranges; and σ ranged from 1.01 to 2.4 in steps of 0.01. Unless otherwise noted, the complex index of refraction is limited to $m = 1.5 + 0.0j$, representing non-absorbing tropospheric aerosols.

To illustrate the general idea, Figure 4.1 shows plots of the contours for 2% uncertainty for different types of optical coefficient ratios (extinction, backscatter and $S = \frac{\alpha}{\beta}$ ratio) generated from a log-normal PSD with $\bar{r}_0$ = 0.65um, σ_0 = 1.4.

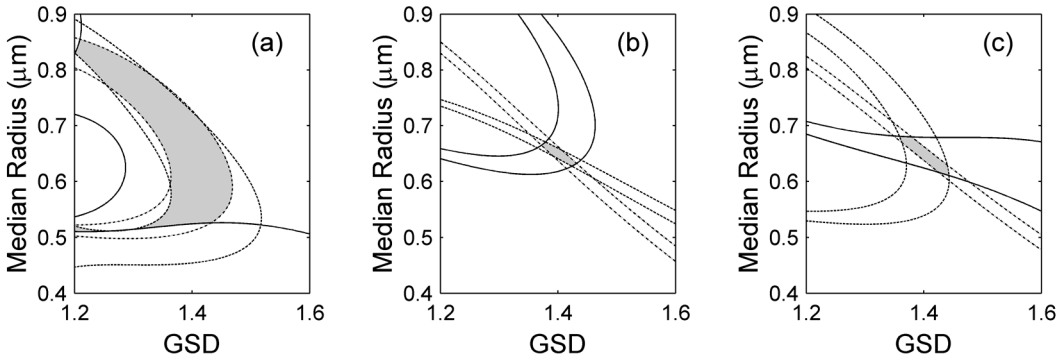

Figure 4.1. Example of graphical determination of PSD parameter solution space from 3 optical coefficients generated from an underlying log-normal distribution with $\bar{r}_0$ = 0.65um, σ_0= 1.4 and with the addition of 2% uncertainty for a) extinction coefficient ratios, b) backscatter coefficient ratios and c) S Ratios.

In this figure, the significant improvement in PSD retrieval using backscatter coefficients over extinction coefficients can be seen, which is well known in scattering theory and pointed out by Post (1996). Finally, while it may seem that the third element in the ratio set is redundant and supplies no independent information, when assessing the complete error space, it can make a significant difference.

4.2.1 Estimates of Confidence Bounds

In order to calculate the correct confidence bounds, it is necessary to first develop the statistics for the ratio of random variables (RVs), and then combine the variances of the resulting set of ratios. Performing tedious but straightforward calculations based on the PDF of bivariate functions (Papoulis 1991), the probability density function for the fractional optical coefficient ratio deviation $\delta = \left(R - \overline{R}\ \right)\!/\overline{R}$ is given by,

$$P(\delta;\Delta_n,\Delta_d) = \frac{1}{2\pi\,\Delta_{\mathrm{n}}\,\Delta_{\mathrm{d}}}\exp\!\left(-D/2\right) \times\left[\frac{2}{A}\exp\!\left(-\frac{A}{2}(B-1)^2\right)+(B-1)\sqrt{\frac{2\pi}{A}}\,\mathrm{erf}\!\left(\sqrt{\frac{A}{2}}(B-1)\right)\right], \quad (4.3a)$$

where,

$$A = \frac{1}{\Delta_{\mathrm{d}}{}^2}+\frac{(1+\delta)^2}{\Delta_{\mathrm{n}}{}^2}\ , \quad (4.3b)$$

$$B = A^{-1}\frac{\delta+\delta^2}{\Delta_{\mathrm{n}}{}^2}\ , \quad (4.3c)$$

and,

$$D = \frac{\delta^2}{\Delta_n{}^2} - AB^2 \quad , \tag{4.3d}$$

where $\Delta_d = \mathrm{std}\left(\frac{\chi_d - \bar{\chi}_d}{\bar{\chi}_d}\right)$ is the fractional standard deviation of the optical coefficient in the denominator and $\Delta_n = \mathrm{std}\left(\frac{\chi_n - \bar{\chi}_n}{\bar{\chi}_n}\right)$ of the optical coefficient ratio in the numerator. Eq. 4.3a is a direct consequence of a generalized formula, Eq. 6.17, derived in section 6.1. To verify the PDF expression in Eq. 4.3a, the above expression is plotted in Figure 4.2 together with a numerically generated frequency plot of the ratio statistics for the case of $\Delta_d = \Delta_n = 0.1$. The PDF shows significant asymmetry about zero, illustrating that Gaussian formulations of confidence bounds is incorrect.

If the ratios are uncorrelated, once the single ratio PDF P is given, the total PDF for *N* ratios scales as $P_N \approx (P)^N$. Similarly, if we define the 1D cumulative distribution function as

$$C(\delta) = \int_{-\delta}^{\delta} P(\delta')d\delta' \quad , \tag{4.4a}$$

and the *N* dimensional cumulative distribution function as

$$C_N(\delta) = \int_{-\delta}^{\delta} \cdots \int_{-\delta}^{\delta} P_N(\delta_1 \cdots \delta_N) d\delta_1 \cdots d\delta_N \quad , \tag{4.4b}$$

then $C_N^{uncorr}(\delta) = C(\delta)^N$, where $C_N(\delta)$ is the probability that all ratios lie between $\overline{R}_k(1-\delta) \le R_k \le \overline{R}_k(1+\delta)$.

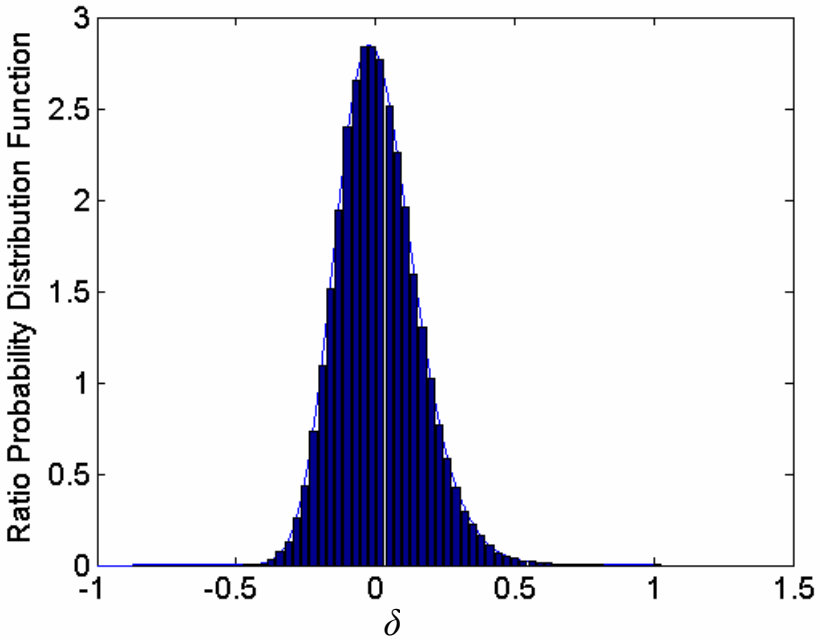

Figure 4.2. Verification of the probability density function formula in Eq. 4.3a with a numerically generated histogram of the ratio statistics for the case of $\Delta_d = \Delta_n = 0.1$.

When each ratio set within the *N* dimensional domain defined by Eq. 4.4b will result in a parameter retrieval inside the contour curves defined in Eq. 4.1 (i.e. when the contours bound a closed area in solution space), then the percentage of ratios inside the *N* dimensional domain will be equal to the percentage of parameters retrievals inside the bounding contour curves. This allows the uncorrelated confidence to be defined bound as the numerical (graphical) solution to

$$C_N^{uncorr}(\delta) = C_{level}, \tag{4.5}$$

where C_{level} is the desired confidence level.

However, because correlations in the optical coefficient ratios are present, $C_N^{corr}(\delta) \neq C(\delta)^N$. To illustrate the effect of correlations between the optical coefficient ratios, the cumulative distribution function for a correlated set of optical coefficient ratios obtained from numerical simulations are illustrated in Figure 4.3. In this figure, the cumulative PDF $C_N^{corr}(\delta)$ of the set of correlated optical coefficient ratios valid for 3, 4, and 5 optical channels are compared to the cumulative PDF generated from the same number of uncorrelated optical coefficient ratios $C_N^{uncorr}(\delta)$. Both curves were calculated using uncorrelated random number generators that directly generate appropriate uncorrelated optical coefficient ensembles with 10% uncertainty. From these, the set of correlated or uncorrelated optical coefficient ratio ensembles were formed and the cumulative distribution function constructed. It should be pointed out that numerically calculating the cumulative PDF is far less intensive than trying to calculate the actual PDF distribution.

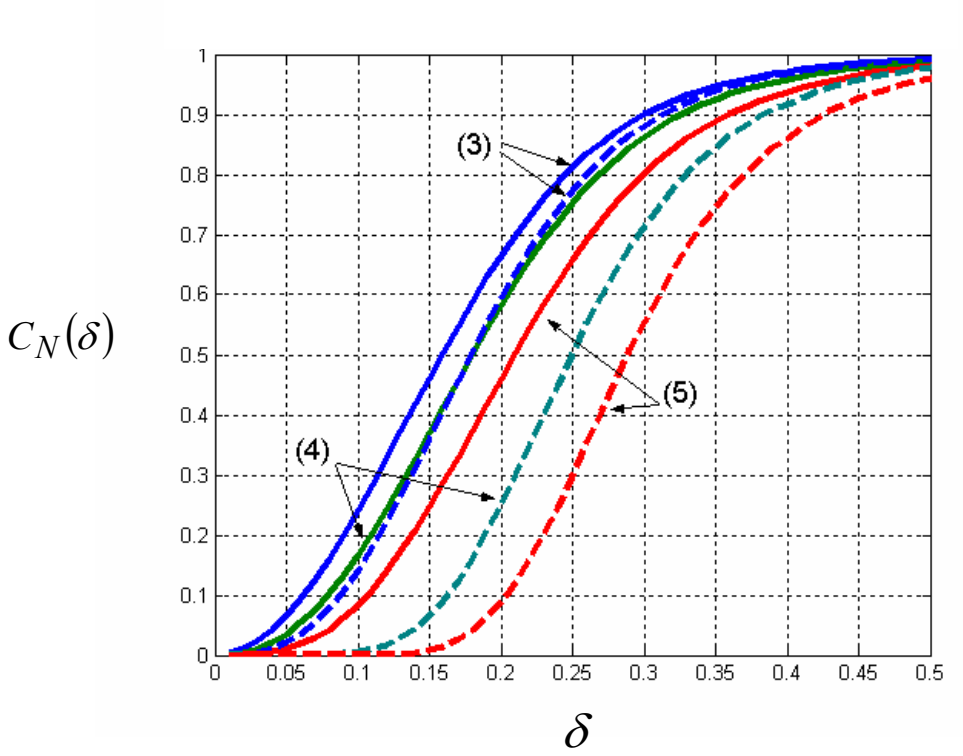

Figure 4.3. Comparison of the cumulative probability distribution function for the case of 3,4 and 5 optical coefficients showing the increased conditional probability for correlated ratios. Solid lines are calculated for correlated ratios while the dashed lines are for uncorrelated ratios.

A significant increase is observed in the difference between the correlated and uncorrelated cases as the number of measurements increases. This implies, as intuitively expected; that a larger set of correlated data has the information content of a smaller set of uncorrelated data. By assuming that all the ratios are uncorrelated, the confidence bounds are overestimated depending on the size of the optical coefficient set (3, 4, or 5).

To verify the general approach, conventional least-squares inversions on an ensemble of three channel backscatter coefficients generated from a-priori PSD parameters were performed, and with equal uncorrelated fractional variances of 10%. The inversions were then compared in Figure 4.4 to the graphical method solution space using the confidence level C_{level}=.65, but with uncorrelated bounds. The scatter points represent the inversion results and the solid lines are the confidence boundaries.

The results of the inter-comparison are shown in Table 4.1, which provides the percentage of numerical inversion solutions within the graphical solution space. As is clear from the table, the graphical estimates were fairly accurate for calculating the confidence region as seen by the agreement between the fraction of inversions within the confidence region and C_{level}=.65. However, a small but significant overestimation of the confidence region results. This is easily explained by observing that the true cumulative distribution $C_N^{corr}(\delta_{uncorr})$ evaluated with the uncorrelated bounds results in $C_N^{corr}(\delta_{uncorr})$=.71 which is in good agreement with all the table entries except case d. For this case, the solution space included the boundary of the look up tables.

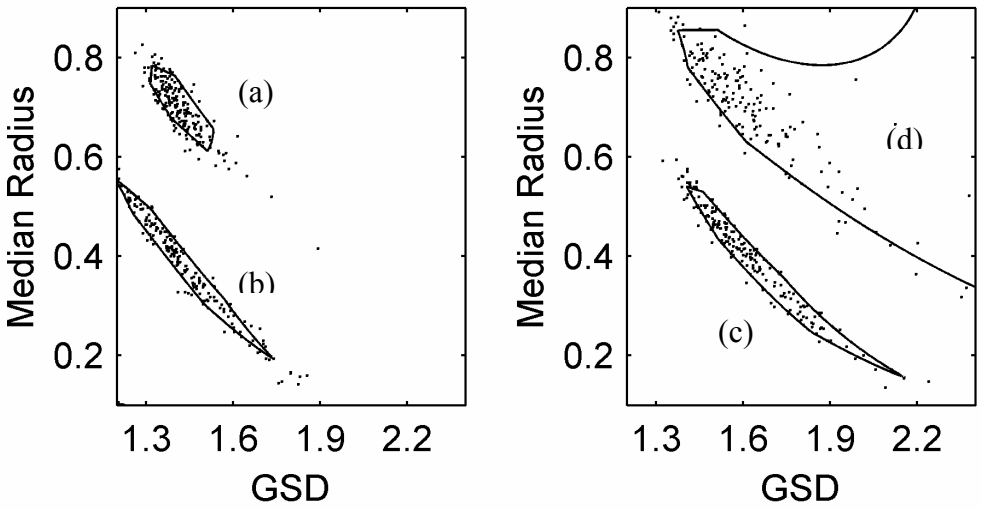

Figure 4.4. Numerical retrievals relative to graphical bounds at $C_{level} = .65$ on ensembles of three channel backscatter coefficients generated from a-priori PSD parameters, with equal uncorrelated fractional variances of 10%: a) $\bar{r}_0 = 0.7 \quad \sigma_0 = 1.4$, b) $\bar{r} = 0.4 \quad GSD = 1.4$, c) $\bar{r}_0 = 0.4 \quad \sigma_0 = 1.6$, d) $\bar{r} = 0.7 \quad GSD = 1.6$

Table 4.1: Percentage of numerical inversions within the correlated graphical confidence bounds versus theoretical value C_{level} =.65. Cases under study are those of Figure 4.4.

Distribution Parameters	% within Symmetric Confidence Bounds	Total # of Points (unclipped)
$\bar{r}_0 = 0.7 \quad \sigma_0 = 1.4$	73.65591	186
$\bar{r}_o = 0.4 \quad \sigma_0 = 1.4$	69.23077	195
$\bar{r}_0 = 0.4 \quad \sigma_0 = 1.6$	69.50000	200
$\bar{r}_0 = 0.7 \quad \sigma_0 = 1.6$	81.70732	164

Therefore, some of the numerical inversions were attracted to the boundary, which is inside the confidence region (see Figure 4.4d). This leads to an additional artificial increase of the percentage of inversions within the confidence zone.

4.2.2 Number Density Estimation

When optical coefficient ratios are used, information on the number density is not directly accessible. However, once a confidence region is determined, the number density distribution may be obtained numerically using the following procedure:

1. For each parameter cell $\mathbf{p}(\bar{r}_i, \sigma_j)$ within the confidence zone, the associated number density $\mathbf{N}_k^{i,j}$ for the k^{th} optical coefficient χ_k is calculated as the ratio $N_0 \chi_k \big/ \boldsymbol{\chi}_k^{i,j}$ where N_0 is the reference number density (=1 for this work) used to generate the optical coefficient look up tables $\boldsymbol{\chi}_k^{i,j}$. Therefore, the relative error $\mathbf{N}_k^{i,j} \big/ N_0$ ascribed to each cell in parameter space is $\chi_k \big/ \boldsymbol{\chi}_k^{i,j}$.
2. From the relative Number Density error, the average number density uncertainty can be obtained for each cell as $\mathbf{N}_{ave}^{i,j} \big/ N_0 = \frac{1}{k} \sum_k \mathbf{N}_k^{i,j} \big/ N_0$ (averaging over all optical coefficient ratios *k*).
3. A histogram is constructed from set of $\mathbf{N}_{ave}^{i,j} \big/ N_0$ values calculated from all the parameter cells that were in the solution space. As an alternative, the histogram can be constructed from $\mathbf{N}_k^{i,j} \big/ N_0$, bypassing step 2.

This method was applied to optical coefficients generated from an example PSD ($\bar{r}_0$=0.5, σ_0= 1.4) with an uncertainty of 10%, and confidence level set to 65%. The

confidence bounds are shown in Figure 4.5a and the resultant fractional number density histogram is given in Figure 4.5b. In particular, besides obtaining the general statistics, it is shown how averaging of the number density over all optical channels compares to taking each channel separately and combining all number density retrievals into one histogram. From the comparison, the averaging is demonstrated to have a positive effect on the large deviation tail, taking probability away from the large deviation events and transferring it to smaller deviating solutions. Of course, once the numerical PDF is obtained, all meaningful statistical moments can be calculated for a given confidence level.

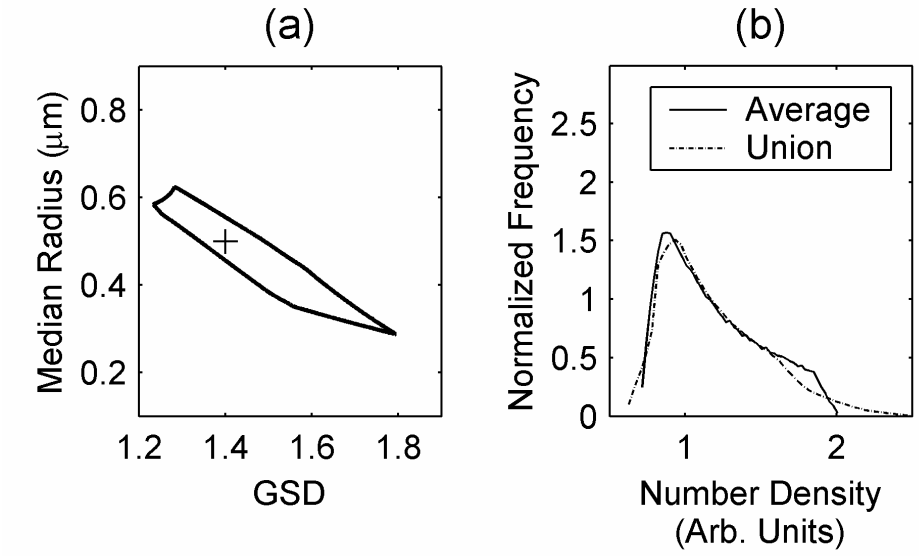

Figure 4.5. a) Confidence domain obtained using optical coefficients generated from an example PSD ($\bar{r}_0$=0.5, σ_0= 1.4 denoted by the + sign) with an uncertainty of 10%, and confidence level set to 65% b) Fractional number density retrieval obtained from the confidence domain using two different approaches, combining the results from different optical channels.

4.3 Applications of Graphical Inversion Assessment

4.3.1 Parameter Retrieval Improvement with Extinction Coefficients

A quick measure of the overall potential for a set of optical coefficients to retrieve the PSD mode parameters is the "size" of the retrieval area A defined by the confidence bound contours in parameter space Besides limiting the aerosol mode mean $\bar{r}$ and variance σ, reducing the retrieval area obviously reduces the uncertainty in the number density retrieval as well. Therefore, it is natural to study how A varies over the parameter space. Since errors are usually reported as fractional errors, it is useful to define a normalized retrieval area $\Delta A_N = \frac{A}{[\bar{r}_0 \sigma_0]}$ where $(\bar{r}_0, \sigma_0)$ are the underlying PSD parameters used to calculate the optical coefficients before the introduction of uncertainty. For completeness, we also normalize deviations of the mode mean and mode variance as $\Delta \bar{r}_N = \frac{\Delta \bar{r}}{\bar{r}_0}, \Delta \sigma_N = \frac{\Delta \sigma}{\sigma_0}$ where $\Delta \bar{r}$ and $\Delta \sigma$ are the difference between the maximum and minimum retrieval of the mode radius and variance within area A, and explore their variation over the entire log normal parameter space for the important reference case of the elastic backscatter signal from the three harmonics of the Nd: YAG transmitter at 1064nm, 532 and 355nm (i.e. three backscatter but no extinction) with an assumption of 10% uncertainty in optical coefficients. This error level is commonly used as the benchmark by the community but is still quite difficult to achieve

The Raman Lidar technique allows for additional optical data in the form of extinction coefficients at the nitrogen Raman-shifted wavelengths for generated from laser transmitter output at the 355 nm and 532 nm signal (at 387 and 607 nm respectively) which can be extrapolated to the elastic lidar wavelengths with appropriate

power law wavelength dependent extinction models. However, the usefulness of the additional extinction data is not clear and seems to depend on whether the complex refractive index is known. It has been shown in a number of studies that a single extinction coefficient is necessary for stabilization when using regularization (Böckmann 2001, Müller et al. 2001, Veselovskii et al. 2002) and that a single extinction is needed for PSD with a known complex index.

To evaluate the information content in a given set of optical coefficients, using the confidence level C_{level} =0.65, we calculate for each value of the parameters $(\bar{r}_0, \sigma_0)$, the set $\Delta\bar{r}_N, \Delta\sigma_N$ and ΔA_N **,** as the relevant performance metrics.

To illustrate the importance of proper definition of confidence bonds, the calculations are initially carried out assuming that the optical coefficient ratios are uncorrelated. Figure 4.6 shows the performance metric for the case of 3-wavelength backscatter lidar. It is immediately apparent that the parameter space can be divided into a region of good parameter retrieval and a region of poor parameter retrieval that occurs for distributions with simultaneously large mode radius and mode width. This is consistent with the general observation that large particle modes cannot be inverted unless the exact shape is known and as illustrated, even then, the uncertainties grow very large.

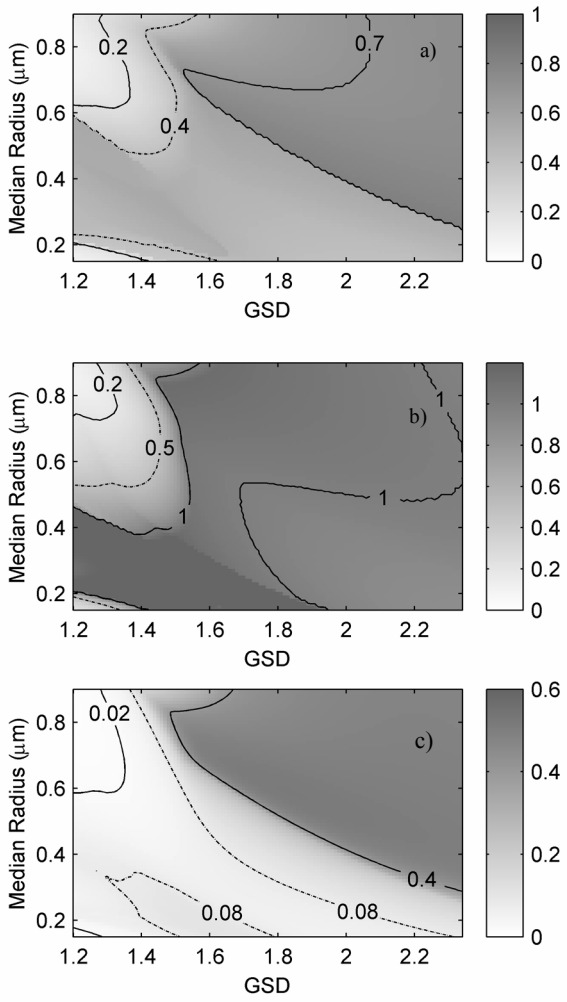

Figure 4.6. Assessment over the log-normal space of parameter retrieval uncertainty span (maximum deviation) for the case of three backscatter coefficients with 10% coefficient uncertainty and confidence level of 65% a) maximum deviation of normalized mode radius $\Delta\bar{r}_N$, b) maximum deviation of normalized mode variance $\Delta\sigma_N$, c) normalized uncertainty area ΔA_N. The gray scale bar to the right is a measure of retrieval uncertainty span (lighter regions have lower retrieval errors)

Unfortunately, a practical problem that needs to be addressed in the graphical method is that finite LUT calculation windows bias the retrieval as the confidence region interacts with the edges of the computation window. Therefore, only the portion of parameter space that is sufficiently far from the calculation window boundary has confidence regions totally contained in the window unclipped while for parameter values near the boundary, the confidence region is clipped. However, in our case, significant clipping occurs mostly for PSDs, which are already determined to be difficult to retrieve (large particle modes) so it does not greatly affect the qualitative division of the parameter space.

While the three backscatter optical coefficient set is capable of retrieving PSDs for a known index of refraction for the case of log-normal distributions for certain well defined regions of parameter space as illustrated by Figure 4.6c, a natural extension is to assess usefulness extra Raman extinction coefficients. It has already been shown with regularization methods (Veselovskii et al. 2002 and 2004) that 2 extinction parameters can stabilize a solution in the presence of unknown refractive index while when considering PSD retrieval with known refractive index, it was shown that a single extinction coefficient is sufficient for monomodal distributions (Müller et al. 2001).

To study the improvements in the retrieval with increased number of optical coefficients, and the effect of correlation in the optical coefficient ratios, we plot in Figure 4.7 the enhancement factors as defined by

$$E_{4,3} = \log_{10}\left(\frac{[\Delta A_N]_3}{[\Delta A_N]_4}\right) \text{ for } 3\beta \;\; vs \;\; 3\beta + 1\alpha \tag{4.6}$$

and

$$E_{5,4} = \log_{10}\left(\frac{[\Delta A_N]_4}{[\Delta A_N]_5}\right) \quad 3\beta + 1\alpha \quad vs \quad 3\beta + 2\alpha \,. \tag{4.7}$$

first using uncorrelated bound estimates. From these definitions, positive enhancement factors result in improved retrieval uncertainty. In particular, we make the following observations that are valid where the refractive index is assumed known a-priori.

1. There is a definite region where $E_{4,3} > 0$ indicating that an additional extinction coefficient helps reduce parameter retrieval uncertainty (Figure 4.7a)
2. However, since $E_{5,4} < 0$ for nearly all parameter values, we come to an anomalous result that suggests that more coefficients may lead to an increased uncertainty in parameter retrieval (Figure 4.7b)

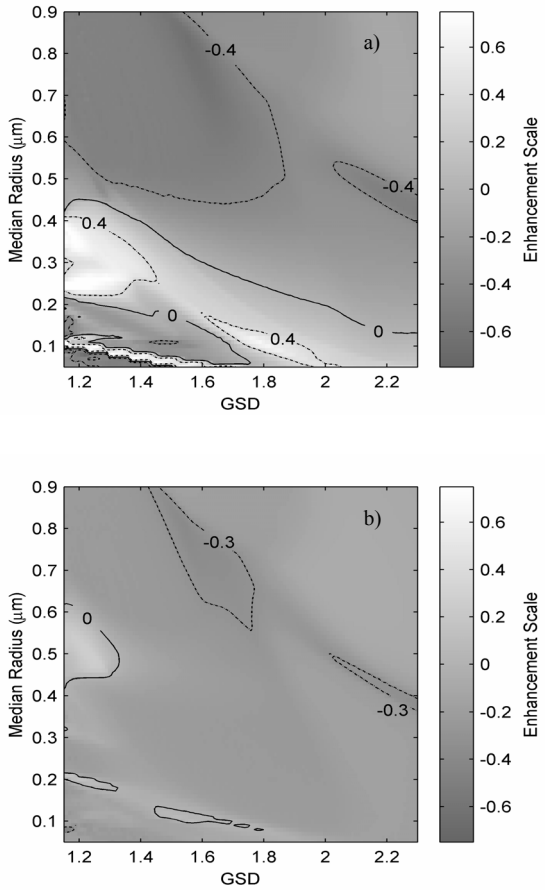

Figure 4.7. Enhancement of parameter retrieval as extinction coefficients are added using uncorrelated bound definition. a) $E_{4,3}$ [$(3\beta + 1\alpha)$ vs (3β)] b) $E_{5,4}$ [$(3\beta + 2\alpha)$ vs. $(3\beta + 1\alpha)$]. Lighter regions have larger enhancement factors.

This observation is counter-intuitive since adding extra optical coefficient ratios should reduce the retrieval uncertainty. This comes about because, as noted earlier, the above calculation was done assuming uncorrelated optical coefficient ratios. In this approximation, the magnitude of the uncertainty bounds applied to each ratio is larger than it should be if statistical dependence is taken into account. To adjust for this affect, we can simply apply a correction factor $f(C) = \frac{\Delta R_{corr}(C)}{\Delta R_{uncorr}(C)}$ to convert the uncorrelated confidence bounds for a given confidence level to the correct correlated bounds utilizing the results in Figure 4.3. Examination of Figure 4.3 shows that very little adjustment is needed for the case of three optical coefficient ratios (i.e. f ~ 1). However, a significant adjustment of confidence bounds is needed when considering more than three optical coefficients. This correction has been verified by calculating the percentage of numerical inversions contained in the correlated confidence bounds for a variety of confidence levels and comparing to the solid curves in Figure 4.3.

Applying this correction to the calculation of enhancement factors $E_{4,3}$ and $E_{5,4}$ to account for correlations leads to a much more reasonable quantitative comparison as illustrated in Figure 4.8. Note that nearly all the anomalous negative enhancement regions have been eliminated. The final observations are:

1. There is a definite region where $E_{4,3} > 0$ where an additional extinction coefficient helps reduce parameter retrieval uncertainty significantly (Figure 4.8a)

2. Since $E_{5,4} \approx 0$ for most parameter values (Figure 4.8b), we see that including a second extinction parameter does not significantly help in reducing the parameter retrieval for the most part, except when considering a fine urban mode (e.g. $\bar{r} \approx .15um \quad 1.4 \le \sigma \le 2.0$), where there is a moderate enhancement going from one to two extinction parameters as seen in the expanded view of Figure 4.9.

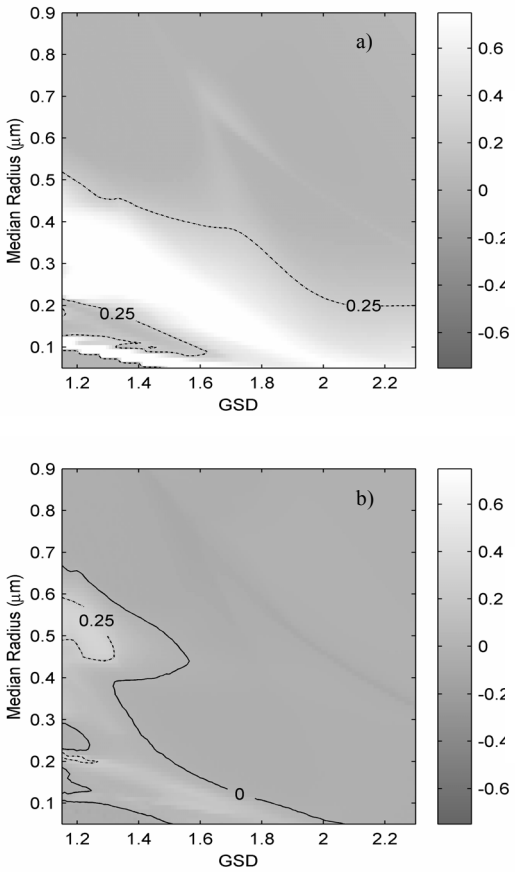

Figure 4.8. Enhancement of parameter retrieval as extinction coefficients are added using correlated bound definition. a) $E_{4,3}$ $[(3\beta + 1\alpha)$ vs $(3\beta)]$ b) $E_{5,4}$ $[(3\beta + 2\alpha)$ vs $(3\beta + 1\alpha)]$. Lighter regions have larger enhancement factors.

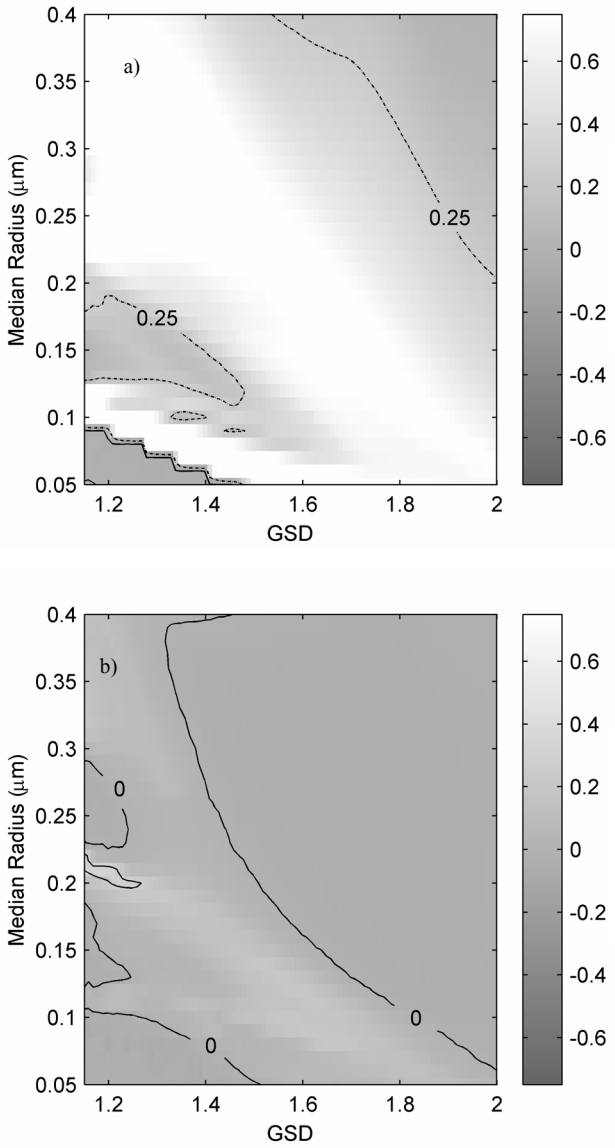

Figure 4.9. Zoom view of parameter retrieval enhancement in the fine urban aerosol mode regime. Calculations performed using correlated bounds. a) $E_{4,3}$ [$(3\beta+1\alpha)$ vs (3β)] b) $E_{5,4}$ [$(3\beta+2\alpha)$ vs $(3\beta+1\alpha)$]. Lighter regions have larger enhancement factors.

While the calculations above were carried out for 10% uncertainty in the optical coefficients, the efficacy of the graphical technique for various uncertainty levels can be easily examined. This is illustrated in Figure 4.10 where the values of the normalized maximum and minimum PSD parameters within the solution space ($\bar{r}_N^{\max,\min} \equiv \frac{\bar{r}_{\max,\min}}{\bar{r}_0}$, $\sigma_N^{\max,\min} \equiv \frac{\sigma_{\max,\min}}{\sigma_0}$) are plotted as a function of the noise uncertainty level. All results were obtained for $m_r = 1.5 \quad m_i = -0.02$. Panel a) representing larger particle modes ($\bar{r}_0 = 0.8 \quad \sigma_0 = 1.8$) shows clearly that the retrieval error rapidly increase for relatively small uncertainty levels and no improvement occurs with the addition of extinction coefficients. For more moderate particle sizes ($\bar{r}_0 = 0.5 \quad \sigma_0 = 1.6$), three backscatter coefficients alone will result in large errors for moderate uncertainty levels in the optical coefficients while significant improvement is observed with the addition of a single extinction channel. In this case, the second extinction channel adds little improvement. Finally, for smaller particles better representative of the urban accumulation mode ($\bar{r}_0 = 0.3 \quad \sigma_0 = 1.4$) the addition of either 1 or 2 extinction coefficients significantly improves the retrieval.

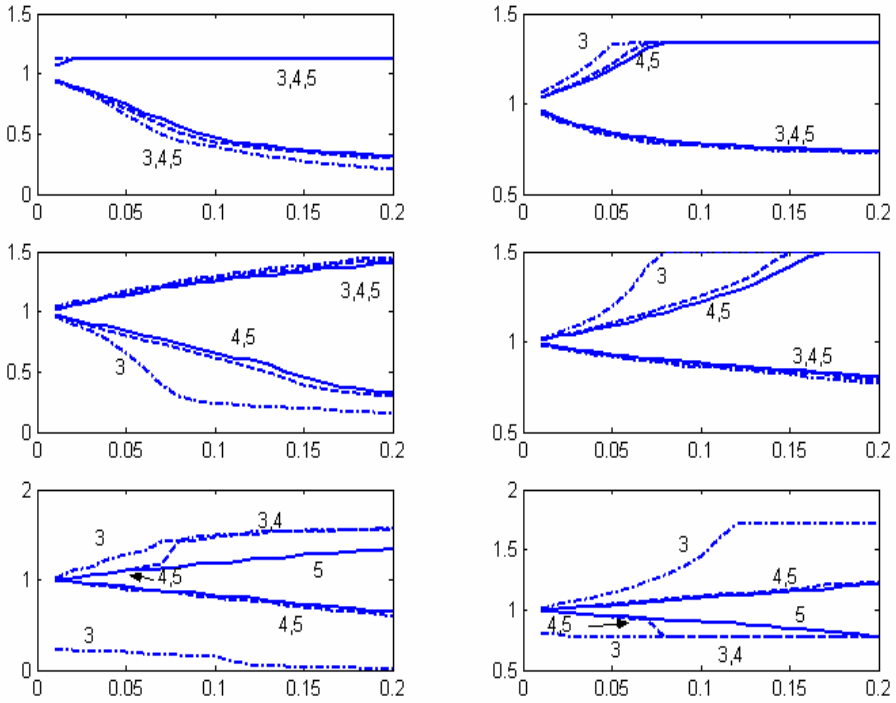

Figure 4.10. Comparisons of normalized parameter retrieval maximum and minimum bounds between different optical coefficient sets as a function of uncertainty level for different parameter sizes. All analyses use $m_r = 1.5 \quad m_i = -0.02$

a) $\bar{r}_0 = 0.8 \quad \sigma_0 = 1.8$ large particle size

b) $\bar{r}_0 = 0.5 \quad \sigma_0 = 1.6$ intermediate size

c) $\bar{r}_0 = 0.3 \quad \sigma_0 = 1.4$ small (accumulation mode) size

4.3.2 Bi-Modal Distributions

While the single lognormal mode is very amenable to graphical analysis, the most realistic aerosol particle size distribution models are bimodal size distributions that combine fine particulate (accumulation mode) aerosols with a coarse lognormal particle mode. This model is consistent with the global cluster analysis from a very extensive set of AERONET (NASA Aerosol Robotic Network) measurement records across the globe (Dubovik et al. 2002). In that study, the analysis of the AERONET data set yielded six distinct clusters described by the column optical (refractive index) and physical (Particle size distribution) properties and were identified as desert dust, biomass burning, rural (background), industrial pollution, marine, and dirty pollution. Each cluster has a unique set of mode parameters total volume, mean radius, GSD and complex refractive index $\left\{(V)_{f,c},(r)_{f,c},(\sigma)_{f,c},(m_r)_{f,c},(m_i)_{f,c}\right\}$ for both fine (f) and coarse (c) parameters,

$$\frac{dV}{d\log(r)} = \frac{V_f}{\sqrt{2\pi}\ln(\sigma_f)}\exp\left(-\frac{\ln(r/r_f)}{2\ln(\sigma_f)^2}\right) + \frac{V_c}{\sqrt{2\pi}\ln(\sigma_c)}\exp\left(-\frac{\ln(r/r_c)}{2\ln(\sigma_c)^2}\right) \tag{4.8}$$

One issue that would be of interest is the possibility of separating monomodal distributions from true bimodal distributions. This problem has been studied using extinction coefficients alone (O'Neill, Dubovik, and Eck 2001) where it was shown that if the coarse mode is not too abundant and the fine mode Angstrom coefficient is sufficiently large (to extract from the coarse flat Angstrom coefficient), such a separation is possible.

In Figure 4.11, we examine the situation where both modes in the bimodal distribution are fine modes with mode 1 given as,

$$\bar{r}_0 = 0.3 \quad \sigma_0 = 1.4 \quad m_r = 1.5 \quad m_i = -.02 \tag{4.9}$$

and mode 2,

$$\bar{r}_0 = 0.7 \quad \sigma_0 = 1.4 \quad m_r = 1.5 \quad m_i = -.02 \tag{4.10}$$

with a 10% uncertainty level. The mixing ratio was designed so both modes would have equal optical depths as 532nm. This situation is comparable to the bimodal distributions considered in Müller et al. 2001. In that work, although the bimodal distribution was successfully inverted, a larger optical data set $(6\beta + 2\alpha)$ was needed. Applying the graphical technique to optical coefficients $(3\beta + 2\alpha)$ derived from the bimodal distribution resulted in a parameter retrieval space whose mean radius is between the individual modes and whose GSD parameter is significantly larger as would be expected by a smeared out bimodal distribution. It is concluded that it is not possible to distinguish an underlying fine particle bimodal distribution using the graphical technique within the five optical coefficients $(3\beta + 2\alpha)$ scheme and more optical backscatter channels are needed.

However, the situation changes drastically if the two modes are sufficiently separated (i.e. coarse and fine mode). To see this, four Aeronet PSD retrieved over New York City were chosen, as illustrated in Figure 4.12. Panels a) and b) represent near single mode distributions while panels c) and d) are highly bi-modal examples. Instead of using the actual retrieved refractive index that has wavelength dependencies that

cannot be implemented within the context of the graphical technique, all retrievals were processed with a complex refractive index of $m_r = 1.5 \quad m_i = -.02$. Figure 4.13 shows the results of the graphical technique applied to optical coefficients derived from the four PSDs using a 10% error level. We note that for the near single mode PSDs, all three lidar schemes $(3\beta + 2\alpha)$, $(3\beta + 1\alpha)$ and (3β) produce reasonable results. However, with the strongly bimodal distributions, (3β) produces a false monomodal distribution in all cases, $(3\beta + 1\alpha)$ produces a false monomodal distribution for the more moderate coarse mode case (Figure 4.13d) while the $(3\beta + 2\alpha)$ does not produce any false monomodal distributions. This suggests that the $(3\beta + 2\alpha)$ scheme can potentially flag bimodal distributions.

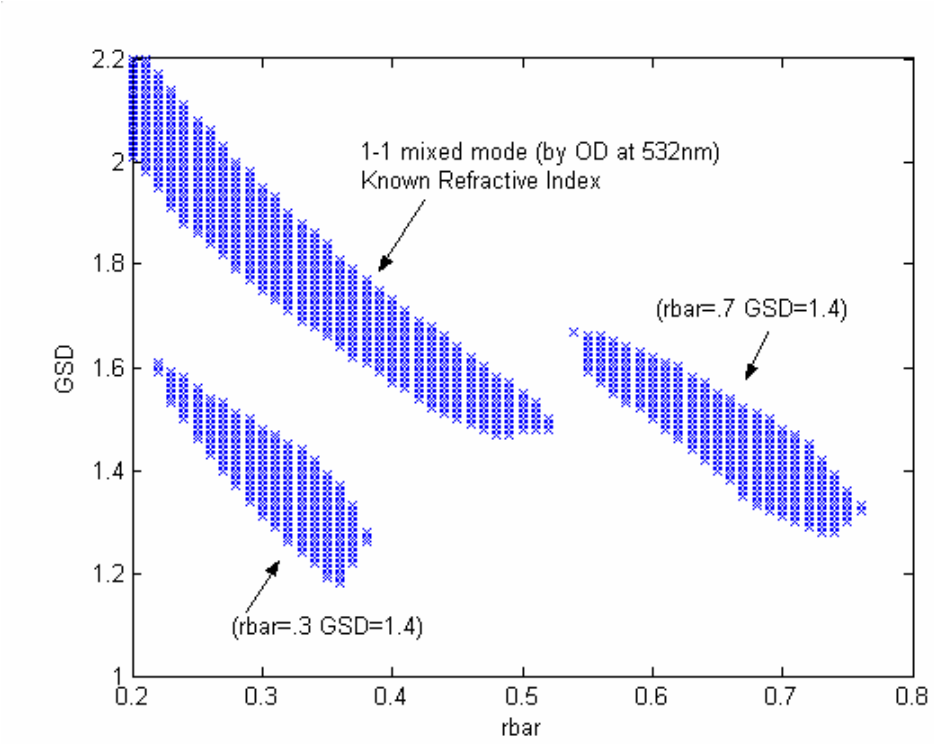

Figure 4.11. False monomodal retrieval from a set of optical coefficients $3\beta + 2\alpha$ generated from an initial bimodal distribution:
Mode 1: $\bar{r}_0 = 0.3 \quad \sigma_0 = 1.4 \quad m_r = 1.5 \quad m_i = -.02$
Mode 2: $\bar{r}_0 = 0.7 \quad \sigma_0 = 1.4 \quad m_r = 1.5 \quad m_i = -.02$
Also shown for reference are the retrievals for each individual mode.

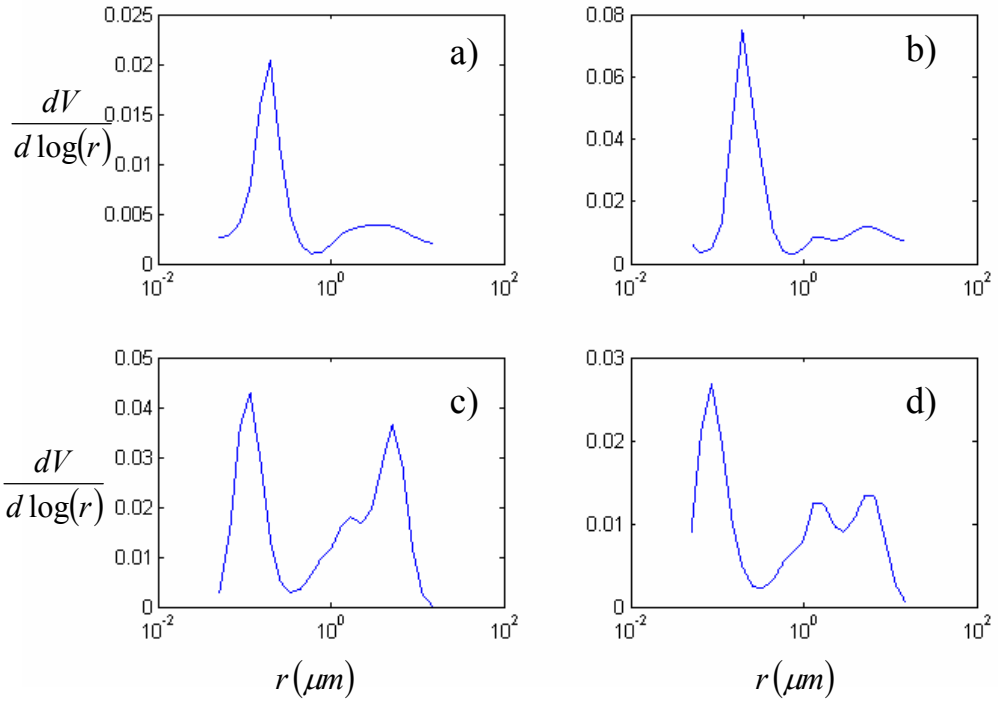

Figure 4.12. Representative bimodal aerosol PSD modes obtained from Aeronet over New York City. a) and b) are representative of near monomodal distributions while c) and d) were chosen for their significant bimodal characteristics.

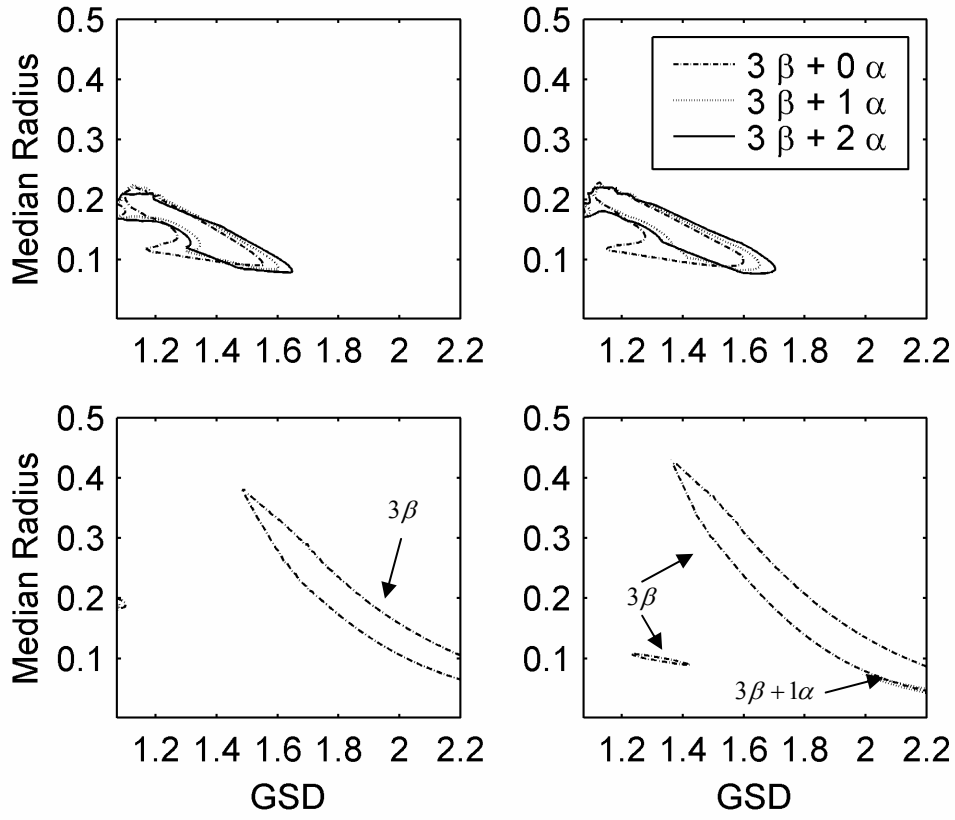

Figure 4.13. Graphical retrieval regions in the lognormal PSD parameter space for the different bimodal distributions displayed in Figure 4.12. The refractive index was chosen to be fixed at $m_r = 1.5 \quad m_i = -.02$. In panels c) and d) where the coarse aerosol mode is significant, the 3β (dashed line) and $3\beta + 1\alpha$ (dotted line) coefficient sets can result in false retrievals while there is no retrieval for the $3\beta + 2\alpha$ (solid line) set.

While the AERONET bimodal distributions were processed with the original a-priori refractive index of $m_r = 1.5 \quad m_i = -.02$, it might be argued that false monomodal distributions may be retrieved if an incorrect refractive index is used. This has been investigated numerically by varying the refractive index of the retrieval LUT and observing that no choice in the complex refractive index will retrieve a false monomodal PSD for the $(3\beta + 2\alpha)$ set. This result is in accord with the observation that the perturbations in the optical coefficient spectra due to a coarse particle mode are wavelength insensitive, but changes in the refractive index would result in very large modifications in the optical coefficient spectra. Therefore, refractive index changes cannot reproduce the modifications to the optical coefficient spectra due to an additional coarse mode.

4.4 Summary

The graphical technique employing log-normal particle size distribution parameters with an assumed refractive index is applied to the intercomparison of typical three wavelength backscatter lidar (1064, 532, 355 nm) and four and five wavelength systems augmented by additional nitrogen Raman channels that can yield estimates of extinction coefficients at 532 and 355 nm. It has been demonstrated that to obtain correct quantitative comparisons, the correlations between optical coefficient ratios used in graphical technique must be included. This is accomplished by refining the confidence bounds to account for such correlations. The graphical technique has also been extended by introducing a procedure for the retrieval of particle number density and its uncertainty. The results of this intercomparison indicate that significant improvement in retrieval error results from the addition of a single extinction channel. The addition of a second

extinction channel seems to mainly impact fine modes. Finally, the application of the graphical technique to bimodal distributions can result in false monomodal solutions. If the coarse mode fraction is not large, the technique can isolate and retrieve the fine mode parameter. When the coarse mode fraction is significant, it is demonstrated that the technique can potentially flag the bimodal distributions, when both Raman optical extinction channels are available, by yielding no false single mode retrieval.

5. Bayesian Paradigm

To begin, it is noted that a thorough treatment of uncertainty in aerosol size distribution retrieval from lidar measurements has not been discussed in the literature. The approach commonly employed shall be described as a forward Monte Carlo approach (Ligon, Gillespie, and Pellegrino 2000, Veselovskii et al. 2004). The forward Monte Carlo method can be described as follows: Starting with a set of values for the optical coefficients, an ensemble of measurement vectors is obtained by adding errors from a random number generator to the original values. An ensemble of solutions is then obtained by performing the inverse method on the ensemble of measurement vectors and the solution ensemble is examined to find its variances. Although it is not explicitly stated, the presumed rationale behind this method is that the ensemble of generated measurement vectors projected onto the range of attainable coefficient vectors is somehow representative of a statistical distribution of possible true measurement vectors. It is shown in section 6.4 B that this is not generally correct when one considers the posterior error distribution that is conditional on the measured values.

In fact, the forward Monte Carlo analysis answers the approximate question of what would the statistical distribution of inversion results be if the true values of the optical coefficients are $\mathbf{y}$,. This is useful in determining the overall performance of the inverse method, but does not answer the question of what is the statistical distribution of true aerosol distributions that could have generated those measurement values if the result from a measurement is $\hat{\mathbf{y}}$. That is the question that should be answered in order to ascribe uncertainty to a solution associated with a particular instance of a measurement.

This problem can be solved by Bayesian analysis, in which the parameters to be determined, traditionally denoted by $\mathbf{\theta}$, are treated as random variables.

In the Bayesian framework, the occurrence of a measurement is modeled as a two step sequence of realizations of random variables. A posterior probability function that expresses the uncertainty of $\mathbf{\theta}$ is derived from this model. For the purposes of this dissertation, it is assumed that the random variables can be modeled with probability density functions (PDFs). First the parameters $\mathbf{\theta}$ arise from an assumed prior PDF, $f_{\text{pri}}(\mathbf{\theta})$. The prior PDF gives a statistical description of the probabilities of the parameters to be retrieved in the absence of any measurements, and is based on prior statistical assumptions including any constraints due to physical unrealizability. For example, with the requirement that the particle density of an aerosol distribution must be positive, one would choose a prior PDF whose values are zero over the region where the particle density is negative. As another example, a prior PDF that is more heavily weighted at smoother aerosol distributions is implicit in the regularized inverse formula that is the basic building block of the methods of Müller et al. (1999a), and Veselovskii et al. (2002). After the realization of $\mathbf{\theta}$, the measurements $\hat{\mathbf{y}}$ then arise from the likelihood PDF, $f_{\text{li}}(\hat{\mathbf{y}} \mid \mathbf{\theta})$ depending on the resulting $\mathbf{\theta}$. The posterior PDF is derived from the prior and likelihood PDFs by

$$f_{\text{post}}(\mathbf{\theta} \mid \hat{\mathbf{y}}) = f_{\text{li}}(\hat{\mathbf{y}} \mid \mathbf{\theta}) f_{\text{pri}}(\mathbf{\theta}) \big/ f_{\text{marg}}(\hat{\mathbf{y}}), \tag{5.1}$$

where $f_{\text{marg}}(\hat{\mathbf{y}})$ is the marginal PDF for the measurements without any constraint on $\mathbf{\theta}$. The marginal PDF is included in the formulation to normalize the posterior PDF, and is given by

$$f_{\text{marg}}(\hat{\mathbf{y}}) = \int f_{\text{li}}(\hat{\mathbf{y}} \mid \boldsymbol{\theta}) f_{\text{pri}}(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} \,. \tag{5.2}$$

This formulation is a consequence of the concept of conditional probability. For two events *A* and *B* the probability of *A* under the condition *B* is defined as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \,, \tag{5.3}$$

thus resulting in

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \quad \text{(Papoulis 1991).} \tag{5.4}$$

In application to uncertainty computations, the condition *B* is the event that measurements $\hat{\mathbf{y}}$ are found to be some value, and *A* is an event that $\boldsymbol{\theta}$ belongs to some subset of $\boldsymbol{\theta}$'s domain.

With the purpose of trying to validate the forward Monte Carlo method, a first attempt to try to assess uncertainty using the conditional probability concept directly, resulted in the development of the reverse Monte Carlo method. In this method a database of outcomes, $(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, from the forward Monte Carlo method is created over a grid that spans a range of the parameter space. Given a measurement vector $\hat{\mathbf{y}}_0$, a retrieval is performed yielding a $\hat{\boldsymbol{\theta}}_0$ and the database is searched for outcomes where $\hat{\boldsymbol{\theta}} \approx \hat{\boldsymbol{\theta}}_0$, and the $\boldsymbol{\theta}$s from this subset of $(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ make up the output ensemble as shown in Figure 5.1.

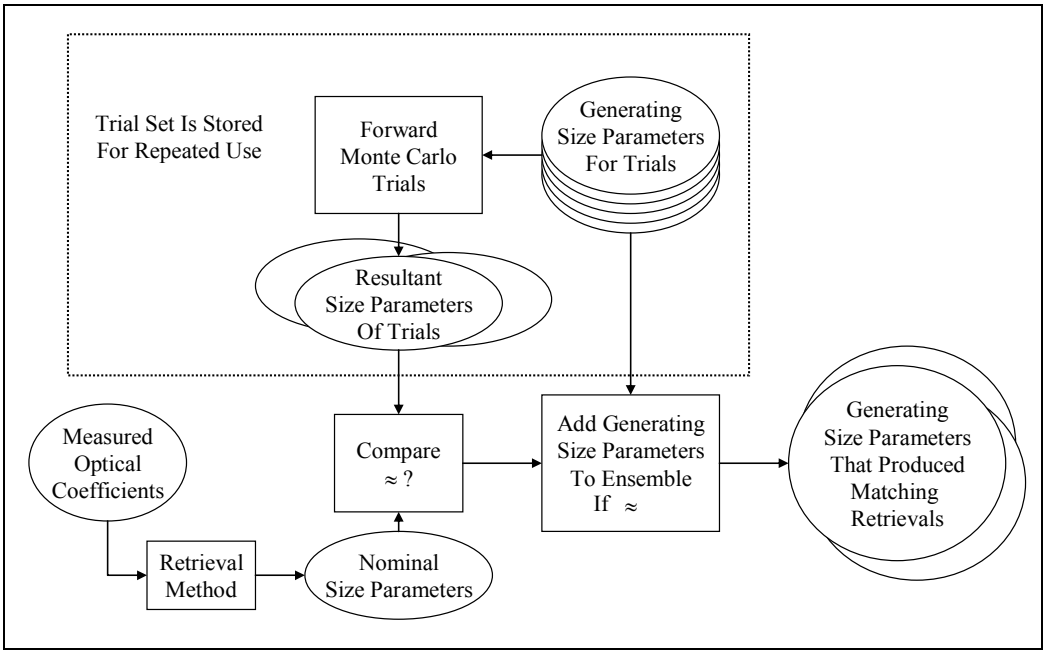

Figure 5.1. Reverse Monte Carlo method

This method was ultimately abandoned since its extension to problems of higher dimensions with the inclusion of index of refraction as an unknown variable appear to be infeasible. Furthermore, there is the potential for loss of information in the measurements since the condition B is $\hat{\boldsymbol{\theta}} \approx \hat{\boldsymbol{\theta}}_0$ rather than $\hat{\mathbf{y}} \approx \hat{\mathbf{y}}_0$. Nonetheless, it serves as an illustration of the differences in the forward and inverse statistical analyses.

In this study the forward and reverse Monte Carlo methods were applied using ratios of backscatter coefficients at 1064nm, 532nm and 355nm with an iterative graphical retrieval method described as follows. A two dimensional array for each wavelength's backscatter coefficients is created, which is derived from a single particle density, with element indexes representing different median radii and GSDs. Two arrays of ratios, R_1 and R_2, are created from the backscatter arrays. Given measured ratios, the ratio arrays are searched to find the elements that lie within an error margin ΔR by simultaneously satisfying the inequalities,

$$R_1 - \Delta R \le \left[R_1\right]_{i,j} \le R_1 + \Delta R \,, \tag{5.5a}$$

and

$$R_2 - \Delta R \le \left[R_2\right]_{i,j} \le R_2 + \Delta R \,. \tag{5.5b}$$

The search is repeated adjusting the error margin until one element remains, yielding the inversion result. In the following computations the arrays were calculated using an index of refraction of 1.5 + 0i over a range of 0.1-0.9μm in median radius and 1.2-2.4 in GSD.

Table 5.1 and Figure 5.2 show the error statistics computed with the forward and reverse Monte Carlo methods from synthetic error free measurements of backscatter

coefficients corresponding to the size parameters r_0 = 0.5μm and σ = 1.6. Although the synthetic measurement values are error free, the error model used for the statistical computation was Gaussian distributed with a standard deviation of 6.97%. The value 6.97% was chosen heuristically to produce a 10% standard deviation in ratio errors. Results from statistical computations using the Bayesian posterior probability density function (PDF) are also shown in Table 5.1 and Figure 5.2. The methodology for computing the PDF will be discussed in section 6. It is clear that there is good agreement between the reverse Monte Carlo method and the PDF analysis, while the forward Monte Carlo statistics underestimate the standard deviation and skew the mean to the initial size distribution parameters. A more extensive analysis of this effect will also be given in section 6.

Table 5.1. Uncertainty statistics as assessed by three methods

Ratios of Backscatter Coefficients		Median Radius	GSD
Initial size distribution parameters		0.6	1.5
Forward Monte Carlo	Mean	0.580	1.538
	Standard Deviation	0.0929	0.164
	Correlation Coefficient	-0.952	
Reverse Monte Carlo	Mean	0.528	1.635
	Standard Deviation	0.113	0.192
	Correlation Coefficient	-0.942	
Posterior PDF	Mean	0.521	1.647
	Standard Deviation	0.116	0.213
	Correlation Coefficient	-0.959	

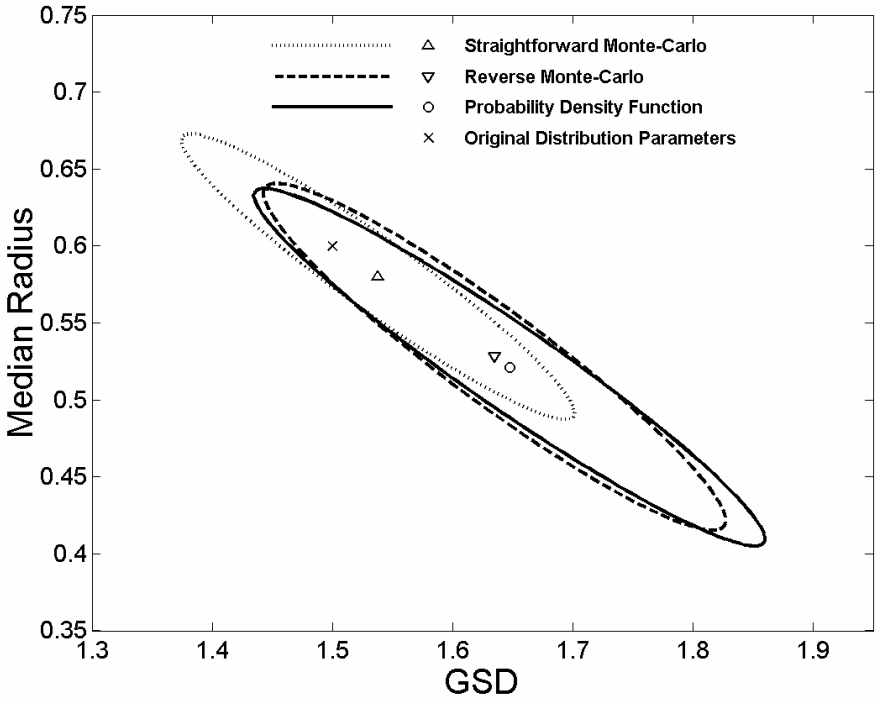

Figure 5.2. Comparison of uncertainty statistics computed with different methods. To facilitate comparisons, mean, standard deviation, and correlation coefficient are displayed as error ellipses which would represent a minimal confidence set of a Gaussian distribution with corresponding statistical parameters. The uncertainty distributions are not Gaussian; the ellipses are shown for comparison purposes only.

6. Bayesian Analysis

6.1 Analytic Error PDF Formulation

In this section, a formula is derived for calculating the posterior probability function in Eq. 5.1 with measured ratios $\hat{\mathbf{R}}$ and size parameters **p** replacing the $\hat{\mathbf{y}}$ and **θ** respectively. The likelihood PDF is derived from the ratio error PDF,

$$f_{\mathrm{li}}(\hat{\mathbf{R}} \mid \mathbf{p}) = f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}(\hat{\mathbf{R}}, \mathbf{p})) J_{\boldsymbol{\varepsilon}/\hat{\mathbf{R}}}(\hat{\mathbf{R}}, \mathbf{p}), \tag{6.1}$$

where $f_{\boldsymbol{\varepsilon}}$ represents the PDF for the error distribution, and $J_{\boldsymbol{\varepsilon}/\hat{\mathbf{R}}}(\hat{\mathbf{R}})$ represents the differential volume transformation function given by the determinant of the Jacobian matrix, i.e.

$$J_{\boldsymbol{\varepsilon}/\hat{\mathbf{R}}} = \left| \det\left(\frac{\partial \boldsymbol{\varepsilon}}{\partial \hat{\mathbf{R}}} \right) \right|, \tag{6.2}$$

where $\partial\boldsymbol{\varepsilon}/\partial\hat{\mathbf{R}}$ is shorthand for the matrix,

$$\frac{\partial \boldsymbol{\varepsilon}}{\partial \hat{\mathbf{R}}} = \begin{pmatrix} \partial\varepsilon_1/\partial\hat{R}_1 & \cdots & \partial\varepsilon_1/\partial\hat{R}_N \\ \vdots & \ddots & \vdots \\ \partial\varepsilon_N/\partial\hat{R}_1 & \cdots & \partial\varepsilon_N/\partial\hat{R}_N \end{pmatrix} \text{ (Papoulis 1991).} \tag{6.3}$$

An additive error model would have $\varepsilon_i = \hat{R}_i - R_i$ resulting in an identity matrix and a constant Jacobian. However, a non-additive formulation will be necessary in order to eliminate the particle number density as a parameter for retrieval, since the error

formulation must not have any dependence on it. This is most easily achieved by formulating the error in optical coefficients as a fractional error of the form

$$\hat{y}_i = y_i\left(1+\delta_i\right). \tag{6.4}$$

Considering a ratio of measured coefficients which are independent of N_0, we have

$$\hat{R}_k = \frac{y_i\left(1+\delta_i\right)}{y_j\left(1+\delta_j\right)} = R_k\left(1+\frac{1+\delta_i}{1+\delta_j}-1\right) = R_k\left(1+\varepsilon_k\right), \tag{6.5}$$

where R_k is the ratio of true optical coefficients y_i/y_j, thus allowing the ratio error to have the same parameter-independent fractional error formulation as the optical coefficients, where the fractional error of the ratio has the form

$$\varepsilon_k = \left(1+\delta_i\right)/\left(1+\delta_j\right)-1. \tag{6.6}$$

The form of $\boldsymbol{\varepsilon}(\hat{\mathbf{R}},\mathbf{p})$ is given from Eq. 6.5 as

$$\varepsilon_i = \frac{\hat{R}_i}{R_i(\mathbf{p})}-1, \tag{6.7}$$

resulting in a diagonal matrix in Eq. 6.3 and therefore

$$J_{\boldsymbol{\varepsilon}/\hat{\mathbf{R}}}(\hat{\mathbf{R}},\mathbf{p}) = \prod_i \frac{1}{R_i(\mathbf{p})}. \tag{6.8}$$

The formula for the ratio error PDF, $f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon})$, is derived by first establishing the set of ratios of the N optical coefficients as

$$R_i = y_i / y_N, i \in \{1,..., N-1\}. \tag{6.9}$$

The relationship between ratio error and optical coefficient error was already given in Eq. 6.6, but to derive the ratio error PDF, $f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon})$, from the optical coefficient error PDF, $f_{\boldsymbol{\delta}}(\boldsymbol{\delta})$, the dimension of the error vector must be maintained, which is done by using an extended ratio error vector by appending the N^{th} optical coefficient error to the ratio error vector so that

$$\varepsilon_i = \begin{cases} \dfrac{1+\delta_i}{1+\delta_N} - 1, & i \neq N \\ \delta_N, & i = N \end{cases}, \tag{6.10}$$

and by inverting these formulae we obtain

$$\delta_i = \begin{cases} \varepsilon_i + \varepsilon_N + \varepsilon_i \varepsilon_N, & i \neq N \\ \varepsilon_N, & i = N \end{cases}. \tag{6.11}$$

The PDF of the extended ratio error vector, $\boldsymbol{\varepsilon}'$ is given by

$$f_{\boldsymbol{\varepsilon}'}(\boldsymbol{\varepsilon}') = f_{\boldsymbol{\delta}}(\boldsymbol{\delta}(\boldsymbol{\varepsilon}')) J_{\boldsymbol{\delta}/\boldsymbol{\varepsilon}'}(\boldsymbol{\varepsilon}'), \tag{6.12}$$

and the matrix elements needed for $J_{\boldsymbol{\delta}/\boldsymbol{\varepsilon}'}(\boldsymbol{\varepsilon}')$ are derived from Eq. 6.11 as

$$\frac{\partial \boldsymbol{\delta}}{\partial \boldsymbol{\varepsilon}'} = \begin{pmatrix} 1+\varepsilon_N & 0 & \cdots & 0 & 1+\varepsilon_1 \\ 0 & 1+\varepsilon_N & \cdots & 0 & 1+\varepsilon_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1+\varepsilon_N & 1+\varepsilon_{N-1} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{6.13}$$

Therefore, the expression for the differential volume transformation is finally calculated as

$$J_{\boldsymbol{\delta}/\boldsymbol{\varepsilon}'}(\boldsymbol{\varepsilon}') = \left|(1+\varepsilon_N)^{N-1}\right| . \tag{6.14}$$

The fractional errors of the optical coefficients are assumed to have a Gaussian distribution and to be independent having standard deviations σ_i, which leads directly to the following formula for the extended ratio error vector PDF,

$$f_{\boldsymbol{\varepsilon}'}(\boldsymbol{\varepsilon}') = \frac{\left|(1+\varepsilon_N)^{N-1}\right|}{(2\pi)^{\frac{N}{2}}\prod_{i=1}^{N}\sigma_i} \exp\left(-\frac{1}{2}\left[\sum_{i=1}^{N-1}\frac{(\varepsilon_i+\varepsilon_N+\varepsilon_i\varepsilon_N)^2}{{\sigma_i}^2}+\frac{{\varepsilon_N}^2}{{\sigma_N}^2}\right]\right) . \tag{6.15}$$

Once the PDF in the extended space is obtained, the PDF can be projected onto the correct subspace by analytically integrating $f_{\boldsymbol{\varepsilon}'}(\boldsymbol{\varepsilon}')$ over ε_N to get $f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon})$, the PDF of the non-extended error vector. Expressing the argument of the exponential function as $-0.5[A(\varepsilon_N - B)^2 + C]$ and making the substitution $u = \varepsilon_N + B$, results in the expression,

$$f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) = \frac{1}{(2\pi)^{\frac{N}{2}}\prod_{i=1}^{N}\sigma_i}\exp(-\tfrac{1}{2}C)\int_{-\infty}^{\infty}\exp(-\tfrac{1}{2}A\,u^2)\left|(1-B+u)^{N-1}\right|\mathrm{d}\,u \tag{6.16a}$$

where A, B, and C are functions of $\boldsymbol{\varepsilon}$ given by

$$A = \frac{1}{{\sigma_N}^2} + \sum_{i=1}^{N-1}\frac{1+2\varepsilon_i+{\varepsilon_i}^2}{{\sigma_i}^2} , \tag{6.16b}$$

$$B = \frac{\sum_{i=1}^{N-1} \left(\varepsilon_i + {\varepsilon_i}^2 \right) \Big/ {\sigma_i}^2}{A} \text{, and} \tag{6.16c}$$

$$C = \sum_{i=1}^{N-1} \frac{{\varepsilon_i}^2}{{\sigma_i}^2} - A B^2 \text{.} \tag{6.16d}$$

Direct integration gives

$$f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) = \frac{\exp(-\frac{1}{2} C) \sum_{k=0}^{N-1} \left\{ \begin{array}{l} \left(\sqrt{\frac{2}{A}} \right)^{k+1} (1-B)^{N-1-k} \binom{N-1}{k} \\ \quad \times \left[\mathrm{I}_k(\infty) + \mathrm{I}_k(-\infty) - 2\,\mathrm{I}_k\left(\sqrt{\frac{A}{2}} (B-1) \right) \right] \end{array} \right\}}{(2\pi)^{\frac{N}{2}} \prod_{i=1}^{N} \sigma_i} \text{,} \tag{6.17}$$

where,

$$\mathrm{I}_k(u) = \int_0^u x^k \exp(-x^2)\, \mathrm{d}\, x \text{,} \tag{6.18}$$

which can be evaluated recursively with

$$\mathrm{I}_{k+2}(u) = \tfrac{1}{2} [(k+1) \mathrm{I}_k(u) - u^{k+1} \exp(-u^2)] \text{,} \tag{6.19a}$$

where

$$\mathrm{I}_0(u) = \tfrac{1}{2} \pi \operatorname{erf}(u) \text{,} \tag{6.19b}$$

and

$$I_1(u) = \frac{1}{2}[1 - \exp(-u^2)] \quad . \tag{6.19c}$$

With this formulation, the likelihood function can now be calculated. In the calculations, it is assumed that the prior PDF is uniform over the domain of distribution parameter values where the optical coefficients and zero have been tabulated elsewhere so that the posterior PDF is simply calculated by numerically integrating over the distribution parameter domain.

6.2 Comparisons

6.2.1 Simulations

In a typical multi-wavelength Nd:YAG laser lidar system, the optical coefficients which are obtained include the volume-backscatter at 1064nm, 532nm, and 355nm. However, previous work (Müller 1999b) has shown that backscatter coefficients are not sufficient to obtain microphysical parameters when the refractive index is unknown and extinction measurements are critical. Therefore, extra nitrogen Raman channels at 607nm and 387nm are included in the receiver, which provide independent measurements of the aerosol extinction at 532nm and 355nm. This $(3\beta + 2\alpha)$ configuration is taken to be the reference system to which other systems can be compared.

To begin, calculations are first performed based on the forward Monte Carlo assessment. In this scheme, the measurement vectors ensembles are numerically generated according to the statistical model governing the noise impressed on the optical data. For each ensemble element, a retrieval of the parameter vector is obtained based on a particular choice of inversion scheme. Finally, the statistics of the resultant retrieved parameter vector ensembles can be evaluated and compared.

The retrieval methods are implemented by searching a table derived from optical coefficients computed over a grid of particle median radii and geometric standard deviations (GSDs). For the computation grid, the median radius discretization is divided into two segments with the first segment being 42 points logarithmically spaced from 10nm to 50nm. The second segment ranges from 50nm to 1μm in increments of 2nm. The GSD discretization was from 1.1-2.4 in increments of 0.002.

In performing the inversion of the measurement data, two retrieval methods were used. In the first method, distribution parameters were retrieved by finding the **p** that minimizes the maximum deviation $\Delta(\mathbf{p})$ among the measurement ratios R_i, i.e.

$$\Delta(\mathbf{p}) = \max_i \{ \left| \frac{R_i(\mathbf{p})}{\hat{R}_i} - 1 \right| \} . \tag{6.20}$$

This first method will be referred to as the minimized maximum deviation (MMD) method. In the second retrieval method, distribution parameters were retrieved by finding the **p** that maximizes the posterior PDF as described in the previous section. The computation time for the maximized posterior PDF (MPPDF) method was substantially greater than in the MMD method, so the median radius and GSD grids were reduced by excluding alternating grid points in both dimensions.

Once the calculations were performed, the metrics used to quantify the differences between the uncertainty assessments were the marginal cumulative distribution functions (CDFs) in both median radius and GSD. For the sake of clarity, the "post" subscript and the $\hat{\mathbf{R}}$ will be omitted when discussing posterior PDFs or CDFs for the remainder of this paper, and subscripts will be used to denote a specific parameter or subset of parameters.

For the Bayesian assessment in this two parameter model, the marginal CDF is derived from the posterior PDF as

$$\mathrm{F}_X(x) = \int_{X\min}^{x} \int_{Y\min}^{Y\max} f_{X,Y}(x', y')\,\mathrm{d}\,y'\,\mathrm{d}\,x' \,, \tag{6.21}$$

with X representing the log-normal distribution parameter that is being considered, either $\bar{r}$ or σ, and Y representing the other parameter. The CDFs for the two forward Monte Carlo assessments were estimated from the ensemble of retrieval outcomes by the formula,

$$\hat{\mathrm{F}}_X(x) = N\{X \le x\}/N_{\text{total}} \,, \tag{6.22}$$

where $N\{X \le x\}$ represents the number of ensemble members with the distribution parameter under consideration less than or equal to x and N_{total} represents the total number of ensemble members.

Synthetic measurements of optical coefficients used as inputs to the forward Monte Carlo assessments for comparison were created over a grid representing true particle size distributions with median radii varying from 0.2–0.8μm in steps of 0.05μm and GSD varying from 1.3–2.1 in steps of 0.05. Optical coefficient error was modeled as in Eq. 6 with each δ_i being zero mean Gaussian distributed with a standard deviation of 0.1. In this analysis we are working with a 5-dimensional optical coefficient vector used to retrieve a 3-parameter aerosol model, therefore real measurements with errors would not be in the range of optical coefficient vectors that correspond to aerosol parameters. To avoid any possible effects of this issue, a single error vector, $\boldsymbol{\delta}_0$, was generated with 5 samples drawn from a Gaussian random number generator with the same statistical

properties as in the error model. This error vector was used to disturb the simulated true optical coefficients using Eq's 6 for each grid point. Figure 6.1 shows some examples of the CDFs derived from the forward Monte Carlo methods compared with Bayesian analytical CDFs. Examples of 2-dimensional scatter plots are shown superimposed on Bayesian PDF contours in Figure 6.2. We note that there are substantial discrepancies between the forward Monte Carlo estimated CDFs and the Bayesian CDFs and give a explanation of this result in section 6.2.2 after giving a theoretical discussion of error PDFs.

The differences in uncertainty between the forward Monte Carlo and Bayesian assessments were quantified using a deviation parameter,

$$\left(\mathrm{dev}\, X\right)_{\mathrm{meth}} = \int \left| \mathrm{F}_X(x) - \hat{\mathrm{F}}_{X,\mathrm{meth}}(x) \right| \mathrm{d}\, x \quad . \tag{6.23}$$

where $\mathrm{F}_X(x)$ represents the Bayesian CDF and "meth" represents the retrieval method used in the forward Monte Carlo assessment. Figure 6.3 shows images of the deviation parameter for both the MPPDF and MMD retrieval methods. Clearly, there are significant differences between the forward Monte Carlo assessment and the proper Bayesian assessment in both implementations. They also show that the forward Monte Carlo method using MPPDF retrievals generally gives an uncertainty assessment closer to the Bayesian assessment than when using MMD retrievals. To show that this is a general trend, the deviation parameter calculations were performed for four other constant error vectors as shown in Table 6.1. These data show that for each $\boldsymbol{\delta}_0$, the majority of computed deviation parameters are larger when using the MMD retrieval method than with the MPPDF retrieval method for both median radius and GSD.

Furthermore, considering all $\boldsymbol{\delta}_0$s combined, 66.2% of the computations among the different $\boldsymbol{\delta}_0$s and modeled true parameters show an increase in median radius CDF deviation by more than a factor of 2. It is also noted that the variation of results shown in Table 6.1 clearly shows the need for the use of the $\boldsymbol{\delta}_0$s.

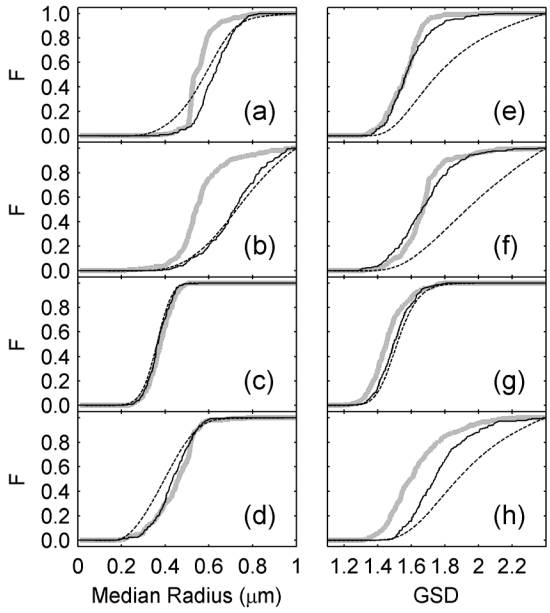

Figure 6.1. CDFs in median radius (a,b,c,d) and GSD (e,f,g,h) with modeled true distribution parameters $(\bar{r},\sigma)$ of: (0.6μm,1.6), (a,e); (0.6μm,2), (b,f); (0.3μm,1.6), (c,g); (0.3μm,2), (d,h). The dotted line represents the posterior CDF. The thin black line represents the estimated CDF using MPPDF retrievals. The thick gray line represents the estimated CDF using MMD retrievals.

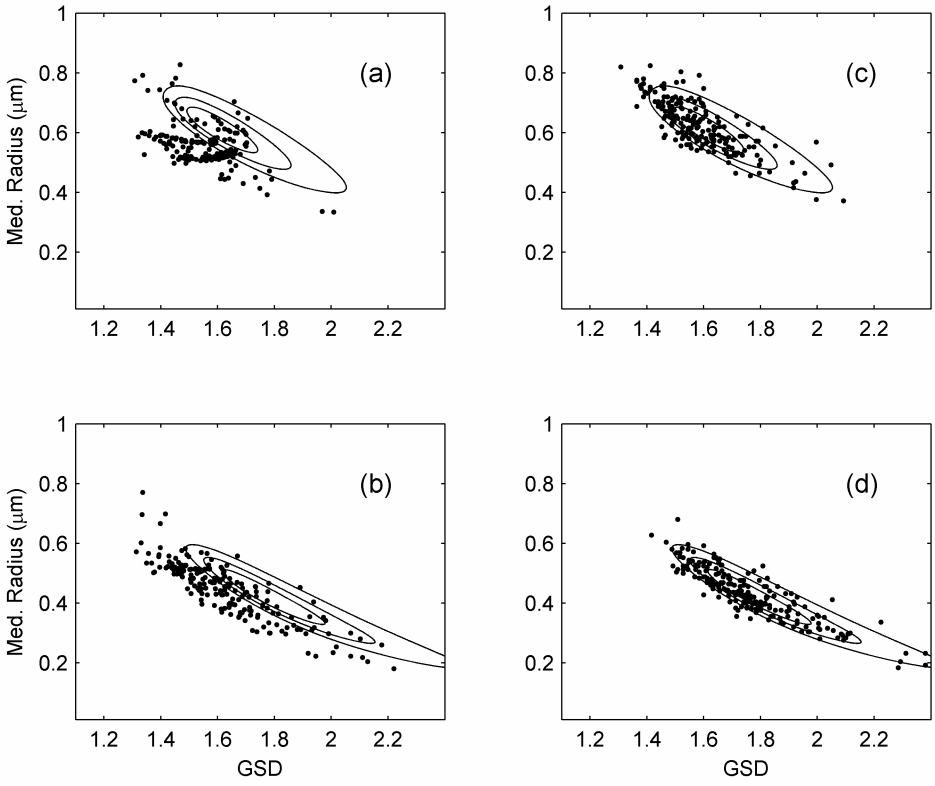

Figure 6.2. Scatter plots of forward Monte Carlo outcomes with contours of Bayesian PDFs. Outcomes using MMD retrievals are shown in (a) and (b). MPPDF retrievals are shown in (c) and (d). The modeled true distribution parameters, $(\bar{r},\sigma)$, are: (0.6μm,1.6), (a,c) [corresponds to Figure 6.1 (a) and (e)]; and (0.3μm,2), (b,d) [corresponds to Figure 6.1 (d) and (h)].

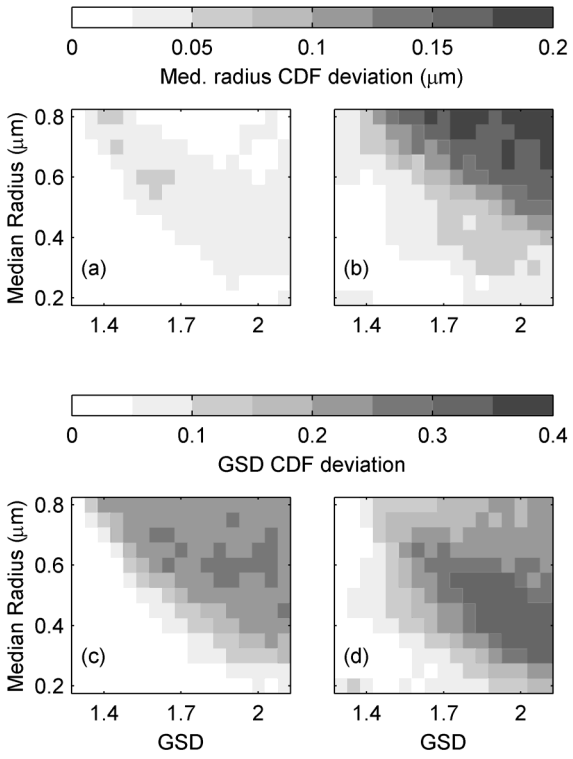

Figure 6.3. Median radius (a,b) and GSD (c,d) CDF deviation from Bayesian CDF as formulated in Eq. 32 using MPPDF (a,c) and MMD (b,d) retrieval methods.

Table 6.1. CDF Deviation Comparison

Error Vector, $\boldsymbol{\delta}_0$	Median Radius Deviation Ratio[a] Threshold		GSD Deviation Ratio[b] Threshold	
	1.0	2.0	1.0	2.0
	Percent above threshold			
1	95.0	60.2	70.6	38.5
2	95.9	74.7	51.6	26.2
3	90.0	63.3	100	67.9
4	89.6	60.6	69.2	46.2
5	93.7	71.9	57.9	36.2
All Combined	92.9	66.2	69.9	43.0

a Ratio of MMD to MPPDF deviation parameters, $(\mathrm{dev}_{\bar{r}})_{\mathrm{MMD}} / (\mathrm{dev}_{\bar{r}})_{\mathrm{MPPDF}}$

b $(\mathrm{dev}_{\sigma})_{\mathrm{MMD}} / (\mathrm{dev}_{\sigma})_{\mathrm{MPPDF}}$

6.2.2 Discussion

The differences in the Bayesian assessment from the forward Monte Carlo assessments can be attributed to a subtle consequence of the transformation between joint probability domains, which is that even if measurement error probabilities are independent from underlying parameters in the forward model, their posterior probabilities conditional to the measurements will have a different PDF than the one in the forward model. To see how this arises, we consider the joint PDF in the error vector and aerosol distribution parameter vector, $(\boldsymbol{\varepsilon}, \boldsymbol{\theta})$, given by

$$f_{\boldsymbol{\varepsilon},\boldsymbol{\theta}}(\boldsymbol{\varepsilon}, \boldsymbol{\theta}) = f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \tag{6.24}$$

so that measurement errors are statistically independent of $\boldsymbol{\theta}$. From this PDF we obtain the PDF in the domain of coefficient true values and measurement errors, $(\boldsymbol{\varepsilon}, \mathbf{y})$. For given error-free coefficient values, it is assumed that there is a unique inverse, $\boldsymbol{\theta} = \mathbf{T}^{-1}(\mathbf{y})$, giving the corresponding aerosol distribution parameters, so that the PDF in $(\boldsymbol{\varepsilon}, \mathbf{y})$ is $f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) f_{\boldsymbol{\theta}}[\mathbf{T}^{-1}(\mathbf{y})] J_{\boldsymbol{\theta}/\mathbf{y}}(\mathbf{y})$, where $J_{\bullet/\bullet}$ represents the differential volume transformation as defined in Eq's 6.13 and 6.14.

This process is now repeated by obtaining the PDF in measurement errors and measured values, $(\boldsymbol{\varepsilon}, \hat{\mathbf{y}})$. Using the fact that the measurement true values can be derived from measured values and measurement errors by an analytic function $\mathbf{y} = \mathbf{M}^{-1}(\hat{\mathbf{y}}, \boldsymbol{\varepsilon})$, then the PDF in this domain is

$$f_{\boldsymbol{\varepsilon},\hat{\mathbf{y}}}(\boldsymbol{\varepsilon}, \hat{\mathbf{y}}) = f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) f_{\boldsymbol{\theta}}\{\mathbf{T}^{-1}[\mathbf{M}^{-1}(\hat{\mathbf{y}}, \boldsymbol{\varepsilon})]\} J_{\boldsymbol{\theta}/\mathbf{y}}[\mathbf{M}^{-1}(\hat{\mathbf{y}}, \boldsymbol{\varepsilon})] J_{\mathbf{y}/\hat{\mathbf{y}}}(\hat{\mathbf{y}}, \boldsymbol{\varepsilon}). \tag{6.25}$$

(Note that $\mathbf{M}^{-1}(\hat{\mathbf{y}},\boldsymbol{\varepsilon})$ would typically be very straight forward; e.g. from Eq. 6.4 one would have $y_i = \hat{y}_i/(1+\varepsilon_i)$.) The conditional error PDF, $f_{\boldsymbol{\varepsilon}|\hat{\mathbf{y}}}(\boldsymbol{\varepsilon}\,|\,\hat{\mathbf{y}})$, is simply $f_{\boldsymbol{\varepsilon},\hat{\mathbf{y}}}(\boldsymbol{\varepsilon},\hat{\mathbf{y}})$ multiplied by normalizing constant for a given $\hat{\mathbf{y}}$. Hence it can been seen from Eq. 6.25 that differences in $f_{\boldsymbol{\varepsilon}|\hat{\mathbf{y}}}(\boldsymbol{\varepsilon}\,|\,\hat{\mathbf{y}})$ and $f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon})$ arise from any of the following three contributing factors which are: 1) from $f_{\boldsymbol{\theta}}\{\mathbf{T}^{-1}[\mathbf{M}^{-1}(\hat{\mathbf{y}},\boldsymbol{\varepsilon})]\}$ if the prior aerosol parameter PDF is non-uniform, 2) from $J_{\boldsymbol{\theta}/\mathbf{y}}[\mathbf{M}^{-1}(\hat{\mathbf{y}},\boldsymbol{\varepsilon})]$ if there is a non-linear relationship between aerosol distribution parameters and measurement true values, and 3) from $J_{\mathbf{y}/\hat{\mathbf{y}}}(\hat{\mathbf{y}},\boldsymbol{\varepsilon})$ if the error model is non-additive.

The second condition is present in our analyses due to the non-linear nature of the log-normal particle size distribution parameterization. Reviewing Figure 6.1, the fact that larger discrepancies between the forward Monte Carlo estimated and Bayesian CDFs are associated with higher uncertainties is indicative of this condition. In current methods based on regularization (Müller et al. 1999a, 1999b, 2001; Veselovskii et al. 2004), a non-uniform prior PDF is implicit by favoring smoother solutions.

The fractional error model that we use makes the third condition present. Alternatively the fractional error model could be viewed as a parameter-dependent error distribution, specifically additive Gaussian distributed error with a standard deviation proportional to the true optical coefficient of the aerosol distribution. In advocating the usefulness of Bayesian methods, it should be noted that if a non-independent error formulation were used, Eq. 6.24 would be $f_{\boldsymbol{\varepsilon},\boldsymbol{\theta}}(\boldsymbol{\varepsilon},\boldsymbol{\theta}) = f_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}\,|\,\boldsymbol{\theta})f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. In that case a forward Monte Carlo computation would not be feasible since one would not have $\boldsymbol{\theta}$

available to be able to implement a random error generator. These issues have not been previously addressed in the context of inferring aerosol distribution uncertainty from lidar measurements.

6.3 Applications to unknown index of refraction

The posterior PDF can be used to assess uncertainty by deriving regions S that have a specific credibility (i.e. *a posteriori* probability) C_0. In 2-dimensions a family of sets can be defined as the regions enclosed by contours of $f(\mathbf{p}) = L$,

$$S(L) = \{\mathbf{p} : f(\mathbf{p}) \geq L\}, \tag{6.26}$$

corresponding to different values of L. (This is a departure from conventional notation in the sense that $S(L)$ refers to a function that maps the set of real numbers into the family of all subsets of the entire parameter domain.) This defines credibility as a function of L by

$$C(L) = \int_{S(L)} f(\mathbf{p})\,\mathrm{d}\,\mathbf{p}, \tag{6.27}$$

and a credible set can be found as $S(L_0)$ with L_0 such that $C(L_0) = C_0$. Sets found this way have the property of being the optimal credible set in the sense that they are sets with the least total area. Although other properties that might be of interest are not necessarily satisfied, such as connectedness or being the credible set having the least maximum distance among the points of the set, these issues are beyond the scope of this dissertation.

With the inclusion of the complex index of refraction, $\mathbf{m} = (m_{\mathrm{Re}}, m_{\mathrm{Im}})$, as unknown parameters, credible sets derived as from Eq. 6.26 in the full 4-dimensional domain are

not easily visualized. One can obtain marginal posterior PDFs in size parameters or index of refraction by

$$f_{\mathbf{p}}(\mathbf{p}) = \int_{\Omega_{\mathbf{m}}} f_{\mathbf{p},\mathbf{m}}(\mathbf{p},\mathbf{m})\,\mathrm{d}\,\mathbf{m} \quad , \tag{6.28a}$$

and

$$f_{\mathbf{m}}(\mathbf{m}) = \int_{\Omega_{\mathbf{p}}} f_{\mathbf{p},\mathbf{m}}(\mathbf{p},\mathbf{m})\,\mathrm{d}\,\mathbf{p} \quad , \tag{6.28b}$$

where $\Omega_{\mathbf{p}}$ and $\Omega_{\mathbf{m}}$ represent the entire subdomains of size parameters and indexes of refration respectively. However, credibilities yeilded in applying Eq. 6.27 to the marginal PDFs represent credibilities of the parameters of one subdomain belonging to that particular $S(L)$ while there is no restriction on the parameters of the other subdomain; thus the 4-dimensional credible set that may be found from either $f_{\mathbf{p}}(\mathbf{p})$ or $f_{\mathbf{m}}(\mathbf{m})$ is either the cartesian product $S_{\mathbf{p}} \times \Omega_{\mathbf{m}}$ or $\Omega_{\mathbf{p}} \times S_{\mathbf{m}}$. Eq. 6.26 may still be applied to the marginal PDFs to define families of sets $S_{\mathbf{p}}(L_{\mathbf{p}})$ and $S_{\mathbf{m}}(L_{\mathbf{m}})$, in which case Eq. 6.27 becomes

$$C(L_{\mathbf{p}}, L_{\mathbf{m}}) = \int_{S_{\mathbf{p}}(L_{\mathbf{p}}) \times S_{\mathbf{m}}(L_{\mathbf{m}})} f_{\mathbf{p},\mathbf{m}}(\mathbf{p},\mathbf{m})\,\mathrm{d}\,\mathbf{p}\,\mathrm{d}\,\mathbf{m} \quad , \tag{6.29}$$

with $C(L_{\mathbf{p}}, L_{\mathbf{m}})$ being monotonicly decreasing in both $L_{\mathbf{p}}$ and $L_{\mathbf{m}}$. Now the solution to $C(L_{\mathbf{p}}, L_{\mathbf{m}}) = C_0$ is a contour in $(L_{\mathbf{p}}, L_{\mathbf{m}})$ rather than a single point. The credible set is

chosen from the $L_{\mathbf{p}}$ and $L_{\mathbf{m}}$ on the contour that minimizes $A_{\mathbf{p}}A_{\mathbf{m}}$, where $A_{\mathbf{p}}$ and $A_{\mathbf{m}}$ represent the areas of $S_{\mathbf{p}}(L_{\mathbf{p}})$ and $S_{\mathbf{m}}(L_{\mathbf{m}})$ respectively.

In implementing this method, the median radius grid was divided into three segments with the first segment being 25 points logarithmically spaced from 2nm to 20nm. The second segment ranged from 20nm to 200nm in increments of 2nm. The third segment ranged from 200nm to 900nm in increments of 10nm. The GSD grid was divided into two segments with the fist segment being 21 points ranging from 1.074 to 1.200 such that σ-1 was logarithmically spaced, and the second segment ranging from 1.2 to 2.4 in increments of 0.01. The method was applied to simulated measurements with true distribution parameters of ($\bar{r}$, σ, m_{Re}, m_{Im}) = (0.3μm, 1.4, 1.5, -0.03) with the errors described in Eq. 6.4 with 10% standard deviation. This was repeated for five possible measurement configurations delineated in Table 6.2.

An example of the total set volume minimization procedure can be seen in Figure 6.4 where the total set volume, $A_{\mathbf{p}}A_{\mathbf{m}}$, is plotted versus $A_{\mathbf{m}}$ with 90% credibility being specified. The minimum values on the plots represent the optimal tradeoff between uncertainty in size parameters and in index of refraction. As the set in one subdomain decreases, the set in the other subdomain must increase to maintain the same credibility level. The minimal set in one subdomain occurs when the set in the other subdomain is the entire subdomain. We refer to this minimal set as the individual credible set. These are represented in Figure 6.4 as the end points of the plots, the leftmost and rightmost points representing the individual credible sets for index of refraction and size parameters respectively. The points where the plots have minimum values represent the optimal 4-dimensional credible set and we refer to the sets in the subdomains as the simultaneous

sets. The resulting credible set components are shown in Figure 6.5 showing the importance of extinction measurements when not assuming an index of refraction. Comparing Figure 6.5 (a) and (f) with Figure 6.5 (c) and (h) demonstrates that the inclusion of only one extinction measurement with three backscatter measurements results in a decrease in uncertainty that is clearly more substantial than improvements in uncertainty that would result from including 5 additional backscatter measurements as seen in Figure 6.5 (b) and (g). These results are in qualitative agreement with previous studies carried out with regularization (Müller 1999b). The band structure seen in the index of refraction credible sets has been demonstrated by Böckmann (2001) in images in **m** of the deviation of measured values from the true coefficient values corresponding to a single solution from the measurements. This structure was also shown by Müller et al. (2001) in images in **m** of the number of solutions obtained among different basis supports, whose deviations in true from measured coefficients were below a certain threshold. However in this case this structure is connected directly to a posterior probability.

Table 6.2: Measurement Groups

Group	Backscatter λ (nm)								Extinction λ (nm)	
Index	355	400	416	532	683	710	800	1064	355	532
1	x	x		x		x	x	x	x	x
2	x			x				x	x	x
3	x			x				x	x	
4	x	x	x	x	x	x	x	x		
5	x			x				x		

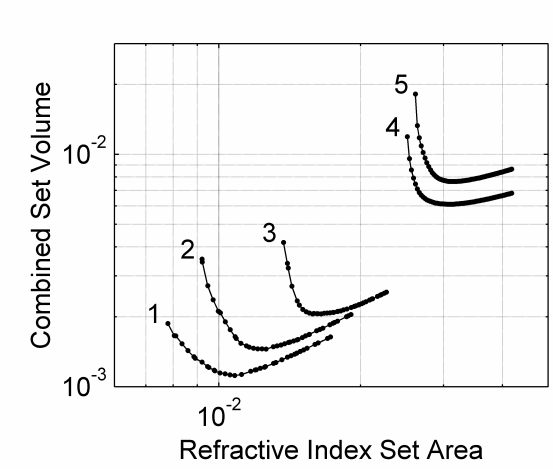

Figure 6.4. $A_{\mathbf{p}}A_{\mathbf{m}}$ vs. $A_{\mathbf{m}}$ of sets with 90% credibility for five measurement groups.

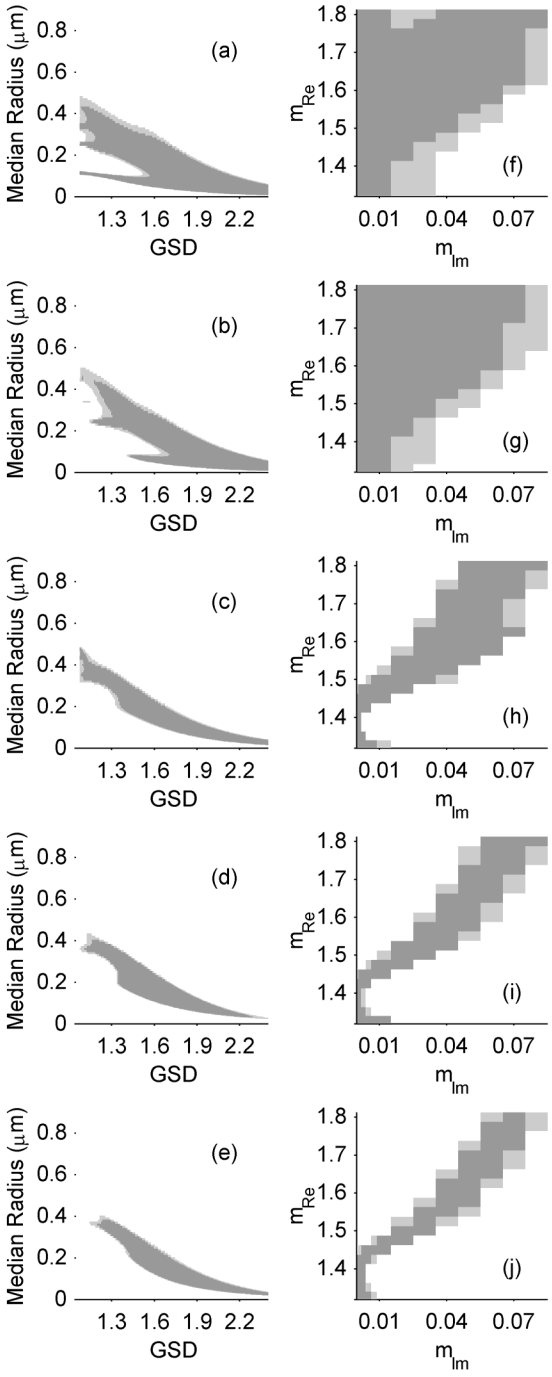

Figure 6.5. Individual (dark gray) and simultaneous (light gray) sets of 90% credibility for measurement groups 5 (a,f), 4 (b,g), 3 (c,h), 2 (d,i), and 1 (e,j).

7. Outlook and Conclusion

7.1 Options for Uncertainty Assessment

As has been shown so far in this dissertation, there are two alternatives for uncertainty assessment for particle size distributions from lidar measurements. A rigorous assessment can be performed with a parameterized particle size distribution, but this is valid only in a highly constrained model. The extension of the parameterization methods to models with higher dimension would seem to be infeasible in computational aspects. The dimensionality of the problem has been reduced analytically by taking advantage of a fractional error model, but providence cannot be guaranteed for arbitrary error models. The other alternative is that a more generalized aerosol model can be used and point estimates of particle size distributions can be obtained from classical linear inverse theory. Even though these point estimates may be based on a Bayesian posterior PDF, it has been shown in the parameterized model that Bayesian based point estimates do not necessarily lead to a Bayesian uncertainty distribution when the estimates are used in forward Monte Carlo assessments—although it has not been demonstrated that such discrepancies extend to classical linear inverse estimates. Furthermore, the Bayesian properties of the linear regularized solution are only valid for Gaussian—or possibly also Gaussian-like—error models with a specific prior PDF, and again an analytic solution may not be derivable for arbitrary forward models.

The relevancy of point estimates of aerosol size distributions in high dimensional models is highly suspect considering the relatively small number of optical coefficient measurements that are taken with lidar. However, it is not unreasonable to suppose that an uncertainty distribution of aerosol macroscopic parameters is well defined. Markov

chain Monte Carlo (MCMC) methods offer an elegant solution to this problem by offering a way of generating samples from arbitrary un-normalized multi-dimensional PDFs. These samples could then give samples of computed macroscopic parameters such as, effective radius, number concentration, surface area or volume concentration, coarse to fine mode ratio, optical coefficients at non-lidar wavelengths, index of refraction, or single-scattering albedo.

7.2 Background on Markov Chain Monte Carlo: Metropolis-Hastings Algorithm

Markov chain Monte Carlo methods are increasingly being used in statistical inverse problems. For example, Tamminen (2004) has applied a MCMC method to the retrieval of gaseous atmospheric constituents (O3, NO2, and NO3) from satellite based stellar occultation measurements. A basic description of a Markov Chain is that it is a sequence of random variables (RVs) where the probability function for each RV in terms of all previous RVs in the chain is dependent only on the RV immediately preceding it and is identical for each RV in the chain. This process is described by a transition kernel $K(\boldsymbol{\varphi},\boldsymbol{\theta})$. The kernel is the conditional PDF for the next RV, $\boldsymbol{\Phi}$, in the chain given that the previous RV, $\boldsymbol{\Theta} = \boldsymbol{\theta}$. Thus if $\boldsymbol{\Theta}$ has a PDF of $f_n(\boldsymbol{\theta})$, the PDF of $\boldsymbol{\Phi}$ is given by

$$f_{n+1}(\boldsymbol{\varphi}) = \int K(\boldsymbol{\varphi},\boldsymbol{\theta}) f_n(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \quad . \tag{7.1}$$

The goal of MCMC in this application is to find a kernel so that given an initial PDF, the PDFs of subsequent RVs converge to a target PDF, $\pi(\boldsymbol{\theta})$, as the RVs are further down the chain. In this case $\pi(\boldsymbol{\theta})$ would be the posterior PDF. This has been achieved by

the Metropolis–Hastings algorithm whose essential descriptions as described in chapter 1 of Gilks, Richardson, and Spiegelhalter (1996) are summarized in this and the following paragraphs. The generalized Metropolis–Hastings algorithm is implemented as follows: One begins with an initial outcome of $\boldsymbol{\Theta}_0$. At each link in the chain a candidate outcome, $\boldsymbol{\Phi}$, is obtained with a random generator that is generally dependent of the current outcome, $\boldsymbol{\Theta}_n$. The candidate outcome may then be accepted as the next outcome in the chain with a probability of

$$\alpha(\boldsymbol{\Phi},\boldsymbol{\Theta}) = \min\left\{1, \frac{\pi(\boldsymbol{\Phi})q(\boldsymbol{\Theta}\mid\boldsymbol{\Phi})}{\pi(\boldsymbol{\Theta})q(\boldsymbol{\Phi}\mid\boldsymbol{\Theta})}\right\}, \tag{7.2}$$

where $q(\boldsymbol{\varphi}|\boldsymbol{\theta})$ represents the conditional PDF for generating $\boldsymbol{\varphi}$ when $\boldsymbol{\Theta}_n = \boldsymbol{\theta}$. If the candidate is accepted then $\boldsymbol{\Phi} \rightarrow \boldsymbol{\Theta}_{n+1}$, otherwise the candidate is rejected and $\boldsymbol{\Theta}_n \rightarrow \boldsymbol{\Theta}_{n+1}$. The remarkable thing about this algorithm is that it has been proven that $f_n(\boldsymbol{\theta})$ always converges to $\pi(\boldsymbol{\theta})$ as $n \rightarrow \infty$ regardless of the shape of q if all regions of the domain of $\boldsymbol{\Theta}$ have a non-zero probability of being generated (see chapters 3 and 4 of Gilks et al. 1996).

The types of candidate generators that are relevant to the proposed work are the random walk and the independence sampler. In the random walk, candidate outcomes are distributed with a mean close to the current outcome. In the independence sampler q is chosen to be a function only of the candidate outcome. To obtain a higher acceptance rate it is chosen so that it is close to the target distribution.

Figure 7.1 shows the results of using a Metropolis-Hastings algorithm to sample the posterior PDF in a log-normal parameterized aerosol distribution model compared with the forward Monte Carlo analysis. Synthetic measurements of backscatter at 355,

532, and 1064 nm, and extinction at 355 and 532 nm were created for four distribution parameter pairs of median particle radius and geometric standard deviation (GSD) at (0.6μm,1.6), (a,e); (0.6μm,2), (b,f); (0.3μm,1.6), (c,g); and (0.3μm,2), (d,h). Ratios of optical coefficients were used to eliminate the particle number density as a parameter for retrieval. The figure shows cumulative distribution functions (CDFs) in Median radius (a-d) and GSD (e-h), estimated from Metropolis-Hastings ensembles (thin solid line), and forward Monte Carlo ensembles (thick gray line), compared with computed posterior CDFs (thin dotted line). For the forward Monte Carlo analysis, solutions were obtained by finding the parameters that maximized the posterior PDF. The majority of the Metropolis-Hastings CDFs are barely distinguishable from the analytic CDFs while the forward Monte Carlo CDFs show significant deviations.

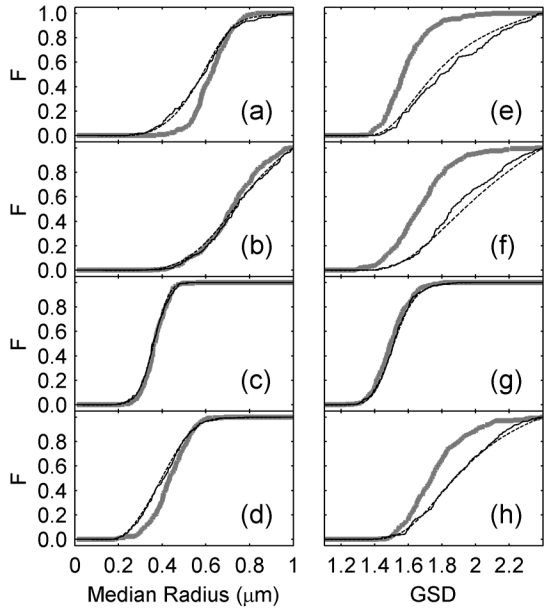

Figure 7.1. Estimated cumulative distribution functions (CDFs) of Monte Carlo outcomes and Bayesian CDFs

7.3 Proposed Implementation

For the final part of this paper, a robust Metropolis-Hastings based algorithm for generating samples from a multi-dimensional PDF is proposed. In the algorithm, the solution ensemble would be taken as a terminal link of an ensemble of ~500 parallel Markov chains. The ensemble would start out generated from an initial distribution, $f_0(\boldsymbol{\theta})$, rather than a single value for $\boldsymbol{\Theta}_0$. The most important reason for this approach is that the algorithm would use the ensemble to adapt the transition kernel to expedite convergence. After several links with a random walk candidate generator, ensemble members would be divided into clusters. The transition kernel would then be customly hybridized for the chain by allowing sampling from a randomly selected Gaussian distribution with a mean and covariance of a corresponding cluster as an alternative with a specified probability to a random step. At this stage in the chain the candidate generator PDF would then be

$$q(\boldsymbol{\varphi} \mid \boldsymbol{\theta}) = p_W q_W(\boldsymbol{\varphi} \mid \boldsymbol{\theta}) + \sum_{i=1}^{N_c} p_i q_i(\boldsymbol{\varphi}), \tag{7.3a}$$

with

$$\sum_{i=1}^{N_c} p_i = 1 - p_W, \tag{7.3b}$$

where p_w is the probability that a random step is chosen, $q_w(\boldsymbol{\varphi}|\boldsymbol{\theta})$ is the random step PDF, $q_i(\boldsymbol{\varphi})$ is the PDF corresponding to the i^{th} cluster, N_C is the number of clusters, and p_i is the probability of sampling from $q_i(\boldsymbol{\varphi})$. To expedite convergence the algorithm would

include an initial period of iterations using larger p_w and smaller p_i values in the acceptance probability formula than would be used in the actual generation. This was a technique that was used in the work that has been presented in section 7.2, where $p_w = 1$ and $p_i = 0$ was used in the acceptance formula. The reason for this technique is that in a teleportation, the value of $q(\boldsymbol{\Theta}|\boldsymbol{\Phi})/q(\boldsymbol{\Phi}|\boldsymbol{\Theta})$ would generally be very small since the term would typically be dominated by $q_n(\boldsymbol{\Phi})$ in the denominator, where n is the selected index, due to large distances between $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$. After initial iterations p_w and p_i values would return to their generation values, since otherwise the chain would not be convergent to $\pi(\boldsymbol{\theta})$.

Clustering would be done by a *k*-means based method (Jain, Murty, and Flynn 1999) with an initial cluster center, C_1, chosen as the member with the highest target PDF value. Since the goal is to find separated regions with appreciable probabilities, the next center, C_2 would be chosen as the member that maximizes $d(\boldsymbol{\Theta},C_1)^m\pi(\boldsymbol{\Theta})$, where d represents the Euclidean distance. The parameter m is included so that more or less weight can be given to distances from centers relative to PDF values. Subsequent centers would then be chosen as the member that maximizes $d_{\min}(\boldsymbol{\Theta},\{C_i\})^m\pi(\boldsymbol{\Theta})$, where

$$d_{\min}(\boldsymbol{\Theta},\{C_i\}) = \min_i\{\, d(\boldsymbol{\Theta},C_i) : i \in [1..N_C]\,\}, \tag{7.4}$$

where N_C is the number cluster centers determined so far. Centers would then be recomputed with a weighted mean, weighted with the target PDF, and members would then be reassigned according to the new centers.

This algorithm would use many of the setup characteristics as used by Veselovskii et al. (2002, 2004). Aerosol distributions would be expressed as a linear combination of

triangular basis functions with a variable index of refraction, and a smoothness prior with positivity constraints would be used. For the generation of the initial ensemble different methods would be tried. The first method would generate rather arbitrary aerosol distributions to obtain the best chances of finding regions of interest. This would be necessary at first to determine if and when isolated regions of interest exist since it would be beneficial to computation time to minimize the use of the hybrid sampler if it is not necessary. In such cases the initial ensemble could be generated by adding random disturbances to Tikhonov solutions from kernel matrixes with randomly selected indexes of refraction.

Using an ensemble would allow the chain to be monitored for convergence by observing the progression of statistics such as the ensemble means and variances of individual basis weights or those of sums of groups of adjacent weights. Once these statistics appear to reach their steady-state values, convergence would be checked by disturbing the ensemble in various ways. These disturbances would be just large enough to push the ensemble statistics beyond their variations along the chain. The algorithm would then check to see that further iterations move the statistics towards the same values of the undisturbed ensemble. Lienert, Porter, and Sharma (2001) used a genetic algorithm which is conceptually similar to the proposed algorithm but does not have a rigorous uncertainty assessment.

7.4 Conclusion

Although this dissertation is mostly written from a Bayesian perspective, it must be noted that Bayesian analysis is not the be-all and end-all to statistical inverse problems. Ideally, if one has a method of deriving Bayesian credible sets, membership of

true retrieval parameters should occur with a probability equal to the prescribed credibility. Kaipio and Sommersalo (2005) make the argument that the credibility is not really a probability by using an example where $\boldsymbol{\theta}$ is a value with a low prior PDF value and show that $\boldsymbol{\theta}$ is rarely a member of the derived credible set. However they neglect that Bayesian credibility is the probability of membership when $\boldsymbol{\theta}$ is a random variable, and that membership probability is not claimed to be equal to the credibility under the condition of a constraint on $\boldsymbol{\theta}$. Nonetheless, their example does bring up an important issue, which is that there may be some $\boldsymbol{\theta}$s that tend not to get retrieved well when Bayesian analysis is applied. It may be that $\boldsymbol{\theta}$s with low prior probability density are the instances that one particularly wants to detect. After all, how interesting is it to detect something that is already expected? These issues are of profound importance and are relevant not just to the areas of optical remote sensing and electrical engineering, but to larger fields of study as well, since they address questions concerning the nature of knowledge itself.

References

D. Althausen, D. Müller, A. Ansmann, U. Wandinger, H. Hube, E. Clauder, and S. Zörner, "Scanning 6-wavelength 11-channel aerosol lidar," *Journal of Atmospheric and Oceanic Technology* **17,** 1469–1482 (2000).

A. Ansmann, U. Wandinger, M. Riebesell, C. Weitkamp, and W. Michaelis, "Independent measurement of extinction and backscatter profiles in cirrus clouds using a combined Raman elastic-backscatter Lidar," *Applied Optics* **31**, 7113-7131 (1992).

D. R. Bates, "Rayleigh scattering by air," *Planetary and Space Science* **32**, 785–790 (1984).

L. R. Bissonnette, G. Roy, and N. Roy, "Multiple-scattering-based lidar retrieval: method and results of cloud probings," *Applied Optics* **44**, 5565-5581 (2005).

C. Böckmann, "Hybrid regularization method for the ill-posed inversion of multiwavelength lidar data in the retrieval of aerosol size distributions," *Applied Optics* **40,** 1329–1342 (2001).

C. Böckmann, U. Wandinger, A. Ansmann, J. Bösenberg, V. Amiridis, A. Boselli, A. Delaval, F. De Tomasi, M. Frioud, I. V. Grigorov, A. Hågård, M. Horvat, M. Iarlori, L. Komguem, S. Kreipl, G. Larchevêque, V. Matthias, A. Papayannis, G. Pappalardo, F. Rocadenbosch, J. A. Rodrigues, J. Schneider, V. Shcherbakov, and M. Wiegner, "Aerosol lidar intercomparison in the framework of the EARLINET Project. 2. Aerosol backscatter algorithms," *Applied Optics* **43**, 977-989 (2004).

C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, (Wiley, New York, 1983).

A. Bucholtz, "Rayleigh-scattering calculations for the terrestrial atmosphere," *Applied Optics* **34**, 2765-2773 (1995).

M. Del Guasta, M. Morandi, L. Stefanutti, B. Stein, and J. P. Wolf, "Derivation of Mount Pinatubo stratospheric aerosol mean size distribution by means of a multiwavelength lidar," *Applied Optics* **33**, 5690-5697 (1994).

C. Dellago and H. Horvath, "On the accuracy of the size distribution information obtained from light extinction and scattering measurements I Basic considerations and models," *Journal of Aerosol Science* **24**, 129-142 (1993).

D. P. Donovan and A. I. Carswell, "Principal component analysis applied to multiwavelength lidar aerosol backscatter and extinction measurements," *Applied Optics* **36**, 9406-9424 (1997).

O. Dubovik, B. N. Holben, T. F. Eck, A. Smirnov, Y. J. Kaufman, M. D. King, D. Tanre, and I. Slutsker, "Variability of absorption and optical properties of key aerosol types observed in worldwide locations," *Journal of Atmospheric Sciences* **59**, 590-608 (2002).

F. G. Fernald, B. M. Herman, and J. A. Reagan, "Determination of aerosol height distributions by lidar," *Journal of Applied Meteorology* **11**, 482-489 (1972).

W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (editors), *Markov Chain Monte Carlo in Practice*, (Chapman & Hall, London, 1996).

P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, (SIAM, Philadelphia, 1998).

J. Heinzenberg, H. Müller, H. Quenzel, and E. Thomalla, "Information content of optical data with respect to aerosol properties: numerical studies with a randomized minimization-search-technique inversion algorithm," *Applied Optics* **20**, 1308-1315 (1981).

D. L. Hutt, L. R. Bissonnette, and L. Durand, "Multiple field of view lidar returns from atmospheric aerosols," *Applied Optics* **33**, 2338-2348 (1994).

A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Computing Surveys* **31**, 264-323 (1999).

J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, (Springer-Verlag, New York, 2004).

N. Larsen, B. Knudsen, T.S. Jorgensen, A. Disarra, D. Fua, P. Digirolamo, G. Fiocco, M. Cacciani, J.M. Rosen, N.T. Kjome, "Backscatter measurements of stratospheric aerosols at Thule during January-February 1992." *Geophysical Research Letters* **21**, 1303-1306 (1994).

B. R. Lienert, J. N. Porter, and S. K. Sharma, "Repetitive genetic inversion of optical extinction data," *Applied Optics* **40**, 3476-3482 (2001).

D. A. Ligon, J. B. Gillespie, and P. Pellegrino, "Aerosol properties from spectral extinction and backscatter estimated by an inverse Monte Carlo method," *Applied Optics* **39,** 4402–4410 (2000).

R. B. Miles, W. R. Lempert, and J. N. Forkey, "Laser Rayleigh scattering," *Measurement Science And Technology* **12**, R33-R51 (2001).

H. Müller and H. Quenzel, "Information content of multispectral lidar measurements with respect to the aerosol size distribution," *Applied Optics* **24**, 648-654 (1985).

D. Müller, U. Wandinger, and A. Ansmann, "Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: Theory," *Applied Optics* **38**, 2346-2357 (1999a).

D. Müller, U. Wandinger, and A. Ansmann, "Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: Simulation," *Applied Optics* **38**, 2358-2368 (1999b).

D. Müller, U. Wandinger, D. Althausen, and M. Fiebig, "Comprehensive particle characterization from three-wavelength Raman lidar observations: Case study," *Applied Optics* **40**, 4863-4869 (2001).

N. T. O'Neill, O. Dubovik, and T. F. Eck, "Modified Ångstrom exponent for the characterization of submicrometer aerosols," *Applied Optics* **40**, 2368-2375 (2001).

A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* 3rd Edition, (McGraw-Hill, New York, 1991).

M. D. Post, "A graphical technique for retrieving size distribution parameters from multiple measurements: Visualization and error analysis," *Journal of Atmospheric and Oceanic Technology* **13**, 863–869 (1996).

C. D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*, (World Scientific, Singapore, 2000).

Y. Sasano, E. V. Browell, and S. Ismail, "Error caused by using a constant extinction/backscattering ratio in the lidar solution," *Applied Optics* **24,** 3929–3932 (1985).

B. Stein, M. del Guasta, J. Kolenda, M. Morandi, P. Rairoux, L. Stefanutti, J. P. Wolf, L. Woste, "Stratospheric aerosol size distributions from multispectral lidar measurements at Sodankyla during EASOE," *Geophysical Research Letters* **21,** 1311-1314 (1994).

J. Tamminen, "Validation of nonlinear inverse algorithms with Markov chain Monte Carlo method," Journal of Geophysical Research **109**, D19303 (2004).

I. Veselovskii, A. Kolgotin, V. Griaznov, D. Müller, U. Wandinger, and D. N. Whiteman "Inversion with regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding," *Applied Optics* **41**, 3685-3699 (2002).

I. Veselovskii, A. Kolgotin, V. Griaznov, D. Müller, K. Franke, and D. N. Whiteman "Inversion of multiwavelength Raman lidar data for retrieval of bimodal aerosol size distribution," *Applied Optics* **43**, 1180-1195 (2004).