CHARACTERIZATION OF A GEIGER-MODE AVALANCHE PHOTODIODE DETECTOR FOR HIGH SPECTRAL RESOLUTION LIDAR

By

Ilya Razenkov

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science
(Atmospheric and Oceanic Sciences)

at the

UNIVERSITY OF WISCONSIN-MADISON

2010

Approved by: _____

Steven A. Ackerman Date

Professor of Atmospheric and Oceanic Sciences

Approved by: _____

Edwin W. Eloranta Date

Senior Scientist of Atmospheric and Oceanic Sciences

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Edwin Eloranta for support and patience. He is the person who provided me the opportunity to work with one of the most advanced lidar instruments in the world. I also appreciate all the time he has spent teaching and guiding me through this part of my education.

My thanks also go to Professors Steve Ackerman and Dave Turner for their thoughtful and constructive comments about this thesis to help improve the presentation of this work.

I would also like to thank all the members of the University of Wisconsin Lidar Group who have helped make this work possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

x

ABSTRACT

The studies of cloud properties and modeling of the radiative transfer of the atmosphere requires a knowledge of the optical depth and the scattering cross section profiles of the atmosphere. Lidar is one of the remote-sensing instruments, which provides the required atmospheric profiles. The conventional lidar, based on elastic scattering phenomenon, requires assumptions about the optical properties of the scattering particles in order to solve a lidar equation.

The High Spectral Resolution Lidar (HSRL) avoids the limitation of standard lidars by separating the aerosol and molecular photons in the receiver; the separated molecular signal is used as a reference for calculation of the optical depth and backscatter cross section profiles without a-priori knowledge of the scattering properties of the atmospheric particles.

The eye-safety requirement of modern lidar instruments impose limitations on the transmitted energy density per laser pulse. In order to meet the requirement, the lidar should operate at low transmitted energy in the laser pulse and high pulse repetition rate. However, because of the low transmitted energy in the laser pulse, lidar returns compete with sky noise. The sky noise is suppressed by using a small receiver field-of-view, but requires an accurate maintenance for alignment of the receiver's and transmitter's optical axes. The mechanical stability of transceiver of conventional lidars with separated transmitting and receiving telescopes is the main problem which causes the misalignment of transmitter's and receiver's optical axes.

The HSRL lidar with a beam expansion through the receiving telescope eliminates the problems associated with mechanical stability of transceiver and reduces the energy density in the outgoing beam. A drawback of this design is that some of the transmitted light is scattered from internal optical elements and part of this internally scattered light reaches the detectors. This results in the detector overloading and induces time delayed spurious photon counts that contaminate the measured signal profile and increases signal baseline. This is a so called afterpulsing phenomenon. The other problem associated with peculiarities of avalanche photodiodes (APD), used for detecting the photons in the lidar, is the afterpulsing from the signal photons, which is a significant source of afterpulsing in case of dense clouds, and the pile-up effect, caused by the inherent detector's dead time.

The purpose of this study is to measure the characteristics of the detector to perform a lidar data correction on distorted data profiles. A number of tests with an APD based single photon counting module were made in order to simulate conditions in the real lidar. The tests showed that the baseline light profile caused by internally scattered light could be measured by covering the telescope and then subtracted from the lidar data as range dependent values in order to correct for the data. The impulse response function of the detector was measured by using a low amplitude laser pulse, which was used for the correction on afterpulsing from the signal photons. The test with laser pulse for different light intensities attenuated by a set of calibrated neutral density filters was performed for evaluation of the linear region of afterpulsing and for the measurement of the detector's dead time.

Chapter 1

Introduction

The transfer of a shortwave solar radiation and a longwave radiation through the atmosphere is one of the most important processes for radiation budget. Clouds are important modulators of the radiation and play a key role in the radiative balance of the earth. The studies of the radiative transfer of the atmosphere shows that the presence of clouds produce either a heating or cooling effect depending on the cloud macrophysical and optical properties [1, 2]. A modeling of the radiative transfer requires a knowledge of the aerosol and cloud optical properties. One type of remote-sensing instrument, which provides optical depth and scattering cross section profiles, is a lidar (LIght Detection And Ranging, an optical analog of radar). Lidars are based on a principle of interaction of emitted light with atmosphere. Mie and Rayleigh scattering, differential absorption, Raman scattering, and resonance fluorescence can be used to derive a variety of parameters characterizing the state of the atmosphere such as temperature, humidity, wind speed, aerosols and clouds properties with a high spatial and temporal resolution. These properties makes lidar a valuable instrument for atmospheric scientists.

The first lidars were based on elastic scattering of the emitted laser light (i.e. scattering at the laser wavelength) by the atmosphere (aerosols and molecules). The lidar equation describing the process of elastic scattering of light relates the number of transmitted photons to the number of received photons. It contains a product of two unknowns at each

point: the local backscatter and optical depth (integral of the extinction coefficient, see Section 2.1). The solution of that equation requires an assumption about the medium. That assumption is a fundamental limitation of the classic lidar [3, 4].

The High Spectral Resolution Lidar (HSRL), designed at the University of Wisconsin-Madison, avoids the limitation of standard lidars by discriminating between Mie and a Rayleigh backscattering. The Doppler effect due to the thermal motion of scatterers broadens the spectrum of the light. The spectrum broadening is about 300 times larger for molecules than for the particles responsible for the lidar signal. A molecular iodine filter is used to separate of the molecular and aerosol signal, and the radiosonde temperature profile is used to calculate the molecular density profile for calibration of the lidar. The lidar directly measures the Rayleigh signal and exploits the temperature profile as a calibration target for the calculation of the molecular density of the atmosphere to calculate the optical depth profile and backscatter cross-section profile, and does not require a-priori knowledge of the backscatter cross-section. The HSRL principles are described in the Chapter 2.

A Q-switched laser can produce giant light pulses with a pulse length of a tens of nanoseconds and a laser beam diameter less than a millimeter, thus containing and high energy density in the output beam. Lasers are extremely dangerous for human eyes because of the intense light. The telescope in the lidar transmitter is used to expand the beam to reduce the energy density in the transmitted beam; this also reduces divergence of the laser beam. However, in order to reach an eye-safety threshold usually lidars are operated at high pulse repetition rate and low transmitted pulse energy. The lidar returns becomes comparable to a sky background noise because of the low transmitted energy in the laser pulse. The small field-of-view of the lidar receiver and transmitter reduces the sky noise, but that requires an

accurate maintenance of alignment of receiver's and transmitter's optical axes. Usually, the size of the outgoing beam is several centimeters with a divergence close to the diffraction limit, e.g. for laser beam diameter with a 10 cm diameter, and the divergence is around 10 μrad. In order to reduce a sky background noise, the field-of-view of the receiver is limited either by the size of the laser beam, or is slightly bigger to account the destruction of the laser beam by turbulent effects. Such a narrow field-of-view requires an accurate alignment and high mechanical stability of the transceiver. The larger beam expansion of the transmitted beam through the receiving telescope keeps lidar eye-safe and maintains the transceiver optical alignment.

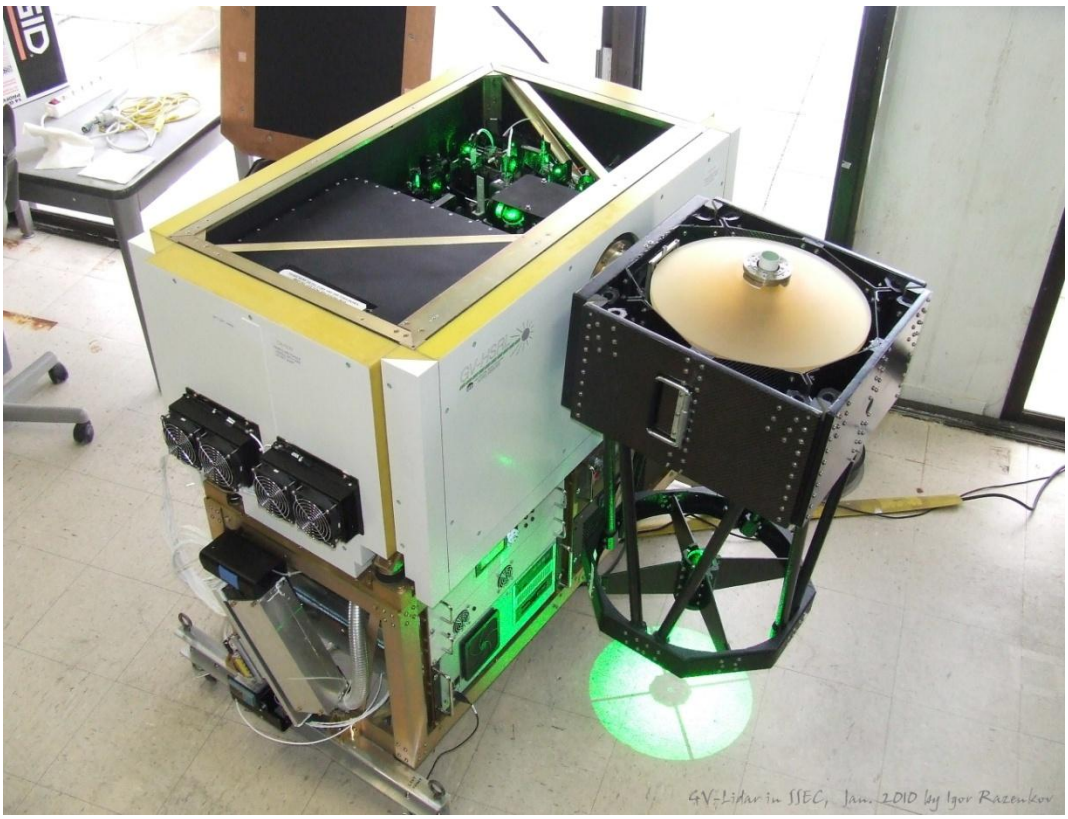

Figure 1. High Spectral Resolution Lidar for NCAR research aircraft Gulf Stream (by Igor Razenkov).

The University of Wisconsin High Spectral Resolution lidar (HSRL) operates through a single telescope to transmit and receive light [5]. This reduces the energy density in the transmitted beam to eye safe levels and improves mechanical stability allowing it to operate with a small field-of-view. However, a drawback of this design is that a portion of the transmitted light pulse is scattered from internal optical elements back to the receiver, which overloads the detectors at the moment of a laser firing and induces spurious signal pulses at a later time. This produces the characteristic decaying tail and contaminates the signal. This is a so called afterpulsing affect. However, the internally scattered light is not the only source of afterpulsing; the atmospheric signal photons can also produce this effect, which can be significant in case of dense clouds.

An example of an aerosol backscatter cross section profiles as a function of altitude versus time is presented in Figure 2. The time interval from 0:00 to 1:30 UT corresponds to a relatively clear atmosphere, and the interval from 1:30 to 4:00 UT shows a cloud at 7 km. However, above the cloud (9 km) the atmosphere remains clear. For the clear atmosphere the backscatter ratio is usually constant with a typical value of 0.2. The average value of the scattering ratio for time intervals from 0:00 to 1:00 UT and from 3:00 to 4:00 UT is shown in Figure 3. The plot shows the influence of the cloud on the data profiles, and, as can be seen, the backscatter ratio increases in the presence of the cloud above 9 km due to the afterpulse effect. The afterpulsing also can be seen as bright lines above the densest part of the cloud in Figure 2 (at 3:20 UTC). The other examples of lidar data and a result of its correction are presented in the Chapter 3.

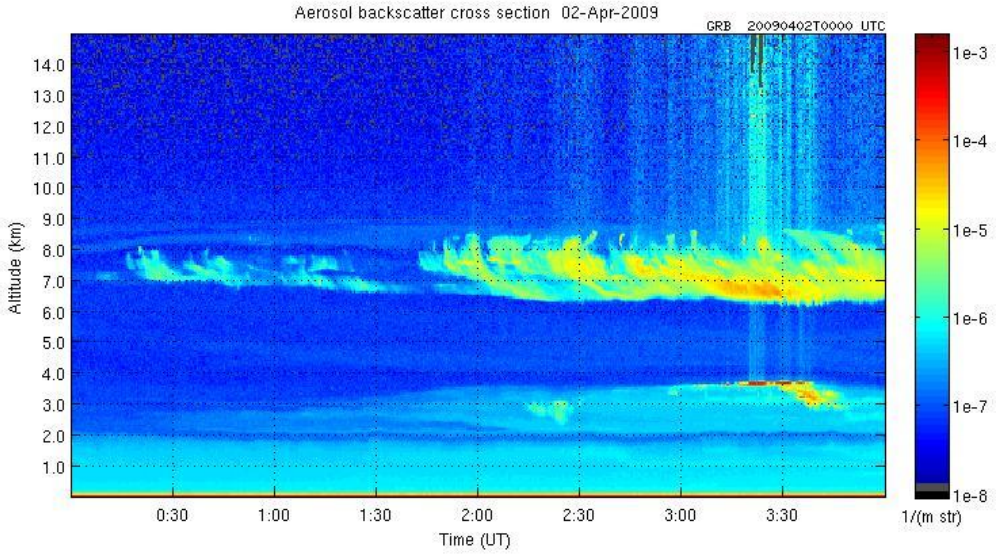

Figure 2. Aerosol backscatter cross section on April 2, 2009, 0:00-4:00 UT, Madison.

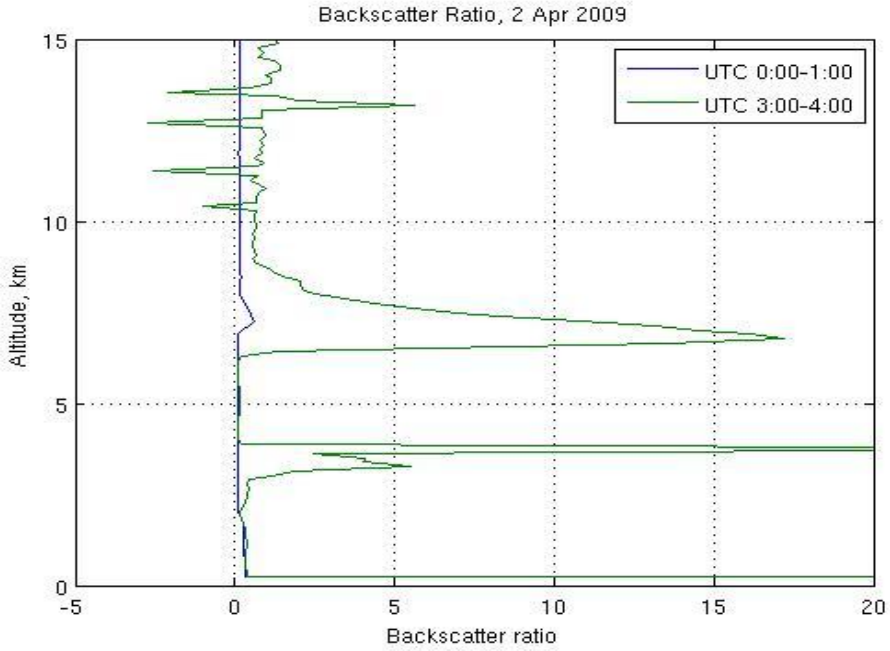

Figure 3. Backscatter ratio averaged over time interval 0:00-1:00 UT(blue line) and 2:00-3:00 UT (green line), April 2, 2009, Madison, WI.

The afterpulsing is a property of a detector and its characteristics depend on many parameters, which are different for different detector types. The receiver's single photon counting modules (SPCM) of the HSRL are based on avalanche photodiodes. Two important

processes occurring in the semiconductor can cause the afterpulsing: (1) the absorption of the photons beneath the p-n junction in neutral region causing a slow drift of the photogenerated electrons towards the junction and a delayed response of the detector; (2) a deep trapping of electrons on the lattice defects from photon triggered avalanche, thus, the afterpulsing is a secondary phenomenon [6].

In order to characterize the detector properties, several tests were conducted with SPCM, which are described in Chapter 4. The results show that the afterpulse signal from internally scattered light can be measured as a function of time (range) separately by covering the telescope and then subtracted from the data as range dependent values, this is called a baseline correction (Section 5.1). The model, which characterizes the detector response, is shown in Section 4.4. A method to correct for the afterpulse contribution from cloud signals is based on measuring the detector impulse response function, which is then deconvolved with the lidar data profile.

Another distortion of data is associated with the inherent detector dead time (a time needed for recovery of photodiode after avalanche breakdown) which causes the pile-up effect in data. Data profiles can be corrected by measuring the detector's dead time and by applying a pile-up correction, which encounters missed photons due to the recovery time (Section 4.4). The examples of lidar data and results of these corrections are presented in Chapter 5 of the thesis.

Chapter 2

Background: Lidar instruments and the HSRL

2.1 Classic Elastic scattering lidar

The conventional lidar is based on the effect of elastic scattering from molecules and aerosols, and the scattered photons are detected by a single channel. The lidar equation of a monostatic classic lidar system can be presented in terms of the number of received photons:

$$\frac{\delta N(r)}{\delta t} = G(r)N_0 \frac{cA}{2r^2}\left(\beta_a(r)\frac{P_a(\pi,r)}{4\pi} + \beta_m(r)\frac{P_m(\pi,r)}{4\pi}\right)e^{-2\tau(r)} + M(r) + b, \qquad (1)$$

where

$\delta N(r)$ = the instantaneous number of photons incident on the receiver from distance *r* per time interval δt, or count rate, $\frac{photons}{s}$;

$G(r)$ = a geometrical factor (overlap function of transmitted laser beam and receiver field-of-view);

N_0 = the number of transmitted photons;

A = area of the receiving antenna, m^2;

c = speed of light, *m/s*;

$\beta_a(r)$, $\beta_m(r)$ = aerosol and molecular volume scattering coefficients, m^{-1};

$\frac{P_a(\pi,r)}{4\pi}$, $\frac{P_m(\pi,r)}{4\pi}$ = aerosol and molecular backscatter, sr^{-1};

$\tau(0,r) = \int_0^r \beta_e(r')dr'$ = optical depth;

$\beta_e(r)$ = volume extinction coefficient, m^{-1};

M = multiple scattering signal,

b = background.

This equation has a product of two unknowns at each point, of the local backscatter and optical depth: $\beta_a(r)\frac{P_a(\pi,r)}{4\pi}e^{-2\tau(r)}$, and, there is not enough information to measure both the atmospheric extinction and backscatter cross-section.

The lidar equation solution with a single scatter approximation proposed by Klett [3] assumes a power law relationship between the extinction and backscatter coefficient:

$$\beta_a(r) = C\beta_a(\pi,r)^k \tag{2}$$

where *C* and *k* depend on the lidar wavelength and the aerosol properties; and backscatter phase function and boundary conditions. However, that relation works properly only if the backscatter phase function is constant with altitude, so that the aerosol size distribution does not change with altitude. That is usually not true in the real atmosphere. The solution is sensitive to the chosen boundary conditions and to the optical depth, and to the value of constant *C* [7-9].

2.2 High Spectral Resolution Lidar Diagram

The High Spectral Resolution Lidar (HSRL) was designed to overcome the limitation of the classic lidar. The instrument transmits a narrow spectral laser pulse to the atmosphere, and the outgoing wavelength is tuned to the iodine absorption line. When the light is scattered on aerosols, its spectrum slightly broadens because of the slow motion of aerosols determined by the wind (~10 m/s) and turbulence (~1m/s), which is ~30 MHz and ~3 MHz, respectively (see Appendix A). The thermal motion of molecules is much more intense (~300

m/s), and the spectrum Doppler broadening of the molecular scattered light produces frequency shifts of the order of ~1 GHz. Thus, the spectrum of the scattered light consists of a narrow spike near the frequency of the laser transmitter caused by particulate scattering laying on a much broader distribution produced by molecular scattering used to separate molecular return. The HSRL uses the iodine absorption cell to reject the return from aerosols in the molecular channel. A separate channel measures a combined return. A molecular density profile calculated from the atmospheric temperature profile is used as a calibration target [10].

A simplified diagram of the HSRL transmitter and receiver is presented in Figure 4. A collimator (Col) expands the laser beam up to 20 mm diameter. Then it passes through a cross polarized pickoff (CPP), a thin film polarizer (TFP), and a quarter-wave plate (QWP) converts initial linear polarization in to a circular polarization. A combination of a thin film polarizer with a quarter wave plate forms an optical transmitter-receiver switch, allowing use of one telescope to transmit and receive light. A second pass of the QWP by a light after scattering by atmosphere converts the circular polarized light back into linear. A double pass of the QWP rotates a polarization of the light by 90^o relative to the transmitted light (parallel component), so that parallel component (not depolarized) is reflected by the TFP, and Cross Polarized Pickoff (1/10 ratio) reflects 10% of cross-polarized component. After that a beam polarizing cube (PC1) combines these components.

The combined light passes the collimator with a field stop (FS) in the foci of the lenses, which defines the field-of-view for all lidar channels. In order to cut off the background noise and make system operational in day time the light is filtered by interference filter (IF) and Fabry-Perot etalon. Then, a second polarizing cube (PC2) directs

the cross component to the cross-channel, and the remaining parallel component is split by a beam splitter BS1 20/80 with 80% of light directed to the molecular channel, where it passes through the iodine absorption filter to remove the aerosol signal, and 20% to the combined channel.

The second beam splitter is used to enhance a dynamic range of the aerosol channel by dividing it into two parts: combined high (an aerosol channel with high sensitivity), and combined low channel (an aerosol channel with low sensitivity). That allows detection of an aerosol signal with a high dynamic range from dense clouds, so that when one of the detectors is saturated another detector is still able to detect the signal avoiding detector pile-up and saturation.

Note, that the actual optical diagram of the lidar is 3-dimensional, and the transmitted beam, and beams reflected by the thin film polarizer and cross polarized pickoff are mutually orthogonal.

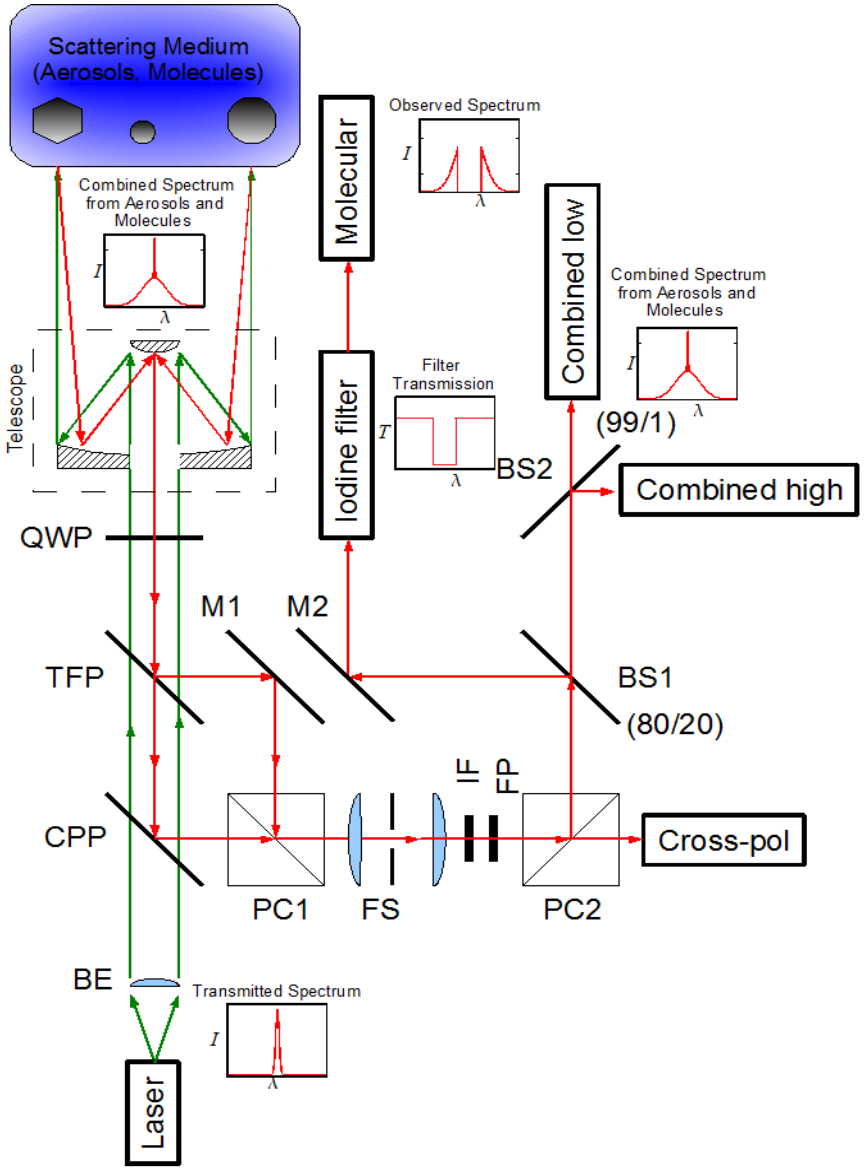

Figure 4. A simplified optical diagram of the HSRL transmitter and receiver with light spectrum. A laser beam is expanded up to 20 mm diameter by beam expanding telescope (BE). A combination of a Thin Film Polarizer (TFP) with a Quarter Wave Plate (QWP) forms an optical transmitter-receiver switch. The narrow spectrum laser pulse is transmitted into the atmosphere and the Doppler broadened backscattered signal is returned into the same telescope. A double passing of the QWP rotates a polarization of the light by 90^{o}, so that parallel light component (undepolarized) is reflected by the TFP, and Cross Polarized Pickoff (1/10 ratio) reflects 10% of cross-polarized component, and then these beams are combined by a polarizing cube (PC1). The combined light passes the collimator (C) with a field stop (FS) in the foci of the lenses. A PC2 directs cross component to the cross-channel, and the remaining parallel component is split by a beam splitter BS1 20/80 with 80% to the molecular channel, where it passes through the iodine absorption filter to remove the aerosol signal; and 20% to the combined channel. The 1/100 BS2 enhances the dynamic range of the combined channel.

2.3 High Spectral Resolution Lidar principles

Two lidar equations for separated an aerosol and a molecular signals can be written for the "HSRL". For the aerosol signal:

$$\frac{\delta N_a(r)}{\delta t} = G(r)N_0 \frac{cA}{2r^2}\beta_a(r)\frac{P_a(\pi,r)}{4\pi}e^{-2\tau(r)}. \quad (3)$$

For the molecular signal:

$$\frac{\delta N_m(r)}{\delta t} = G(r)N_0 \frac{cA}{2r^2}\beta_m(r)\frac{P_m(\pi,r)}{4\pi}e^{-2\tau(r)}, \quad (4)$$

where the molecular backscatter phase function is $\frac{P_m(\pi,r)}{4\pi} = \frac{3}{8\pi}$ from the Rayleigh scattering theory. Note that a scattering ratio of the signals does not depend on an overlap function or the optical depth:

$$SR(r) = \frac{\delta N_a(r)}{\delta N_m(r)}. \quad (5)$$

An atmospheric temperature profile is used to calculate an atmospheric density profile and a molecular scattering cross section from the Rayleigh theory. The molecular scattering cross section per unit volume is

$$\beta_m(r) = N_m(r)\frac{d\sigma_R(\pi)}{d\Omega}, \quad (6)$$

where $N_m(r)$ – a concentration of molecules. σ_R – a scattering cross section. For a mixture of atmospheric gases below 100 km altitude [11,12]

$$\frac{d\sigma_R(\pi)}{d\Omega} = 5.45\left[\frac{\lambda(\mu m)}{0.55}\right]^{-4}10^{-28}cm^2sr^{-1}. \quad (7)$$

The number of gas molecules per unit volume can be calculated from the ideal gas law using atmospheric pressure $P(r)$ and temperature $T(r)$ as:

$$N_m(r) = \frac{P(r)N_A}{T(r)R_a}, \quad (8)$$

where the N_A is the Avogadro constant, and the R_a is the universal gas constant. Thus, the molecular scattering coefficient is proportional to the atmospheric pressure and temperature:

$$\beta_m(r) = C_{air}\frac{P(r)}{T(r)}, \quad (9)$$

where $C_{air} = 3.786 \cdot 10^{-6}\frac{K}{hPa\ m}$ at 532 nm.

Using a lidar equation with no extinction in the medium ($\tau = 0$) and expressions for number of molecules and molecular scattering coefficient, we can calculate a theoretical molecular signal in the absence of attenuation:

$$\delta M_T(r) = N_o A\frac{cA}{2r^2}\beta_m(r)\frac{3}{8\pi}, \quad (10)$$

where N_o is a number of transmitted photons. Thus, taking the logarithm of the ratio of the theoretical to the measured molecular photons we derive the optical depth between ranges r_1 and r_2:

$$\tau(r_1, r_2) = \frac{1}{2}log_e\frac{\delta M_T(r_1)G(r_2)\delta N_m(r_2)}{\delta M_T(r_2)G(r_1)\delta N_m(r_1)} = \frac{1}{2}log_e\frac{{r_2}^2\delta N_m(r_2)\beta_m(r_1)}{{r_1}^2\delta N_m(r_1)\beta_m(r_2)}, \quad (11)$$

$G(r)$ = a geometrical factor (overlap function of transmitted laser beam and receiver field-of-view).

The average value of total extinction coefficient is

$$\overline{\beta_e\ (r)} = \frac{\partial\overline{\tau(r)}}{\partial r} = \frac{\tau(r_2)-\tau(r_1)}{r_2-r_1}. \quad (12)$$

The aerosol backscatter coefficient can be calculated by using the lidar backscatter ratio and the molecular backscatter coefficient:

$$\beta_a(r)\frac{P_a(\pi,r)}{4\pi} = SR(r)\beta_m(r)\frac{3}{8\pi}. \quad (13)$$

The examples of data and derived values of the backscatter ratio, optical depth and backscatter coefficient are presented in Chapter 5.

2.4 Detection and calibration.

The HSRL diagram, presented in Figure 4, shows optical separation of the received light into four channels. The cross-polarized channel detects only the cross-polarized component of depolarized light scattered by aerosols and molecules. Signals detected by two combined channels are merged together to a combined signal S'_c which contains only vertically polarized backscattered light (the same polarization as a transmitted light) from aerosols and molecules. The molecular iodine filter in the molecular channel absorbs most of the backscattered light from aerosols and some of it from molecules letting the rest of the molecular signal pass through the filter (S'_m). Thus, the signals detected in the combined and molecular channels can be described as a linear combination of number of photons scattered by aerosols (N_a) and molecules (N_m) incident to the lidar receiver:

$$S'_c = C'_{ac} N_a + C'_{mc} N_m, \tag{14}$$

$$S'_m = C'_{am} N_a + C'_{mm} N_m, \tag{15}$$

where C'_{ac} and C'_{mc} - are relative contributions of aerosol and molecular photons to the combined channel correspondingly; C'_{am} describes a relative contribution of the aerosol photons to the molecular channel due to leakage of the filter; C'_{mm} represents the transmission of molecular photons through the iodine absorption filter. As long as the deriving quantities are in the form of ratio, the above two equations can be normalized by the coefficient C'_{ac}. These coefficients include corresponding channel efficiencies and represent

gains of the channels for aerosol and molecular photons. The inverted equations for the relative number of aerosol and molecular photons incident on the system relative to the gain of aerosol photons combined channel are:

$$N_m = \frac{S_m - C_{am} S_c}{C_{mm} - C_{am} C_{mc}}, \tag{16}$$

$$N_a = \frac{C_{mm} S_c - C_{mc} S_m}{C_{mm} - C_{am} C_{mc}}. \tag{17}$$

The coefficients C_{am}, C_{mm}, C_{mc} are determined from lidar calibration procedure described by Eloranta [10].

The laser is tuned to a peak of the iodine absorption line. The iodine filter rejection efficiency to aerosol photons is less than 100%, which allows a leakage of aerosol photons through the cell. The leakage of the iodine cell to aerosol backscattered photons is measured at the absorption peak of the iodine cell as a ratio of two calibration signals in the channels:

$$C_{am} = \frac{S_m(\lambda_I)}{S_c(\lambda_I)}, \tag{18}$$

where

$S_m(\lambda_I)$ = molecular channel signal at the iodine absorption peak when exposed to the laser spectrum;

$S_c(\lambda_I)$ = combined channel signal at the iodine absorption peak when exposed to the laser spectrum.

The spectral distribution of the aerosol backscatter can be assumed to be similar to the spectral distribution of the transmitted laser light, since the Doppler-broadening of the aerosol backscatter is negligible. Hence, the calibration signals can be presented as a convolution between laser spectral distribution and a transmission spectrum of each channel.

The transmission of molecular photons through the iodine absorption filter of the molecular channel (C_{mm}) is calculated by convolving the molecular (S_m) channel signal scan (i.e. measured filter function) with calculated from an atmospheric temperature profile molecular spectrum M_{spect} at lock point λ_I:

$$C_{mm} = \sum_{n=N_o}^{N_L} S_m(\lambda_n) M_{spect}\, \delta\lambda, \tag{19}$$

where

N_o, N_L = Starting and the Ending point of the calibration scan;

λ = wavelength;

$\delta\lambda$ = the wavelength difference between two points in the calibration scan;

Similar, the transmission of molecular photons through the combined channel (C_{mc}) is calculated by convolving the combined (S_c) channel signal scan with molecular spectrum M_{spect} at lock point λ_I:

$$C_{mc} = \sum_{n=N_o}^{N_L} S_c(\lambda_n) M_{spect}\, \delta\lambda. \tag{20}$$

The C_{mm} and C_{mc} are pressure and temperature dependent, they are computed as functions of altitude.

The HSRL type lidar has a very robust calibration and overcomes the limitation of conventional lidars, and does not require a-priori assumptions in deriving backscatter coefficient and optical depth.

2.5 Internally Scattered Light and Afterpulsing

The operation wavelength and the output beam diameter of lidar allows formation of a beam with a divergence

$$\theta_L = 1.22\,\frac{\lambda}{D} = 1.22\,\frac{532\ nm}{0.4\ m} = 1.6\,\mu rad, \tag{21}$$

limited by diffraction of light on the telescope aperture. However, atmospheric turbulence and forward scattering broadens the beam width in the atmosphere, and in order to detect more backscattered light, the field-of-view of the receiver is set 100 μrad. This small field-of-view requires an accurate alignment and mechanical stability of lidar with two telescopes for transmitter and receiver, otherwise, a small temperature gradient of the thermally expanding parts can cause a misalignment of a transmitter's and a receiver's optical axes. That problem for the HSRL lidar at University of Wisconsin–Madison was solved by using one telescope to transmit and receive light (see Figure 4). However, this design causes a saturation of detectors at the moment of laser firing, because some of the transmitted light is scattered from internal optical elements and part of this internally scattered light reaches the detectors (internally scattered light).

Since the light incidence angle on the secondary mirror is close to the normal, a significant part of the stray light is the light scattered on the secondary mirror of the telescope including a mirror block (about 50%). The center of the secondary mirror is blocked with a small mirror block with a reflecting surface tilted 45^o relative to the beam, which reflects a central part of the laser beam to a trap. That prevents a specular reflection of transmitted light to the receiver. The quality of that block also affects the amount of scattered light, which is around 25% of the total amount of the scattered light. The other 50% of light is the light

scattered from other optical elements of the transmitter. An influence of the internally scattered light on the signals from the atmosphere was investigated by performing a test with two sources of light, a laser light and a light from light emitting diode described in Chapter 4.

The internally scattered light increases a baseline in data profiles caused by the afterpulse effect. That is produced by trapped electrons in the detector's p-n junction (afterpulse effect); and it is the main, but not the only source of afterpulsing in the detected lidar signals. The origin of afterpulsing in avalanche photodiodes is discussed in Chapter 3. The largest amount of the internally scattered light in the lidar system exists in the combined high sensitivity channel with approximately $3 \cdot 10^5$ photons per 50 ns time interval (bin) per laser shot (photons/bin/shot). The combined low sensitivity channel has around $3 \cdot 10^3$ photons/bin/shot of the scattered light.

The cross polarized channel has $7 \cdot 10^4$ photons/bin/shot. The cross polarized signal is very sensitive to the accuracy of correction and, in particular, to the internally scattered light correction (baseline correction). Even though the baseline signal is small, the error is large because the signal must be divided by the cross-pol beam splitter ratio. For that reason, it is ten times more sensitive relative to the aerosol and molecular signals.

The beam splitter 1 (BS1) directs 80% of the internally scattered light to the molecular channel. An iodine absorption filter absorbs most of the scattered light, thus, the total number of photons in the molecular channel is around 1000 photons/bin/shot.

At a time of laser firing the magnitude of internally scattered light in the various channels are up to four orders of magnitude higher than the values of the highest signals from the atmosphere, and the afterpulsing from that light is larger than the afterpulse contribution from atmospheric photons.

Several range dependent effects such as a background correction of a molecular signal and overlap function of the lidar complicates the afterpulse correction.

Chapter 3

Detector Theory

3.1 Overview

In photon counting experiments the detector plays a key role, and the quality of data significantly depends on its characteristics. The process of detecting signals of any origin is usually accompanied by distortions due to a background noise and a detector nonideality. Distortions of a signal wave-form introduced by a detector are a function of numerous effects including nonlinearity of the detector response, afterpulsing, and internal detector noise, spectral, temporal, and spatial characteristics of device. One of the ways to eliminate these problems is to work in linear region of device characteristics. However, this can significantly restrict dynamic range and, as a result, limit the potential of instrument.

For a long time a photomultiplier tube (PMT) was the most widely used light detection device. Some modern PMTs have a gain of 10^8. PMT consist of a photocathode, a system of dynodes, and an anode under a high voltage. When the incoming photon hits the photocathode it knocks out an electron. This electron hits a series of dynodes inducing an avalanche of electrons from each dynode as it travels under the force of electric field to the cathode, and knocking out more electrons when hitting it. Even though PMTs have a relatively wide linear operation region and have a short dead time (recovery time), it still has

a small quantum efficiency (typically around 10% for 532nm) relative to the quantum efficiency of avalanche photodiodes (APD).

Recent developments provide avalanche photodiodes with quantum efficiency of 65% in visible part of the electromagnetic spectrum. The other advantages of APD are small size, small energy consumption and high robustness which allow compact devices to be designed. Both of these detectors have a common drawback of an afterpulsing effect. A probability of appearance of detector output pulse after the pulse from detected photon is called afterpulsing probability, and the phenomenon is called afterpulsing (effect) [13, 14]. The afterpulsing in the PMTs is well investigated [15, 16]. Unlike PMTs', afterpulsing phenomenon in APDs is poorly investigated, because of the complex dependency of detector properties from operational conditions. For example, the stability of the APD (gain) is very sensitive to the applied external voltage and to operation temperature. Moreover, every semiconductor device is unique and it possesses unique properties caused by degree of purity of a crystal and other technological factors [17-19].

An approach for afterpulse signal correction is based on the experimental measurements of an afterpulse probability distribution, defined as a detector *impulse response function***.** The distribution is deconvolved with a detected signal wave-form containing an afterpulse signal in order to eliminate the afterpulse effect. Another type of distortion is caused by detector's inherent dead time (a time needed for recovery of the detector), and is called a pile-up effect, which is also common for both detector types. This distortion can be minimized through measuring a detector dead time and by calculating a correction coefficient for a given count rate [4, 13, 14].

3.2 Operation principles of an avalanche photodiode

A photodiode is a type of semiconductor device (photodetector) that converts light into a current. The operation principle of the photodiode is based on a photoelectric effect: when a photon hits a p-n junction it excites a negatively charged electron, and a positively charged electron hole (photogenerated carriers). In order to excite that pair of charges, the photon has to possess a sufficient amount of energy, which has to be more than a band gap (or energy gap) of a semiconductor: it is the amount of energy required to free an outer shell electron from its orbit around a nucleus to a free state. If the photodiode is operated slightly above a breakdown voltage, a single photon (or a single dark current electron) can induce a significant avalanche of electrons. This operation regime is called a single-photon avalanche mode, and the device is called a Single-Photon Avalanche Diode (SPAD) or a Geiger-mode detector. In this mode the diode is used as a trigger device which allows detection of low intensity light (down to the single photon) [20, 21].

A strong electric field is induced in a semiconductor when a high bias voltage is applied to a p-n junction. In that condition a single charge carrier injected in the depletion layer can trigger a self-sustaining avalanche breakdown producing a rapid current rise flowing through the diode (picosecond rise-time). The strength of the electric field depends on applied voltage and electrical conductance of the semiconductor material. The threshold of the electric field, required to induce an avalanche breakdown, greatly depends on the semiconductor's material.

In avalanche photodiodes incident light generates free electrons. If that happens in the region with an electric field the electron obtains kinetic energy from the electric field. As the electrons move in the crystal, they strike the lattice. They get absorbed by the atoms and the process halt if their kinetic energy is not sufficient for avalanche breakdown. However, if the necessary energy has been achieved, then the electron will knock out another free electron from the atom and will ionize it producing a hole. This process is called impact ionization. Then both these two electrons are accelerated by an electric field and strike other atoms and produce additional electrons in the same way.

The ionized atoms (or holes) are moving in the direction opposite to electrons' motion. If field strength is high enough, then holes can also initiate an avalanche multiplication, producing a secondary electron avalanche. In reality, holes are not moving. Instead electrons are moving by "jumping" from one atom of a lattice to another, and the path of these electrons is much shorter than that of free electrons. Therefore, in order to initiate an avalanche multiplication of holes, it requires a stronger electric field than for electrons [22, 23].

A number of free electrons in the material increases exponentially as the electrons moves through the material, producing a flow of a large current, which reaches its maximum value within a few picoseconds. An external electric circuit quenches the diode by lowering the bias voltage below the breakdown voltage and drops the current. In order to be able to detect another photon the idling state of the diode have to be recovered by raising the bias voltage again above breakdown. Active current quenching quickly limits the current and let the diode to recover quickly. For SPADs it is necessary to provide a sufficiently low dark count rate. The intensity of light obtained in the single photon counting mode is proportional

to the number of output pulses within a measurement time slot. The temporal distribution of the output pulses gives a waveform of the signal [22, 23].

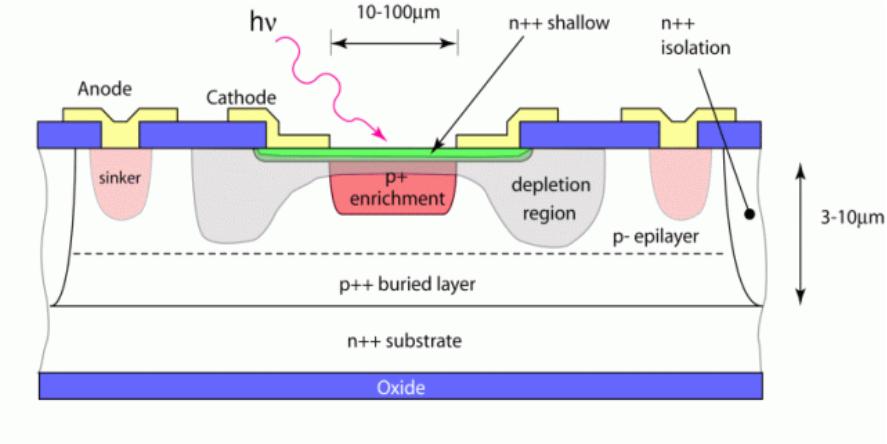

Figure 5. Thin SPAD cross-section.

An example of a single photon avalanche diode is presented in the Figure 5. A photon incident on the sensitive area (p-n junction in the figure colored with a green and red colors) gets absorbed in the p-n junction to generate a primary carriers. The efficiency of the detector is defined how well incident photons are absorbed in this region. The applied electric field across the entire structure accelerates primary carriers and induces the avalanche below the depletion layer (multiplication region).

3.3 Review

In photon counting experiments using detectors that are operated in the Geiger mode, genuine output pulses may be followed by an afterpulse (afterpulse effect). The origin of the afterpulse phenomena and its characteristics depend on the detector type. For a photomultiplier the most frequent causes for afterpulsing are ionized atoms of the residual

gas that are accelerated towards the photocathode and generate delayed photoelectrons. Other causes include fluorescent effects of dynodes and luminescence of the residual gas [15].

An afterpulsing effect in a photomultiplier tube (PMT) was investigated in [15]; the afterpulsing in a PMT is caused by the ionization process of residual gases by electrons. These ions are heavier than electrons, and under the electric field force in the PMT these ions move to the anode slower than the electrons to the cathode, thus, inducing afterpulses (i.e. the pulses induced by the gas ions come after a few microseconds later than the pulses from the electrons). The PMT contains a mixture of gases, which typically are the ions He+ , O2+, H2+, H+. As long as the ions have different weights, each type of the ion produces afterpulses at different times. In order to estimate the ions arrival time, the electric field for a given configuration of the dynodes and a cathode was calculated. For experimental validation they used a PMT and an acquisition system with time resolution less than 2 ns based on registration of only correlated pulses, and the pulse from the photoelectrons was a triggering pulse. They also showed that PMTs are exposed to aging and for 10 year old tube an afterpulse effect was 5 times higher.

The afterpulse effect in an avalanche photodiode (APD) is a different origin than in PMT. Instead of the spikes from arrived ions in the impulse response function (or response function), the response function of an APD has a slowly decaying tail. That behavior of detector response function can be explained by two processes: a delayed detector response (primary pulses) and an afterpulsing (a secondary phenomenon) [18].

The delayed detector response in semiconductor devices is caused by photons that are not directly absorbed in the thin active junction of the SPAD but in the neutral regions beneath the depletion layer (junction). The photon generated carriers slowly diffuse towards

the active region, triggering with a certain probability a delayed avalanche. The resulting slow tail in the detector response function depends on the geometry of the device and on the locations of charge carrier generation regions. As long as the optical absorption coefficient has strong wavelength dependence, the devices with deep neutral layers have a wavelength-dependent diffusion tail [24, 25].

The afterpulsing is a secondary phenomenon and it is correlated to an initial output pulse. In semiconductor avalanche photodiodes, a photoelectron is produced by an incident photon absorbed in the p-n junction. That photoelectron initiates a chain of ionizations that causes a breakdown pulse at the detector output. The raising edge of the current pulse is synchronous to the photon absorption with very small jitter, down to 20 ps. However, some of the generated charge carriers are temporarily trapped on the lattice defects [17-19,25]. Any material defect of the multiplicative areas of the APD may be centers that capture current carriers. The carriers caught in the junction depletion layer in the trap levels are subsequently released with an exponentially decaying probability. When the carriers are released by thermal excitation, new free carriers are created that can lead to afterpulses which are correlated with the initial event.

The probability of afterpulsing depends on many different parameters such as a purity of the semiconductor crystal, an operational temperature, and a breakdown voltage. The breakdown voltage must be uniform over the entire p-n junction [18, 25]; hence, this requires the absence of material defects and a temperature stabilization, which is also can cause an increase of a dark-count rate [13]. Thermal generation of carriers and trapping phenomena in the semiconductor also contribute to avalanches. Thermally generated avalanches cause a constant background that can be separately measured and then subtracted from the signal.

Moreover, an afterpulse itself may produce subsequent afterpulses. The avalanches, which are triggered by the release of trapped electrons, populate trap centers again; therefore, a "second generation" of afterpulses is present, which may cause a third generation and so on [19, 24].

The delayed detector response, the afterpulsing from the trapped charge carriers, and "second generation" afterpulses contribute to the detector response function producing a slowly decaying tail. As long as these three phenomena introduce distortion in a detected signal, we do not distinguish them and call them afterpulse effect.

The dead time effect is caused by the fundamental limitations of the semiconductors. In the idle state the p-n junction of the avalanche photodiode has a high applied voltage and the photon coming to the junction produce a photoelectron which induces an avalanche of the electrons. The time needed for recovery of the p-n junction is characterized by the detector's dead time. Non-paralyzable detectors "ignore" photons arriving in time period less than the dead time interval since the previously detected photon. In non-paralyzable detectors, after detecting a photon any other arriving photon is ignored and does not increase the overall dead time. Thus, two photons will be detected only if they are separated in time more than a dead time. For the system with a dead time t_d and for the measured count rate N_m, the actual count rate N_a is determined by:

$$N_a = \frac{N_m}{1 - N_m\, t_d}, \tag{22}$$

where the term represents the total fraction of the uncounted photons, so the rate at which event is lost is $N_a\, N_m t_d$ [4, 25]. The detector dead time also can affect the afterpulse: the shorter the dead time, the more visible the afterpulsing effect becomes.

Chapter 4

Measurement of Geiger-mode APD Characteristics

4.1 Experimental setup

A series of tests with a single photon counting module were performed in order to characterize the detector; the relative contribution of the afterpulse effect and the pile-up effect to the signal were investigated to develop lidar signal corrections. A block diagram of the experimental setup is illustrated in Figure 6. A single photon counting module from Perkin Elmer model SPCM-AQR-12 which has a dark count rate of 500 Counts/second and a dead time value of 53ns was employed for the experiments [26].

The detector is illuminated with a short laser pulse and with a light pulse from a light emitting diode (LED). The Nd:YAG second harmonics (532 nm) laser is used as a source of short high intensity light pulses. An intense green LED (HLMP-CM15) with a peak wavelength of 524 nm is used to generate long rectangular light pulses.

In the experiments, the LED is driven by an electric pulse generator HP 8082A. The generator, along with a data acquisition board, is triggered by the electric pulse from a laser Q-switch (laser trigger) at 4 kHz repetition rate. The laser light is guided by a multimode optical fiber to the module containing the LED and neutral density filters.

The expanding beam of laser light from the fiber and the LED light passes through two converging lenses, where the first lens collimates the light and the second one focuses it

into another optical fiber, which couples the light from both sources and the detector. In a gap between the two lenses a wheel with a set of calibrated neutral density filters is installed, which attenuates the light to a suitable intensity level. The light intensity of LED pulse relative to the laser pulse, delay between them, and LED pulse length are adjusted by the generator. The background noise is well suppressed and is around the detector dark count rate of 500 Hz.

The acquisition system frequency is 20 MHz which corresponds to the 50 ns accumulating time interval between sampling points denoted as bins.

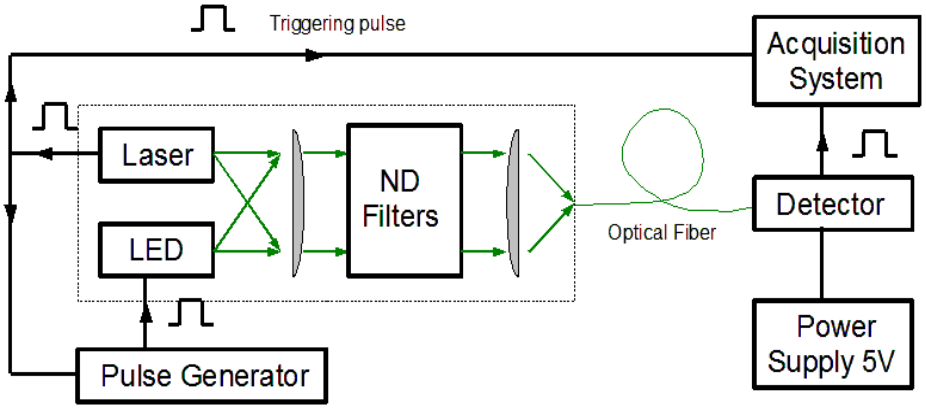

Figure 6. A block-diagram of the experimental setup.

4.2 Impulse response function.

In the first test the detector was illuminated by a short laser pulse containing a small number of incident photons in order to measure a detector impulse response function. A full width half maximum (FWHM) of the pulse is 35 ns and a peak value is 0.1 photons/bin/shot. As long as the bin width is 50 ns, the largest part of the laser pulse in time-space is contained in one bin, and, therefore, the laser pulse light with low intensity is considered as a single photon source and a response of the detector represents it's impulse response function. The

detector impulse response function contains information about the effects distorting the signals such as the laser pulse shape, jitter of the laser pulse along with a clock board of the acquisition system, and afterpulse effect. The response function gives a total probability distribution of all of the listed effects. The signal recorded by the acquisition system (the detector output) is a convolution of the incident light pulse with a detector impulse response function:

$$S_m(t) = \int_t^{t+dt} S(t')h(t-t')dt', \tag{23}$$

where S_m - is a measured signal, S - a received signal, $h(t)$ – normalized by the total number of incident photons detector impulse response function. Thus, the total number of incident photons is the number of photons that would be detected in the pulse by an ideal detector with the same quantum efficiency as the actual detector.

The detector impulse response function normalized by the total number of received photons represents a probability distribution function of the signal distorting effects. If we assume that the signal distorting effects are linear processes and are proportional to the amplitude of the signal (or the number of incident photons), then the detector response function could be deconvolved with signals in order to eliminate the distortions. A detector response function is shown in Figure 7.

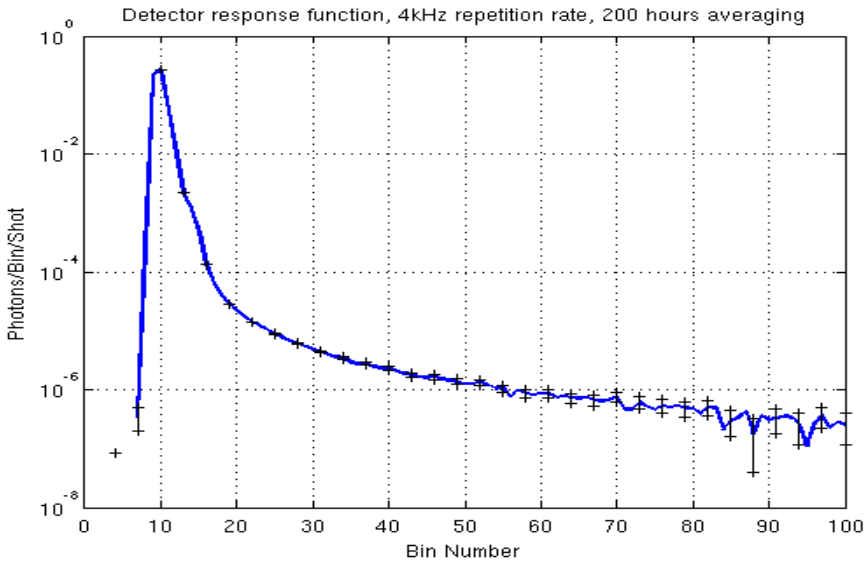

Figure 7. A Detector response function $h(t)$ (1 Bin = 50 ns).

The measurements of the detector impulse response function were performed for relatively low number of incident photons in the pulse with a peak value of 0.1 Photons/bin/shot. The detector response function was measured in a slightly nonlinear region in order to obtain a better statistics, which is proportional to $\frac{1}{\sqrt{N}}$, where N is a number of measurements. For the presented response function, the data were averaged over 70 hours.

The normalized detector response function is:

$$\int_o^\infty h(t')dt' = 1. \tag{24}$$

4.3 Dead-time estimation

The detector response is linear when it is illuminated with a pulse containig a small number of incident photons. Thus, the number of detected photons is proportional to the number of incident photons. As the number of photons increases, a pulse pile-up effect decreases the number of detected photons. This limits the detector to counting ~1 photon per

dead-time interval. The detector's dead-time was measured using a short laser pulse and LED pulse with a length of 0.5 μs and the results are presented in Figure 8. The measured values (number of counts) N_m were approximated by the function:

$$N_p = \frac{N_m}{1 - N_m\ t_d}, \tag{25}$$

where N_p is the number of incident photons. The equation 25 was used for the pile-up correction of the non-paralyzable detector [4, 25]. The best fit for the measured data was found to be for the dead-time 50.4 ns using the least square method.

The main source of errors in the dead time estimate are caused by uncertainty in optical density of filters used to attenuate light, which is $\pm$4% of the filter's optical density (OD) (specified by a manufacturer). The error for the largest filter (OD = 2.5) produces the error of 26% in the signal measurements while for the smallest (OD = 0.04) filter error is around 1.5% of the signal. The blue error bars in Figure 8 represent filter errors.

For the approximation of the measured data for each of the measured point was given a weight proportional to the uncertainty due to the filter error and the points were normalized by the attenuation factor (10^{OD}). It was assumed that the value of the smallest signal, which corresponds to the highest OD filter, is a true value, and the filter error was ignored. Therefore, some of the measured points (blue curve in Figure 8) lies above the line $y = x$ and the theoretical pile-up curve (green curve), because the error of the largest filter was assumed to be zero. The green and blue curves almost coincide on the plot.

The theoretical pile-up curve and the line $y = x$ were calculated using a smallest signal (which corresponds to the highest filter) as a reference point with nonlinearity coefficient of 1, because detector's nonlinearity weakly affects this small signal. The error of

a theoretical pile-up curve is defined by the error for the smallest signal value (largest filter) multiplied by the error of the filter.

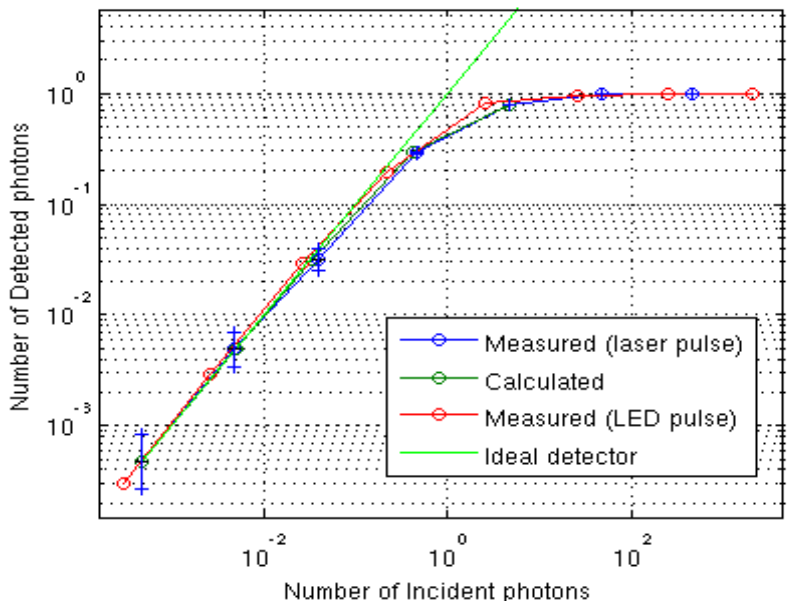

Figure 8. The pile-up curve for short laser pulse (30 ns) and long step function light pulse (0.5 μs), and calculated for 50.4 ns dead-time.

The measurement of the laser pulse with a limited time resolution, which is 50 ns for the current system, averages the values over this time interval and the measured peak value is smaller than the actual value. As the number of incident photons increases the raising edge of the laser pulse may trigger the detector. This causes the uncertainty in the number of incident photons and affects the dead time measurement. It is obvious that the dead time measurement by using LED avoids this problem for moderate pulse length (the length of several bins). The dead time measured with LED is 50.4 ns versus 53 ns for the laser pulse averaged over a 3 bins.

The sensitivity of the pile-up correction to different detector dead time values from 48 ns to 53 ns were estimated by calculating a correction factors for the photons count rate from

.01 to 1 Counted Photon/Bin (photon count rate) using equation 25. The deviation of the corrected count rate for different dead time values from that for the 48 ns dead time as a function of count rate is shown in Figure 9. Separate curves represent a deviation for a certain dead time value, which is shown in the legend. The curves from the bottom to the top are calculated for 49 to 53 ns. The plot shows that the deviation value is less than 1 % for the count rate of 0.1 Counted Photons/Bin for the dead time difference of 5 ns and 2% for the difference of 1 ns at the same count rate.

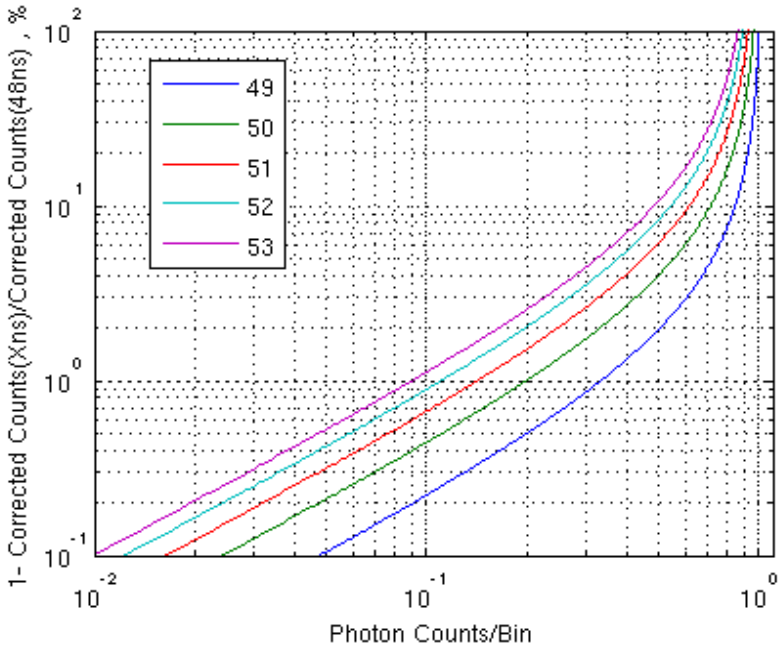

Figure 9. The deviation of the pile-up corrected Photon Count Rate at a certain dead time (49 – 53 ns) value from the Photon Count Rate at 48 ns dead time.

4.4 Characterization of a single photon counting module under overload conditions

4.4.1 Response on a short light pulse

The test for different numbers of incident photons in the laser pulse (light intensities), attenuated by a set of calibrated neutral density filters, was performed in order to determine a linear range for signal contaminating (distorting) effects and to estimate the dead-time for the detector. In the experiments, the detector was illuminated with 35 ns light pulses. The tests were started from the light pulses containing 0.02 detected photons.(the dark count level was 10^{-5} photons per 50 ns) and were increased up to $2 \cdot 10^4$ photons per pulse with an increment factor of 10. The detected signals were normalized by the number of incident photons (see Figure 10). The number of incident photons was calculated as a sum of three bins using the largest values from the signals (bin numbers 49-51). This sum was then multiplied by the attenuation of a neutral density filter. In order to avoid the pile-up effect, the test for the smallest number of incident photons was used (highest neutral density filter) $\sum_{k=49}^{51} S_k(OD_1)10^{OD_1-OD_j}$ (denominator):

$$S'_i(OD_j) = \frac{S_i(OD_j)}{\sum_{k=49}^{51} S_k(OD_1)10^{OD\,1-OD\,j}}, \tag{26}$$

where $S'_i(OD_j)$ is the normalized signal in the *i*th point (bin) of the profile (time moment) for *j*th test (with a *j*th filter). Thus, the total number of incident photons is the number of photons that would be detected in the pulse by an ideal detector with the same quantum efficiency as the actual detector.

Figure 10 shows the photon count normalized by the number of incident photons as a function of time. The background noise is subtracted from the signals, which is calculated by averaging the signals from bin number 3800 to 4000 (from 190 μs to 200 μs). The upper four curves in Figure 10 are averaged over 40 hours; the noise of the signals is caused by the low light pulse intensity (small number of incident photons). The signal spike at bin number 50 corresponds to a peak of the light pulse. The "plateau" seen in signals in Figure 10 shows the detector saturation.

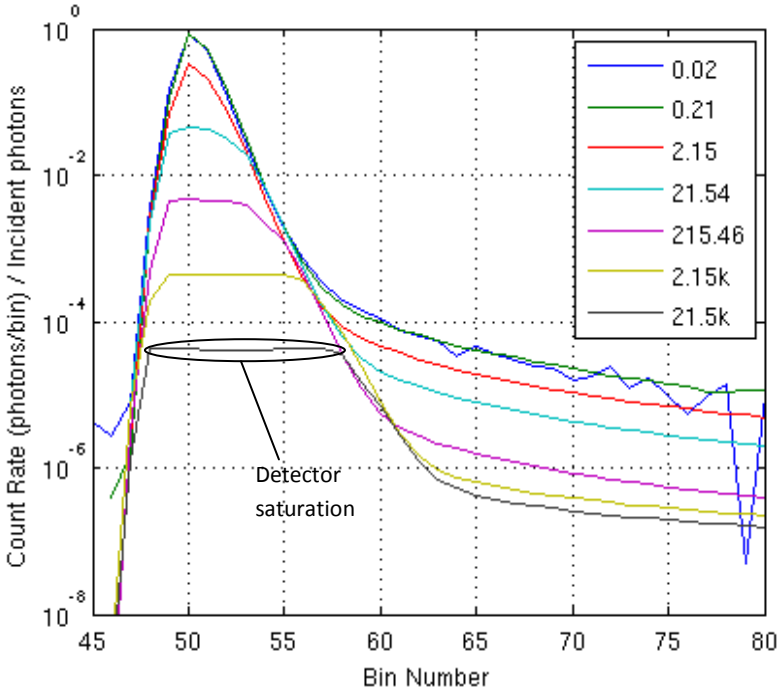

Figure 10. The detected signals normalized by the number of incident photons vs. time (1Bin = 50ns). Legend shows the total number of incident photons.

Figure 11 shows normalized photon counts of one bin as a function of the number of incident photons in the illuminating pulse. Different curves correspond to different bins. The upper curve in Figure 11 represents the peak of the light pulse (50^{th} bin). The decrease that occurs when more than one photon is incident is due to pulse pile-up. The sequence of curves below

corresponds to a single bin numbers and to averaged over various number of bins, as is labeled.

For tests with a small number of incident photons, curves are parallel to the x-axis in Figure 11. This corresponds to the linear region of the detector response where the number of detected photons is proportional to the number of incident photons. As the number of photons increases pulse pile-up decreases the number of detected photons. This limits the detector to counting ~1 photon per dead-time interval. At short time delays the processes producing time delayed counts also saturate the detector and all of the curves are limited by pulse pile-up.

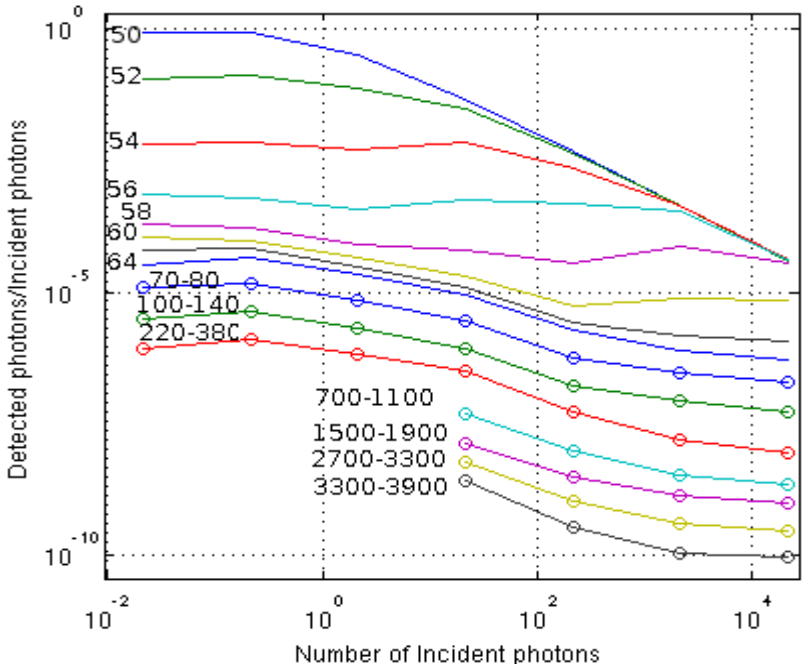

Figure 11. Normalized photon counts of one bin as a function of the intensity of the illuminating pulse. The colored solid curves serially from the top to the bottom represent experimental results for separate bin numbers (50, 52...64), and for averaged values over different bin ranges shown in the picture.

Signatures indicating processes with different time constants can be seen in Figures 10 and 11. Also, we some of these processes are proportional to the number of incident photons on the detector sensitive area (P-group) and processes proportional to the number of counted photons (C-group). The decaying portions of the normalized signals immediately following the laser peak in Figure 10 coincide with a straight sloping line when they are plotted in log-scale on the y-axis. This exponentially decaying signal with a time constant of ~40 ns is the trailing edge of the laser pulse. It is labeled as a process P1. The time constant of the process changes for different pulse repetition rates, which proves that this is the laser pulse. This process initially produces a high contribution to the signal, but decreases rapidly. For that reason, the upper curves in Figure 11 have a larger spacing in the linear region in comparison with the lower curves. After the peak, as the contribution of this process decreases, the signal photons in the trailing edge of the pulse become less affected by the pile-up effect. It then extends to the linear regions for the curves below the pile-up curve.

At larger time delays, greater than bin number 53, the curves in Figure 10 starts to separate at different bin numbers (from 53^{th} to 61^{th} bins). At the same time, the bottom curves in Figure 11, for tests involving between 1 and 100 photons in the incident pulse, tend to follow the pile-up limited curves for short time delays for bin numbers 50 and 52 in the Figure 11. This indicates that the process producing these time delayed counts is proportional to the number of counted photons rather than the number of incident photons. This is labeled as a C-group process.

Measured data profiles have a complex wavefront and, in order to explain the non-linear decrease of spacing between the curves in Figure 11, the C-group includes three

processes proportional to the number of counted photons with different time constants. The values for time constants of the C1 and C2 processes were derived empirically (see Table 1).

The curves in Figure 10 stay separated as time increases. For higher bin numbers in Figure 11, the spacing between the curves decreases slower than for the bins following the laser pulse and the shape of the curve in the bottom is preserved. This shows the C3 process. The time constant for the process C3 (0.5ms) was estimated from the measurements at lower repetition rate of 500 Hz for the time interval from 0.2 ms to 1.4 ms after the laser pulse.

The proportionality of processes to the number of counted photons is usually attributed to the classical afterpulse effect from the electrons trapped on the defects in the semiconductor. When the incident on the detector photon initiates an electron avalanche after the diode break down, some of the electrons are trapped in the avalanche region on the crystal defects. These are released by thermal motion and produce afterpulses.

The spacing between the lower curves in Figure 11 in a region for small number of incident photons is larger than that for higher numbers where they are more parallel to the x-axis. The signals for two tests with the highest number of incident photons have similar values and the three lowermost curves in Figure 10 get closer over time. The P2 and P3 processes with different decay times are proportional to the incident light intensity. These processes with time constants larger than the trailing edge of the laser pulse (P1) were used to explain this behavior.

The P2 and P3 may be fluorescence processes or a result of the thermal effect in the detector. The dark count rate of the detected signals increases with the number of incident photons. We assume that prolonged detector saturation results in the detector heating and increases the baseline of the signal. Thus, the time required for the detector thermoelectric

cooler to dissipate the heat is large. The time constant for the thermal process was estimated from the raw data which includes the dark counts and the background noise and is equal to ~35ms. So, the thermal effect keeps the lower curves close to each other indicating the linearity of the process to the number of incident photons.

4.4.2 Response model

A model describing the detector response was used to qualitatively interpret the experimental data. The two groups of processes were included in the model: processes proportional to the number of incident photons on the detector sensitive area (P-group) and processes proportional to the number of counted photons (C-group). Each of the groups includes three processes which are represented with an exponentially decaying function over time with different time constants and amplitudes:

$$F(t) = Ae^{-\frac{t}{\tau}} \tag{27}$$

where t is time, A is an amplitude, and τ is a time constant of a process.

A light pulse profile is derived by combining the points for bin numbers from 47 to 51 from the detected profiles for the test with the smallest number of incident photons with an approximation of the trailing edge with exponentially decaying process (P1) and scaled by the attenuation factor from the measurements. The inverse pile-up correction is then applied to light pulse profiles, which are scaled by the number of incident photons N_p:

$$N_p(t) = \frac{N_p(t)P1(t)}{1+N_p(t)P1(t)t_d} \tag{28}.$$

The detector dead-time value of 53 ns was used for the corrections, which is derived from the measurements. The corrected light pulse profiles $N_c(t)$ are then convolved with the sum of the C-processes by using a fast Fourier transforms. The P-processes are scaled by the total number of incident photons and then added to the convolved profiles. The signals are presented by the equation:

$$S(t) = N_p\big(P2(t) + P3(t)\big) + N_c(t) * (C1(t) + C2(t) + c3(t)), \quad (29)$$

where N_p is the number of incident photons.

Table 1: Parameters for C- and P-group processes.

Process	A	$\tau, \mu s$
P1	1	0.04 measured
P2	$1.3 \cdot 10^{-5}$	2.25
P3	$4 \cdot 10^{-10}$	500 measured
C1	0.07	0.5
C2	0.007	20
C3	$5 \cdot 10^{-4}$	300

The model photon count normalized by the number of incident photons as a function of time is shown in Figure 12, which is similar to Figure 10. Figure 13 shows normalized photon counts of one bin as a function of the number of incident photons in the illuminating pulse derived from the model (dashed lines). The colored curves represent the measured detector response (same as Figure 11) and dashed lines are derived from the model.

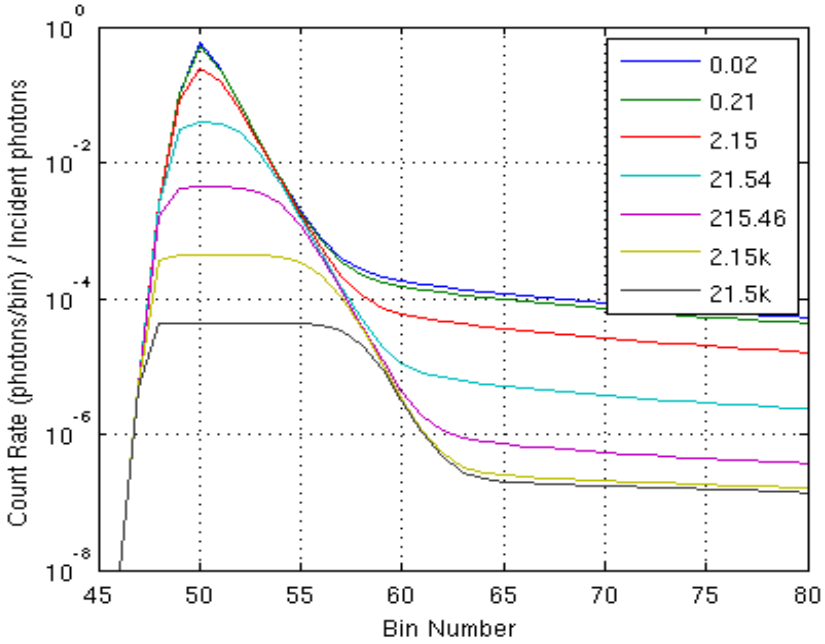

Figure 12. The model signals normalized by the number of incident photons vs. time (1Bin = 50 ns). Legend shows the total number of incident photons.

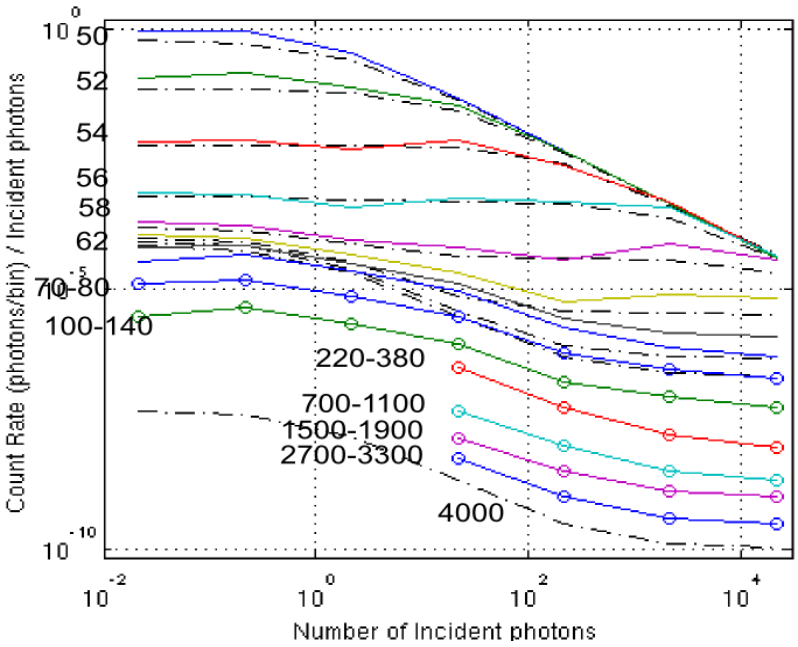

Figure 13. Normalized photon counts of one bin as a function of the intensity of the illuminating pulse. The colored solid curves serially from the top to the bottom represent experimental results for separate bin numbers (50, 52...64), and for averaged values over different bin ranges shown in the picture. The dashed lines are the model results for separate bin numbers 50, 52…64 and 4000.

The trailing edge of the laser pulse was found to have a 40 ns time constant. The difference between the measured detector response $S_m(t)$ and the exponentially decaying trailing edge of the laser pulse $N_p P1(t)$ divided by the total number of counts (denominator) is the total afterpulse probability p:

$$p = \frac{\sum_{t=0}^{200\ \mu s} S_m(t) - N_{pP\,1(t)}}{\sum_{t=0}^{200\ \mu s} S_m}. \quad (30)$$

The total afterpulse probability calculated for the measurement with the smallest number of incident photons (in order to avoid the pile-up effect) is about 1 %.

The model is able to simulate the features of the measured detector response. The processes representing the trailing edge of the laser pulse and the processes with large time constants containing the measured time constants have a good correspondence with the experimental data. However, the actual detector response has a more complicated shape and the current model overestimates the response at large time delays for the pulses with a small number of incident photons (less than 1 photon in the pulse). One possible explanation is that the model does not account the secondary afterpulse effect produced from the afterpulse counts. Another possible reason is an increase of the time constant for the thermal effect with a number of incident photons in a light pulse.

4.4.3 Response linearity test

A linearity test was conducted for the detector under overload condition. The increment of light intensity was a factor of 1.25 and the number of photons was varied from 2,500 to 23,000 photons per laser pulse. In order to define the deviation of the detector response from the linear response, the photon counts as a function of the intensity of the

illuminating pulse were approximated with exponential functions for various ranges of bin numbers. The example for the photon count averaged over a bin range 600 to 800, its approximation and the function for linear detector response (proportional to the number of incident photons) is shown in Figure 14.

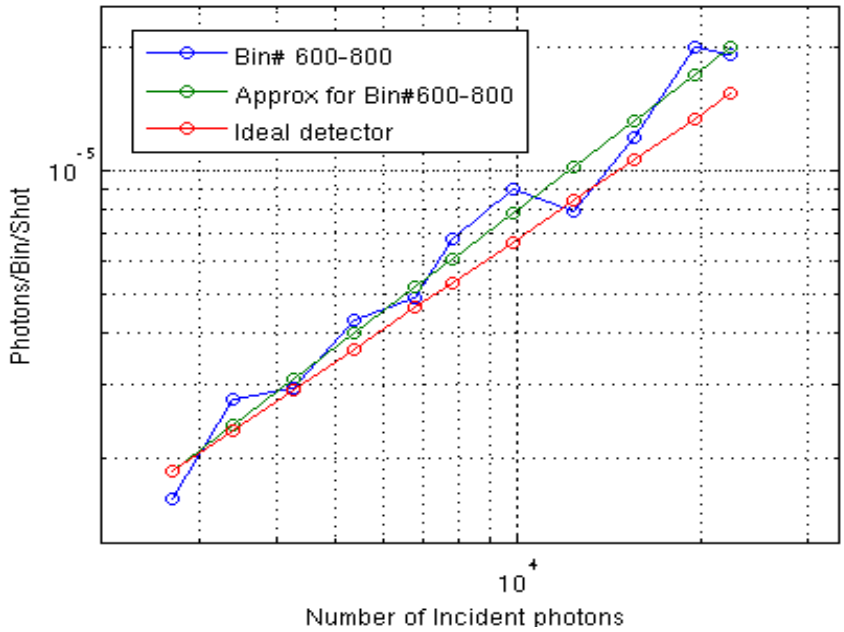

Figure 14. A photon counts averaged over bin numbers 600-800 as a function of the number of incident photons (blue), its approximation (green), and a photon count for an ideal detector with the same quantum efficiency as the actual detector (red).

The approximations for measured detector response were compared to the linear response and the deviation of the real detector response from the linear one for various bin ranges is presented in Figure 15. It shows that the deviation of the detector afterpulsing from the linear response tend to increase with a bin number (with time delay). For the number of incident photons above 1000 in the light pulse the process P3 with a large time constant (a process linear with respect to the number of incident photons, described in Section 4.4.2) has a large contribution to the raised baseline of the signal for large time delays. In the real lidar

system the transmitted energy can be changed within 30%; and, according to Figure 15, the energy change corresponds to the deviation around 3% for bin ranges from 600 to 1000 in the baseline profile from internally scattered light. This result shows that in the real lidar signals the scaling of the baseline profile to a transmitted energy produces a relatively small error. Intervals for larger bin numbers are not presented here due to large noise in the data. The result of a similar test performed on the lidar baseline for different transmitted energies in the light pulse are presented later in Section 5.1.

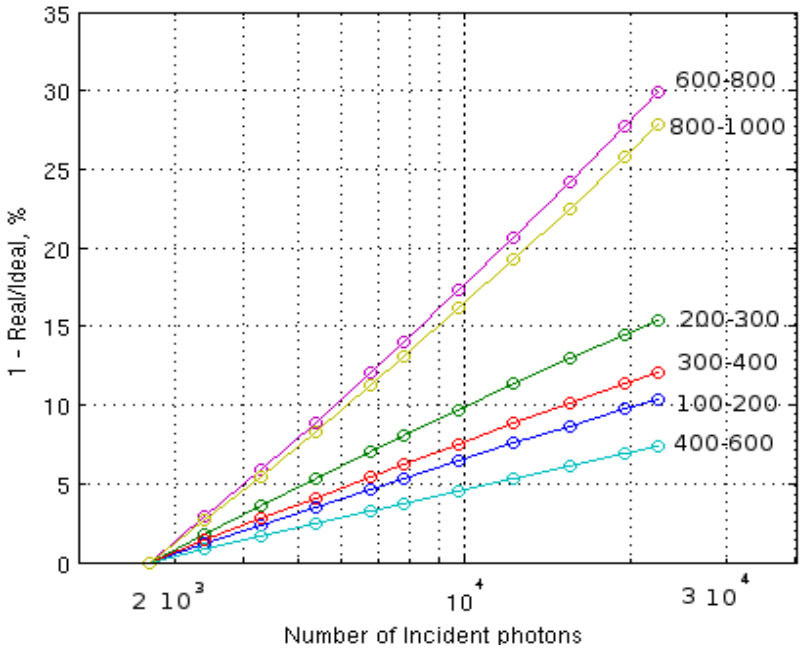

Figure 15. A deviation of the afterpulse detector response from linear response.

4.4.4 Additivity test

A large amount of internally scattered light in the lidar system appears in the detected profiles as an increased baseline of the signals (afterpulse effect). The test with two sources of light (a laser light and a LED light) was performed in order to investigate an influence of the internally scattered light on the receiving signals, and to evaluate the additive afterpulse properties in the presence of the highly intense light pulse saturating the detector.

The laser light was used for simulation of the scattered light producing a large afterpulse signal, which is presented by a green curve in Figure 16. The light from LED was used as the second source of light, which appears in the region for bin numbers 170-210 in Figure 16. The signal from the laser and LED were measured separately (the red and the green curves in the Figure 16 respectively) and for both sources simultaneously (blue curve). The sum of separately measured laser and LED signals has a good correspondence with a simultaneously measured two sources (Figure 17). The test shows that the afterpulsing from the scattered light can be reduced by measuring the scattered light as a function of time-range (baseline profile) with a covered telescope and subtracted from the lidar returns.

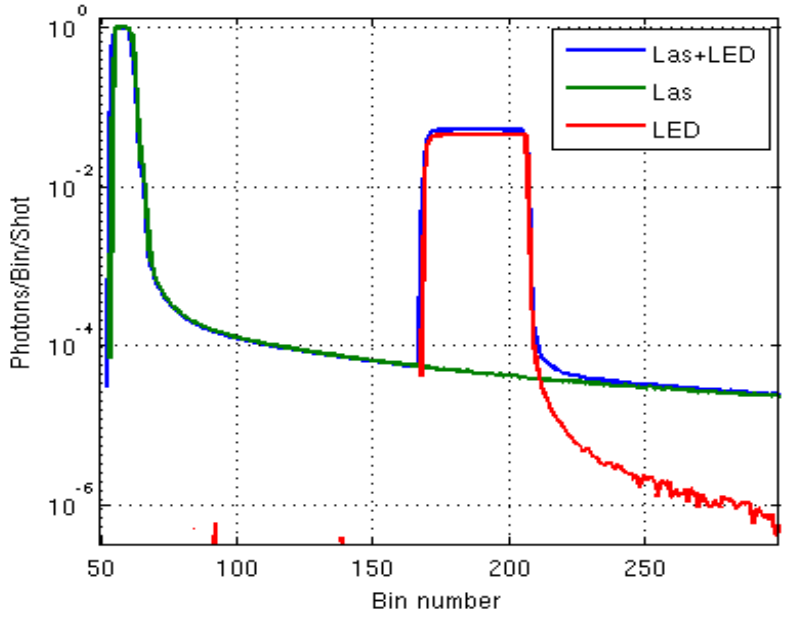

Figure 16. Laser pulse, LED pulse, Laser+LED pulse.

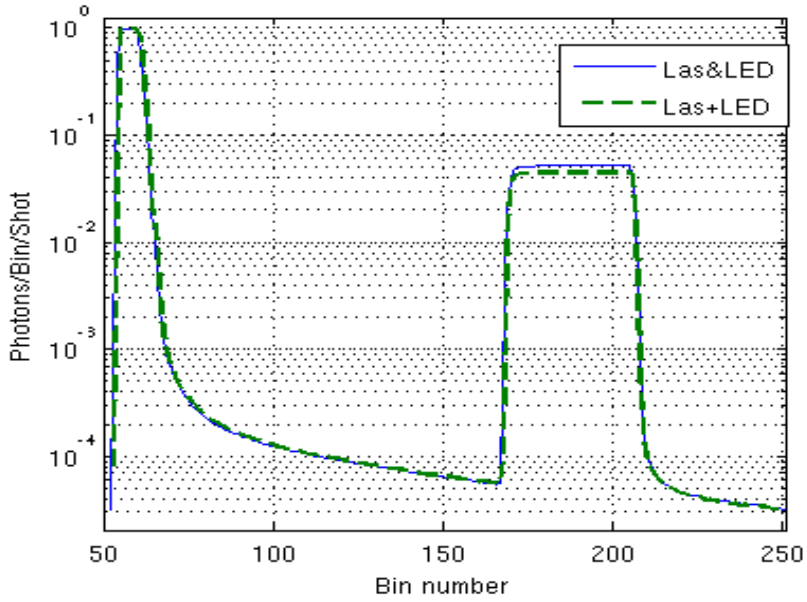

Figure 17. Laser pulse & LED pulse (measured for both sources turned on), Laser+LED (a sum of independently measured profiles) pulse.

The generated by a LED light pulse with a sharp raising and trailing edges best demonstrates a significance of the afterpulse effect shown in Figure 18. Even thought the pulse amplitude is low, the trailing edge of the detected signal is not that sharp as the raising edge. However, the electrical pulse which drives the LED is a rectangular pulse. The deconvolution of the detected LED pulse signal shows a significant correction for the detected signal. However, the origin of the raised trailing edge is not purely attributed to the afterpulsing, so part of it is caused by the fluorescence in the LED p-n junction. For that reason, it is difficult to define the afterpulse correction accuracy, but it is evident that the signal in the decaying signal following the pulse should not decrease down to the background noise immediately after the pulse. The fluorescence of the LED following the electric pulse was observed with a fast photo diode and oscilloscope, so the uncertainty in the amount of afterpulse remains in this test. The trailing edge of the detected light pulse increases with a pulse width which indicates an increasing contribution of LED fluorescence.

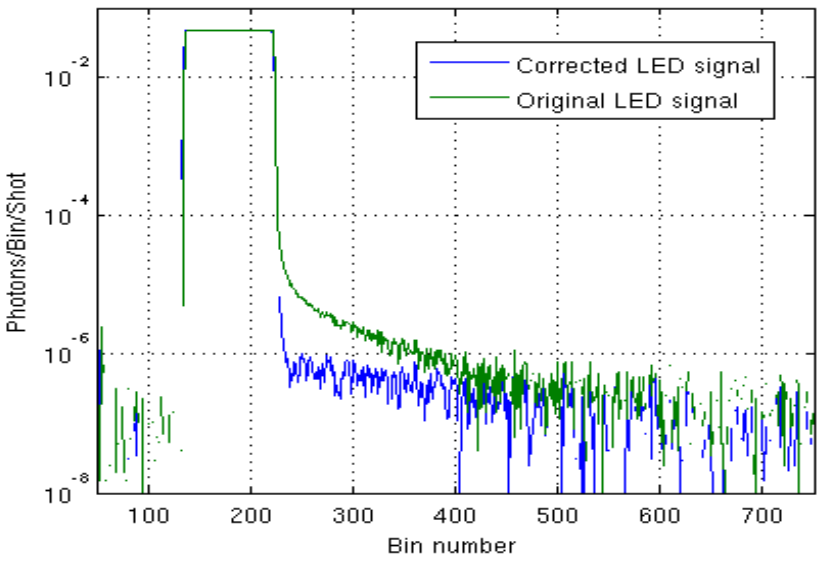

Figure 18. Correction of a LED pulse by deconvolution with detector response function.

The difference between the two signals in Figure 18 shows that the afterpulse correction removes around 50 % of the raised trailing edge from the signal. The breaks of the blue curve at the trailing edge are caused by high frequencies in the signal spectrum (a Fourier transform of the signal) due to the sharp edges of the pulse. A Blackman window was applied to the signal spectrum for the deconvolution with the response function in order to suppress the amplification of high frequency noise. The analysis of the measured detector response function shows that the deconvolution removes afterpulsing effectively for about 150 bins following the signal light pulse. The deconvolution of the data with measured laser pulse removes not only the afterpulse effect, but also the effect of the laser pulse shape if this is done for longer averaging intervals.

Chapter 5

Lidar Data

5.1 Lidar data and corrections

On April 14, 2009 the HSRL collected data from 12:30 UT until 13:30 UT, in Madison, Wisconsin (43.07 N, 89.38W). Water clouds at heights 1.5 km and 4 km were observed during this period. Figures 19 and 20 show altitude versus time image of attenuated backscatter and aerosol backscatter cross-section derived from inverted HSRL data.

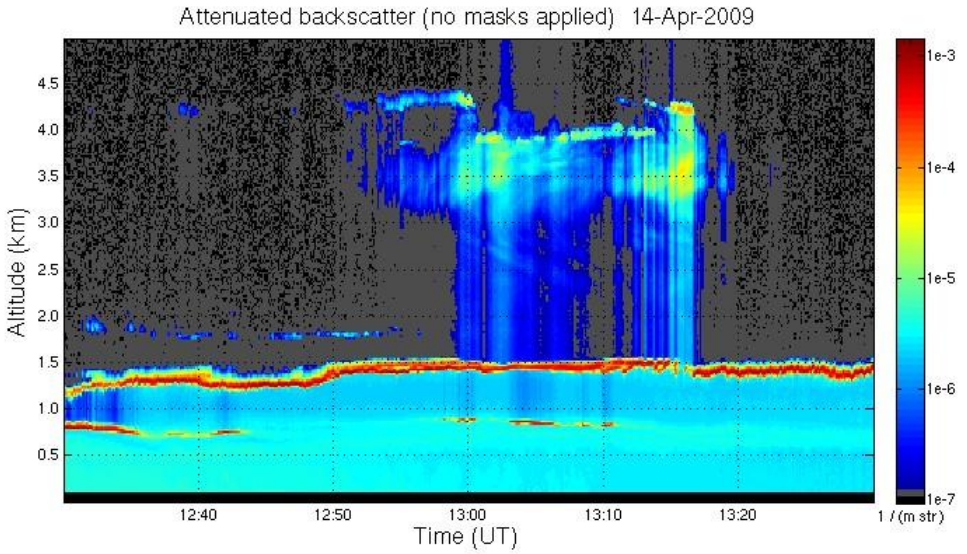

Figure 19. Altitude vs. time image of attenuated backscatter showing stratus water clouds from 12:30 to 13:30 UT on April 14th, 2009. The period used to correct the data is from 13:00 to 13:10 UT.

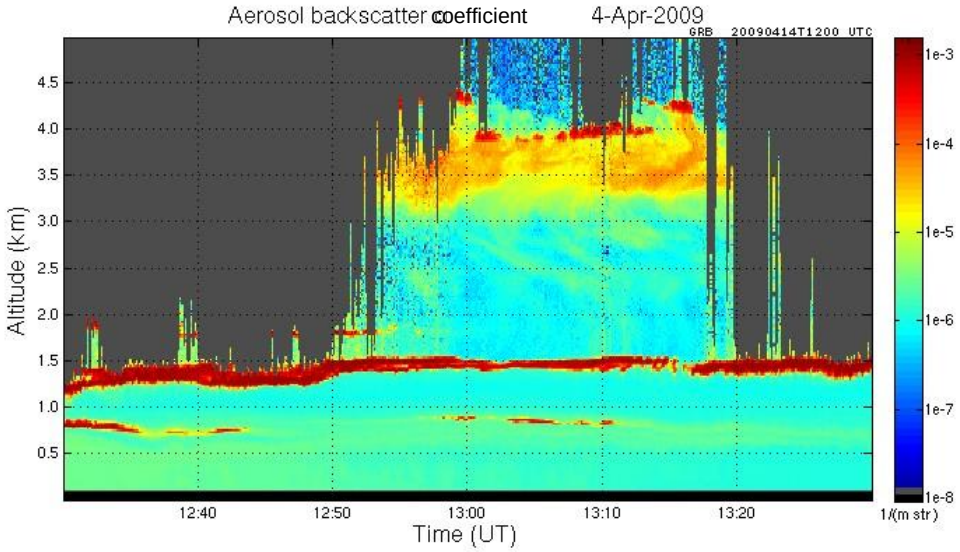

Figure 20. Altitude vs. time image of aerosol backscatter coefficient showing stratus water clouds from 12:30 to 13:30 UT on April 14th, 2009. The period used to correct the data is from 13:00 to 13:10 UT.

Figure 21 shows an altitude versus time image from the same period of inverted aerosol depolarization.

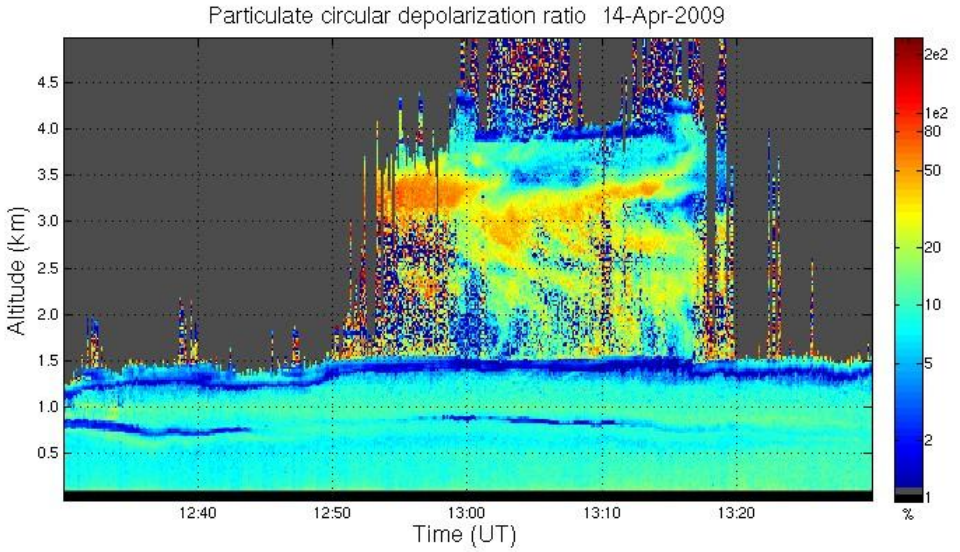

Figure 21. Altitude vs. time image of aerosol depolarization ratio from 12:30 to 13:30 UT on April 14th, 2009. The period used to correct the data is from 13:00 to 13:10 UT.

A temperature profiles from radiosonde measurements are presented in Figure 22 closest for the shown period.

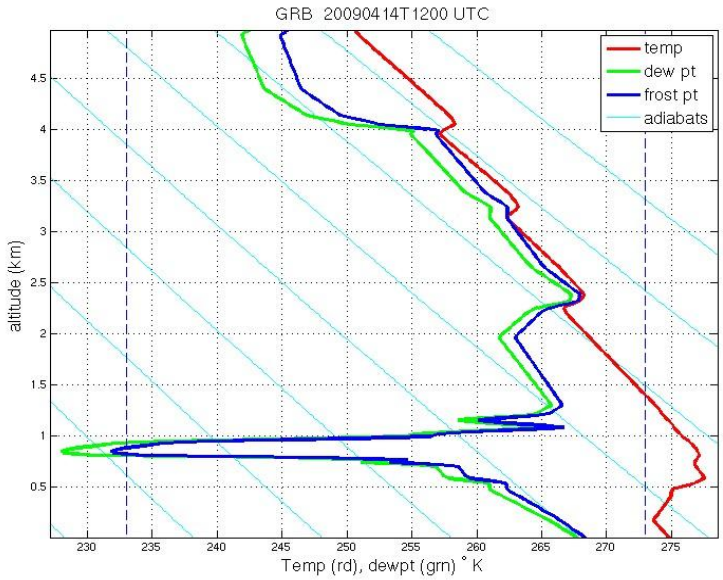

Figure 22. Radiosonde temperature profiles from 12:00 UT on April 14th, 2009, Green Bay, WI used for system calibration

Figure 23 shows raw lidar returns (background noise subtracted) averaged over the selected time period on April 14, 2009, 13:00 – 13:10 UT. The dashed lines on the plot represent measured scattered light profiles (baseline) as a function of time (altitude). The background noise for the data was calculated by averaging 43 data points prior to the laser pulse. These points are a low level signal in front of a high spike from internally scattered transmitted light in the raw data profiles.

As previously discussed, the lidar returns are distorted by the afterpulse effect caused by the light scattered from the internal optical elements at the moment of laser firing (increased baseline), the afterpulse effect from clouds and a pile-up effect due to an inherent detector dead time. The internally scattered light produces a significant impact in the detected

signals for the first few hundred meters and at high altitudes, at which point the signal decreases down to the level of the baseline profile or below. That can be seen from comparison of the baseline profiles (dashed lines) with the detected signals (solid lines) in Figure 23, where the slowly decaying signal above 5 km is mostly produced by the internal scattering effect. The baseline profiles from the scattered light for combined high and low gain channels are labeled as *scat ch* and *scat cl* correspondingly; the baseline for molecular and cross polarized channels are labeled as scat m and *scat cr*. The additivity test (see Figures 16 and 17) showed that the internally scattered light can be minimized in the lidar returns by performing a baseline correction on the signals. This assumes that it can be independently measured as a function of time (range) by covering the telescope and then subtracted from the lidar return as range dependent values.

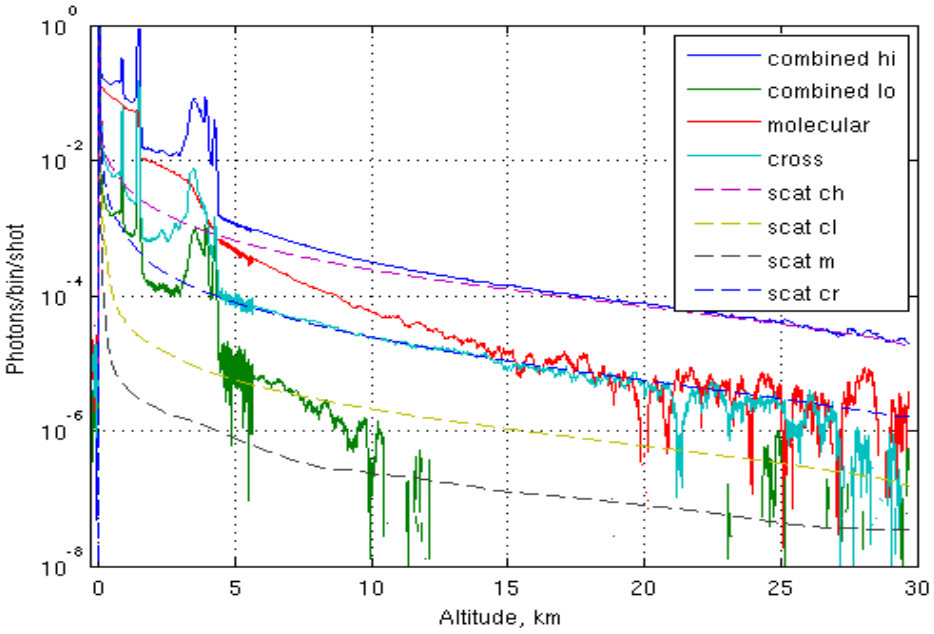

Figure 23. Combined high gain channel (blue line, combined hi), Combined low gain channel (green line, combined lo), Molecular channel (red line, molecular), Cross polarized channel (cayenne line, cross). Case of dense water cloud, 14 April 2009, UT 13:00-13:10, Madison, WI.

A scattered light measurements averaged over one hour (baseline profiles) were smoothed by filtering with a varying window size (weighted average of the points) in order to reduce the noise (dashed lines in Figure 23).

The result of performed baseline correction for the combined high gain channel is presented in Figure 24, in which the green curve is a background corrected signal and the red curve represents the signal after baseline and background corrections. As can be seen from Figure 24, a baseline correction completely eliminates the spike at ranges below 100 m caused by the intense scattered light in the transmitted pulse, and slightly lowers the profile at ranges below 1 km and from 1.5 km to 3 km where the signal is strong, and substantially at ranges more than 4.5 km where the signal is weak (above the cloud). The baseline correction routinely applied to the lidar data is mentioned in [27, 28].

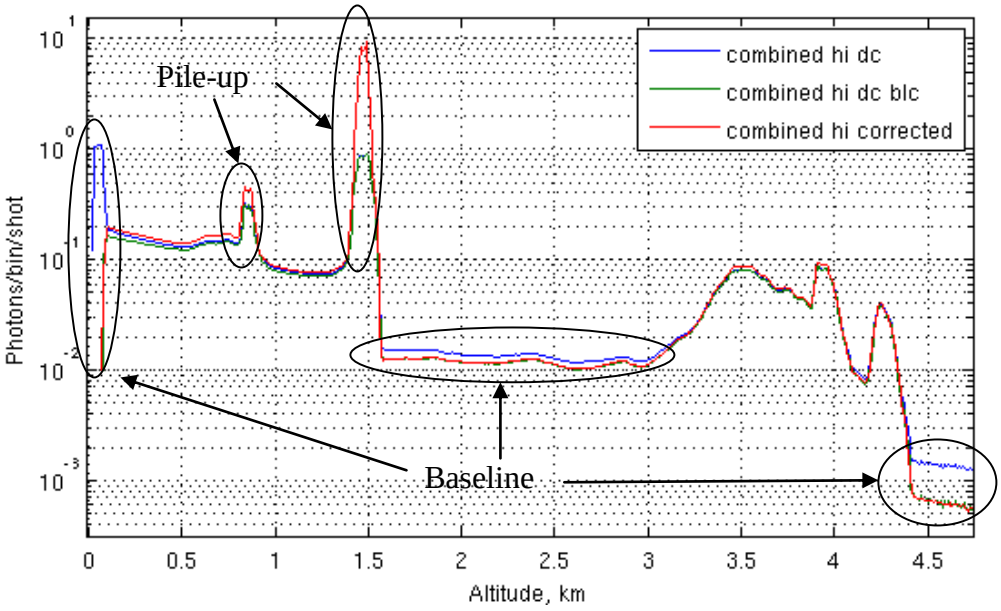

Figure 24. Combined high gain channel fully corrected (blue), Combined high background corrected (green), Combined high background and baseline corrected (red). Case of cloud, 14 April 2009, UT 13:00-13:10, Madison, WI.

The pile-up correction for detector nonlinearity, described in the Section 3.3, was applied to the signals. The red curve in Figure 24, which represents a fully corrected profile, is above the blue curve (signal with subtracted background noise) everywhere from 100 m to 1.5 km. This is the result of a performed pile-up correction for the signals which is stronger for larger signals and, hence, causes a substantial difference at the 0.8 km and 1.5 km between the red and the green (blue) curves in the Figure 24.

The blue and green curves are the background, and background with baseline corrected profiles correspondingly. The red curve in Figure 24 shows a result of a full correction applied to the signal in the following order: baseline correction, afterpulse correction for atmospheric photons, pile-up correction and background correction. Above the cloud at 1.5 km the signal is weak (the pile-up correction is small) and red and green curves are almost coincide.

The afterpulsing in lidar returns from atmospheric photons is significant for dense and geometrically thin clouds and it becomes noticeable after the baseline correction. A case of a thin and optically dense cloud for the time period 13:20 – 13:30 UT on the same day was chosen to demonstrate the afterpulse effect from the atmospheric photons (Figure 25). The cloud at 1.5 km significantly attenuates the transmitted light, so that the detected signal above the cloud (2.3 km height) decreases by a factor of 1000 almost to the level of the baseline profile for the combined channels, and the afterpulse correction reaches 25% of the original signal for this example.

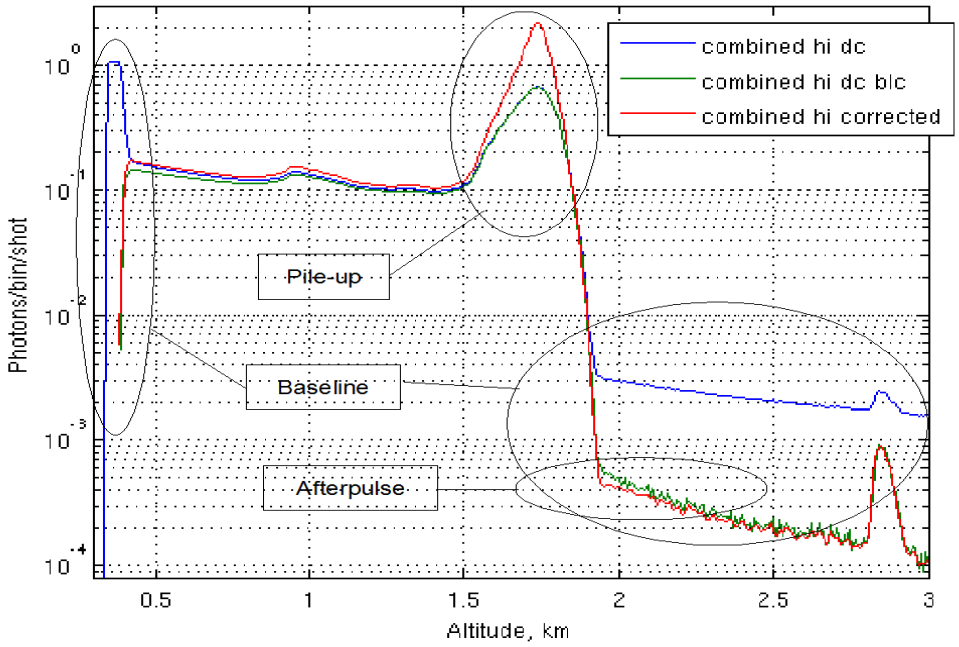

Figure 25. Combined high gain channel fully corrected (blue), Combined high background corrected (green), Combined high background and baseline corrected (red). Case of cloud, 14 April 2009, UT 13:20-13:30, Madison, WI.

The afterpulse effect from atmospheric photons was considered as a linear effect, and is proportional to the number of total counted photons. Therefore, the deconvolution of the signal with the detector response function (afterpulse correction) is done before the pile-up correction. The pile-up effect depends on the signal strength and, hence, it has to be applied to the recorded signal containing background noise. The inverse order of correction will result in the afterpulse overcorrection.

Note that the detector pile-up affects the counts of the total signal including the afterpulse signal and the signal baseline. If we first apply the afterpulse correction, we do not significantly affect the pile-up correction, because of the relatively "weak" afterpulsing from the cloud signals (above 1.5 km in Figure 24, above 2 km in Figure 25). If we then proceed with the pile-up correction and then subtract the signal baseline, we underestimate the

baseline at small ranges which results in a larger signal. Therefore, the baseline correction followed by the afterpulse and pile-up correction allows to derive the profiles which are closer to the "true" value.

The deconvolution of lidar signals with detector response function was done by using the Fast Fourier transforms. Usually, the deconvolution amplifies the high frequency noise, which presents in the signal, and in order to eliminate this effect the Blackman window was applied to the spectrum of the detector response function.

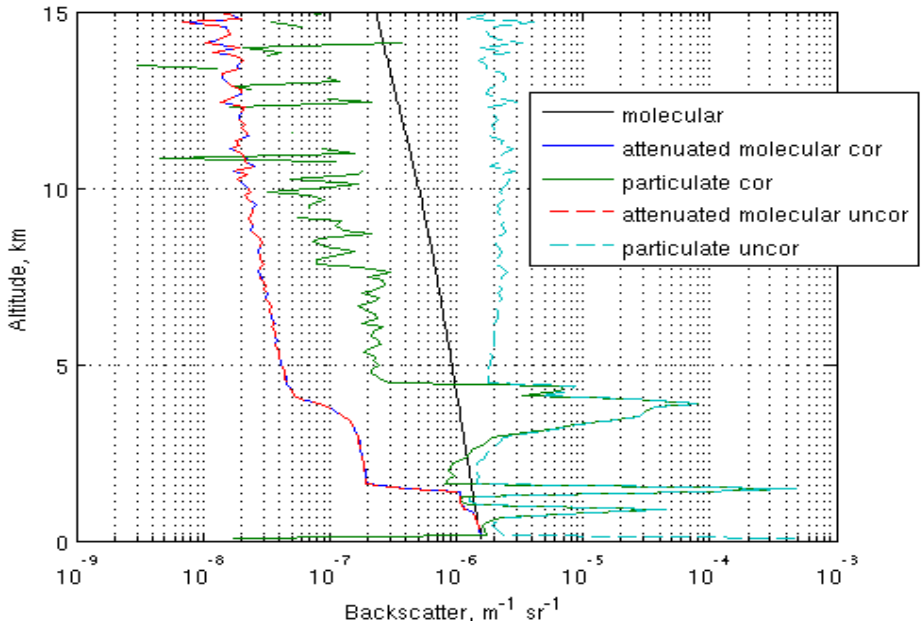

Figure 26. Backscatter profiles. Case of dense water cloud, 14 April 2009, UT 13:00-13:10, Madison, WI.

An example of calculated values of molecular and aerosol backscatter cross sections, scattering ratio, and optical depth profiles for a baseline and afterpulse corrected and uncorrected data are presented in Figures 26, 27, and 28 correspondingly. A comparison of the aerosol backscatter cross section profiles for corrected and uncorrected data demonstrates a significant difference. If a baseline correction is not performed on the signals, a significant increase of the backscatter cross section with height appears for the aerosol channel. This

produces an increase of the scattering ratio (see Figure 27) and decrease of the optical depth as well in case if returns are not baseline corrected. The amount of the internally scattered light in molecular channel is small because the iodine rejection filter absorbs the most of this light and the baseline of the signal in molecular channel is low. For that reason, the optical depth profile for uncorrected data has a very close values and almost indistinguishable from that for corrected data as shown in Figure 28.

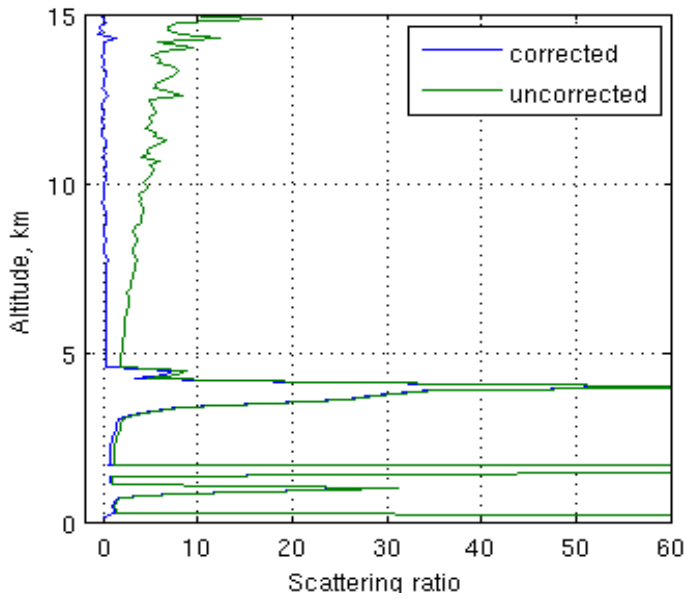

Figure 26. Backscatter ratio profile from water clouds: corrected (blue), uncorrected (green); on 14 April 2009, UT 13:00-13:10, Madison, WI.

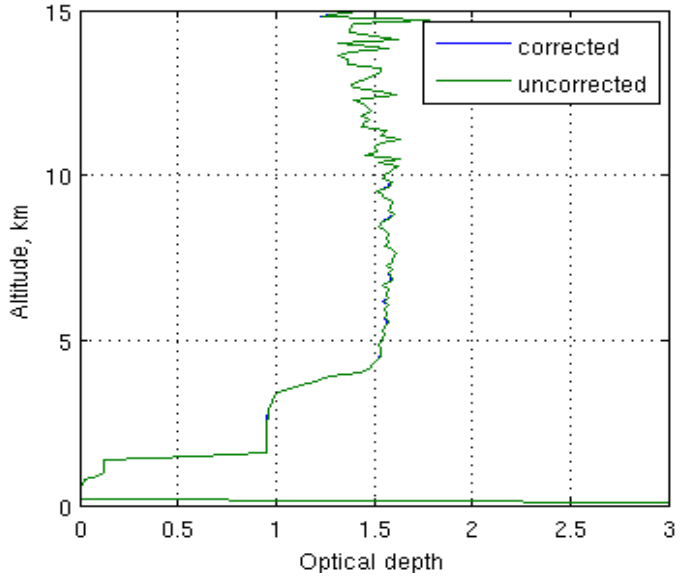

Figure 27. Atmospheric optical depth profile from water clouds on 14 April 2009, UT 13:00-13:10, Madison, WI.

The afterpulse effect from the internally scattered light from the optical elements appears in the lidar signals as an increased baseline. The subtraction of the baseline signal improves the aerosol backscatter cross section profile and allows to derive a physically reasonable values for the backscatter coefficient profile. The correction for the pile-up effect becomes less accurate as the signal increases; the lidar receiver contains two detectors to enhance the dynamic range in the aerosol channel in order to avoid the pile-up effect and detector saturation from strong returns in case of a dense cloud. Thus, for signals below the threshold of .3 counts/bin the detector with a high sensitivity is used to detect the photons, and if the signal rises above the threshold then the low sensitivity detector is used. However, in case of an optically dense cloud the backscattered light can saturate the high sensitivity detector and may produce significant afterpulse effect in a nonlinear region of the detector response, where it is not proportional to the counted photons. This effect is similar to the afterpulsing from the internally scattered light. When the signal becomes below the

threshold, the high gain channel is used again for the signal; however, its significant portion is the afterpulse signal from the photons scattered by clouds and the data are no longer representative. The increased baseline from the large atmospheric signal can be explained by the process proportional to the number of incident photons with a large time constant, labeled as P3 process in the model (see Section 4.3.2).

The results of a test performed for the detector baseline sensitivity for the lidar is presented in Figure 28. The baseline profile was measured for two different values of the transmitted energy corresponding to average power values of: 180 mW and 350 mW; the profiles were normalized by the transmitted energy. The deviation of the measured baseline profile between the 180mW and 350mW for combined high sensitivity channel (which has the largest amount of the internally light) as a function of altitude is presented in Figure 28. It shows that the error associated with the changes of transmitted energy monotonically increases with height and gives the maximum error of 10%. The large values for the altitudes up to 2 km is caused by the detector saturation, because the spikes at small time delays produced by internally scattered light for different light pulse amplitudes are similar. For that reason the difference between the baseline profiles becomes large after scaling them. The linearity test for detector response described in Section 4.3.3 also shows the increase of the afterpulse at larger time delays with number of incident photons. The changes of transmitted energy in the lidar does not exceed 30 %, thus the maximum error associated with changes of the transmitted energy is less than 5%. The sensitivity of the baseline correction to the change of transmitted energy was estimated. The baseline was scaled within range of ±30%, which are reasonable values for system operation over long time periods. For the same case (clear atmosphere) when the baseline is bigger, or transmitted energy is smaller, the error of

the scattering ratio drops below 1% at 400 meters and gets bigger with height, reaching 5 % at 7 km, then becomes around 10% above 25 km. The baseline correction becomes sensitive to the transmitted energy changes when the atmospheric signal becomes small, i.e. at high altitudes and when it is attenuated by clouds.

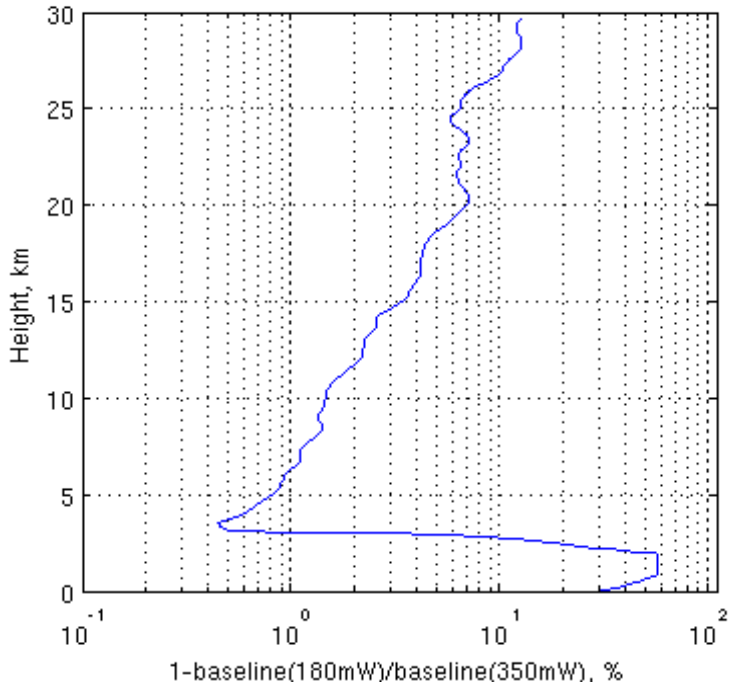

Figure 28. The deviation of the baseline profile for 350 mW and 180 mW as a function of time (altitude).

In the clear atmosphere the signal in the combined channel is mostly produced by molecular scattered photons and only a small fraction of the signal is the aerosol scattered photons. The aerosol channel becomes very sensitive to the baseline correction because after subtraction of the molecular signal from the combined channel, the aerosol signal becomes very small, on the same order as the error in the baseline. For that reason it is very important to have an accurate baseline correction of the signal.

The dust and accumulated precipitation on the window along with the dust on the optical elements may change the amount of the scattered light because the lidar operates through a glass window. The baseline correction becomes sensitive to these changes at high altitudes, where the signal is comparable or less than the baseline. The telescope blocking screen was found to be a good light absorber and does not scatter light back to the receiver. The scattered light in the lidar was attenuated by a neutral density filter with a 10^6 attenuation factor and was measured for a covered and uncovered telescope. The difference between the scattered light in presence of a black screen and without the screen is less than 0.1%. However, for the white screen the scattered light was 4 % larger than that for the black screen.

Also, it was found that the inverted data (scattering ratio) becomes sensitive to the baseline correction when the signal swiftly decreases, like in cases of the cloud when the signal is not completely attenuated, but has a value close to that of the baseline profile.

The instabilities in the laser may change the time delay between laser trigger pulse (Q-switch) and the laser light pulse (build-up time) which changes the position of the signal baseline. It was found that the baseline correction is not sensitive to the laser pulse shift and only significant for the first few bins after the laser firing, where the baseline profile is high relative to the signal. When the baseline decreases relative to the signal (where it is presented by the process with a large time constant ~500 μs), the correction becomes less sensitive to the shift. For clear atmosphere the shift of the laser pulse of 2 bins, which is two times as large as the threshold for the build-up time for a seeded shot, does not produce significant changes for inverted data profiles. It makes difference only for the first few hundred meters and at 400 meters the error drops below 1 %.

The afterpulse correction of the atmospheric photons, by performing the deconvolution of the detector response function, was found to be around 1 % of the incident energy for small number of incident photons. The deconvolution with the measured detector response function produces a phase shift of the signal within one bin. However, it does not affect the data, because the laser pulse build-up time variation is allowed to be larger.

5.2 Discussion

Clouds attenuate the laser beam so that the molecular signal decreases in clouds more rapidly than in the clear atmosphere. At the same time, the aerosol signal decreases faster with height than the molecular signal in the clear atmosphere because of the presence of much less aerosol higher in the atmosphere, and the molecular signal decreases with decreasing air density. The aerosol backscatter cross section is proportional to the ratio of aerosol and molecular lidar returns (lidar backscatter ratio). In a clear atmosphere, the backscatter ratio usually decreases slowly with height, and is used as an indicator of the lidar data quality. The increase of the backscatter ratio with height (or a slower decrease of the aerosol backscatter coefficient than normal), indicates either a molecular signal is small or an aerosol signal is large than the actual value. The afterpulsing from internally scattered light and from atmospheric returns usually causes the increase of the backscatter ratio by producing an extra signal in the lidar returns, which is the largest in the high sensitivity combined channel.

However, there are other range dependent effects such as range dependent calibration coefficients and misalignment of molecular and combined channels, which also can cause the backscatter ratio increase.

A misalignment of the molecular and combined channels leads to different overlap functions. In order to compensate for this effect, a differential geometry correction is applied for the two channels. When the iodine cell is removed from the molecular channel it operates as the combined channel. The ratio between the two channels should be constant with height if the optical elements are well aligned. If this ratio is not constant, the ratio between the two channels as a function of range is applied to the overlap function of one of the channels in order to compensate for the misalignment of the two channels.

Another possible reason of the signal growth is the errors in calibration coefficients. The expressions for the molecular photons N_m and aerosol photons N_a (Section 2.2) describes the HSRL inversion technique of the detected signals. The coefficients C_{mm}, C_{mc}, and C_{am} in these expressions are calculated from calibration. If the C_{mm} (amount of the molecular signal in the molecular channel) is overestimated, then the molecular signal becomes smaller, and the aerosol signal gets larger. The coefficients C_{mm} and C_{mc} are temperature and pressure dependent, and, therefore, they are computed as the altitude dependent values.

Chapter 6

Summary

The University of Wisconsin High Spectral Resolution lidar (HSRL) operates through a single telescope to transmit and receive light. This reduces the energy density in transmitted beam to eye safe levels and improves mechanical stability allowing it to operate with a small field-of-view. A drawback of this design is that some of the transmitted light is scattered from internal optical elements and part of this internally scattered light reaches the detectors. This results in detector overloading and induces time delayed spurious photon counts that contaminate the measured signal profile.

This thesis has presented measurements characterizing the Single Photon Counting Module from Perkin Elmer based on a Avalanche Photo Diode. The detector properties, revealed from the tests, were used in corrections for the lidar data. The baseline correction reduces the influence of the internally scattered light and now it is routinely applied to the lidar data. It assumes that the internally scattered light profiles can be measured with a covered telescope as a function of time (range) and then are subtracted from the lidar returns as height dependent values. That allowed us to significantly improve the data quality and made the lidar instrument more sensitive, so that it is able to detect optically thinner layers of aerosols.

The aerosol signal still remains sensitive to the baseline changes from the transmitted energy in cases of a clear atmosphere. The sensitivity tests showed that the scaling of the baseline to the transmitted energy reduces this effect and improves the data.

In the performed tests the detector was exposed to a wide range of light intensities, from 0.01 photons/shot to 25,000 photons in the laser pulse. An analysis of the data showed that the detector response properties depend on an amount of light it is exposed to. It was found that the afterpulse signal is around 1% of the detected signal for low numbers of incident photons and the detector's dead time is 52 ns, which is consistent with values specified by the manufacturer.

The similarity of the measured curves for photon counts as a function of the number of incident photons at the peak of laser pulse and away from the pulse (about 1μs and more) indicates that the afterpulse signal is proportional to the number of total counted photons for up to ~1000 incident photons in the laser pulse. This result shows that the afterpulse effect is mostly caused by the electrons trapped in the avalanche region (amplification region). Then, as the number of incident photons increases, the signal baseline starts growing linearly with the number of incident photons. We cannot explain the linear growth of the baseline at large number of incident photons. We suspect that it is the result of thermal effect.

The overall response of the detector to large light pulse is complex and can be considered as a superposition of two groups of processes. The processes in the P-group (C-group) are proportional to the number of incident (counted) photons. Each of the groups includes processes with different time constants. The measured time constant for the longest "afterpulse" process is around 500 μs. Because of such a slow decay, its influence on the

lidar profiles is noticeable for far ranges where the lidar returns becomes comparable to the afterpulse from the scattered light.

The largest afterpulse effect from both internally scattered light and from the atmospheric signals in the lidar system is observed in a high sensitivity combined channel. In case of the strong lidar returns from dense clouds the combined high sensitivity channel becomes saturated and when the signal is above 0.3 counts/bin the low sensitivity channel is used. At larger ranges where the signal returns below the threshold, the high sensitivity channel is used again. However, large signals in combined high sensitivity channel when the detector is saturated produces significant afterpulsing. That increases the baseline of the signal similar to the baseline produced by internally scattered light. In this case the afterpulse correction by deconvolution cannot significantly reduce the afterpulse effect in the signal. The future improvement may include a design of electronics which applies the gate signal of a variable length to the high sensitivity combined channel detector when it is saturated by a large signal in real time.

References

1. P.-H. Wang, P.Minnis, M.P. McCormick, G.S. Kent, ans K.M. Skeens, “A 6-year climatology of the cloud occurrence frequency from Stratospheric Aerosol and Gas experiment II observations (1985-1990), J.Geophys. Res.101, 29407-20429 (1996).
2. D.P. Wylie and W.P. Menzel, “Eight years of high cloud statistics using HIRS”, J.Clim. 12, 170-184 (1999).
3. Klett, J.D., 1981. Stable analytical inversion solution for processing lidar returns. Applied Optics, 20: 211-220.
4. Kovalev, V.A., Eichinger, W.E., Elastic lidar : theory, practice, and analysis methods, Hoboken, New Jersey : John Wiley & Sons, Inc., 2004.
5. Eloranta, E. W., I. A. Razenkov, J. P. Hedrick, J. P. Garcia, The design and construction of an airborne High Spectral Resolution Lidar. IEEE Aerospace Conference March 1-8, 2008, Big Sky, MT.
6. S. Cova, A. Lacaita, and G. Ripamonti, "Trapping phenomena in avalanche photodiodes on nanosecond scale," IEEE Electron. Dev. Lett. 12, 685–687 (1991).
7. Hughes, H.G.F., Jerry A.; Stephens,Donald H., 1985. Sensitivity of lidar inversion algorithm to parameters relating atmospheric backscatter and extinction. Applied Optics, 24 : 1609-1612.
8. Frederick G. Fernald Applied Optics, Vol. 23, Issue 5, pp. 652-653.
9. Sensitivity analysis of lidar inversion algorithms, Luc R. Bissonnette, Applied Optics, Vol. 25, Issue 13, pp. 2122-2125.

10. Eloranta, E. W., High Spectral Resolution Lidar in Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere, Klaus Weitkamp editor, Springer Series in Optical Sciences, Springer-Verlag, New York, 2005.
11. Measures R.M.: Laser Remote Sensing (Wiley-Intersciences, New-York 1984).
12. Piironen, P., "A High Spectral Resolution Lidar Based on an Iodine Absorption Filter", Ph.D. Thesis, University of Joensuu, Department of Physics, Finland, 1994.
13. M. Hobel and J. Ricka, “Dead-time and afterpulsing correction in multiphoton timing with nonideal detectors,” Rev. Sci. Instrum. 65, 2326–2336 (1994).
14. H B Coldenstrodt-Ronge et al “Avalanche Photo-Detection for High Data Rate Applications”, J. Phys. B: At. Mol. Opt. Phys. 40 3909-3921 (2007).
15. Torre, S.; Antonioli, T.; Benetti, P. “Study of afterpulse effects in photomultipliers”, Review of Scientific Instruments, Volume 54, Issue 12, December 1983, pp.1777-1780.
16. Akchurin Nural ; Kim Heejong .”A study on ion initiated photomultiplier afterpulses”. Nuclear instruments & methods in physics research. Section A, Accelerators, spectrometers, detectors and associated equipment, 2007, vol. 574, No1, pp.121-126.
17. A. Lacaita, M. Ghioni, and S. Cova, "Double epitaxy improves single--photon avalanche diode performance," Electron. Lett. 25, 841-843 (1989).
18. S. Cova, A. Lacaita, and G. Ripamonti, "Trapping phenomena in avalanche photodiodes on nanosecond scale," IEEE Electron. Dev. Lett. 12, 685–687 (1991).

19. G. Ripamonti, F. Zappa, and S. Cova, “Effects of trap levels in single-photon optical time-domain reflectometry: evaluation and correction”, J. Lightwave Technol., 10, pp. 1398-1402 (1992).
20. F. Zappa, S. Cova; M. Ghioni, A. Lacaita, C. Samori (1996-04-20). "Avalanche photodiodes and quenching circuits for single-photon detection". Applied Optics 35 (12): 1956-1976.
21. Fully ion-implanted p + -n germanium avalanche photodiodes, S. Kagawa, T. Kaneda, T. Mikawa, Y. Banba, Y. Toyama, and O. Mikami, Applied Physics Letters vol. 38, Iss. 6, pp. 429–431 (1981).
22. Breakdown characteristics in InP/InGaAs avalanche photodiode with p-i-n multiplication layer structure, Hyun, Kyung-Sook; Park, Chan-Yong; Journal of Applied Physics, vol. 81, Iss. 2, pp.974–984 (1997).
23. The Art of Electronics — Horowitz & Hill.
24. Guo-Feng Zhang, Shuang-Li Dong, Tao Huang, Yuan Liu, et al. “Photon Statistical Measurement of Afterpulse Probability”, International Journal of Modern Physics B, Vol. 22, No. 12 (2008) 1941-1946.
25. D. P. Donovan, J. A. Whiteway, and A. I. Carswell, "Correction for nonlinear photon-counting effects in lidar systems," Appl. Opt. 32, 6742-6753 (1993)
26. SPCM APD Perkin Elmer Specifications at *www.**perkinelmer**.com.*
27. NOTES AND CORRESPONDENCE. Open Access Micropulse Lidar Signals: Uncertainty Analysis. Ellsworth J. Welton and James R. Campbell. 2089–2094
28. Campbell, J. R., D. L. Hlavka, E. J. Welton, C. J. Flynn, D. D. Turner, J. D. Spinhirne, V. S. Scott, and I. H. Hwang, 2002: Full-time, eye-safe cloud and aerosol

lidar observation at Atmospheric Radiation Measurement Program sites: Instrument and data processing. J. Atmos. Oceanic Technol., 19, 431–442.

Appendix: HSRL Specification

The lidar operates tuned to the center of iodine absorption line 1109. The transmission of Brillouin scattered light through the edge of line 1105 is used to lock the laser wavelength [10].

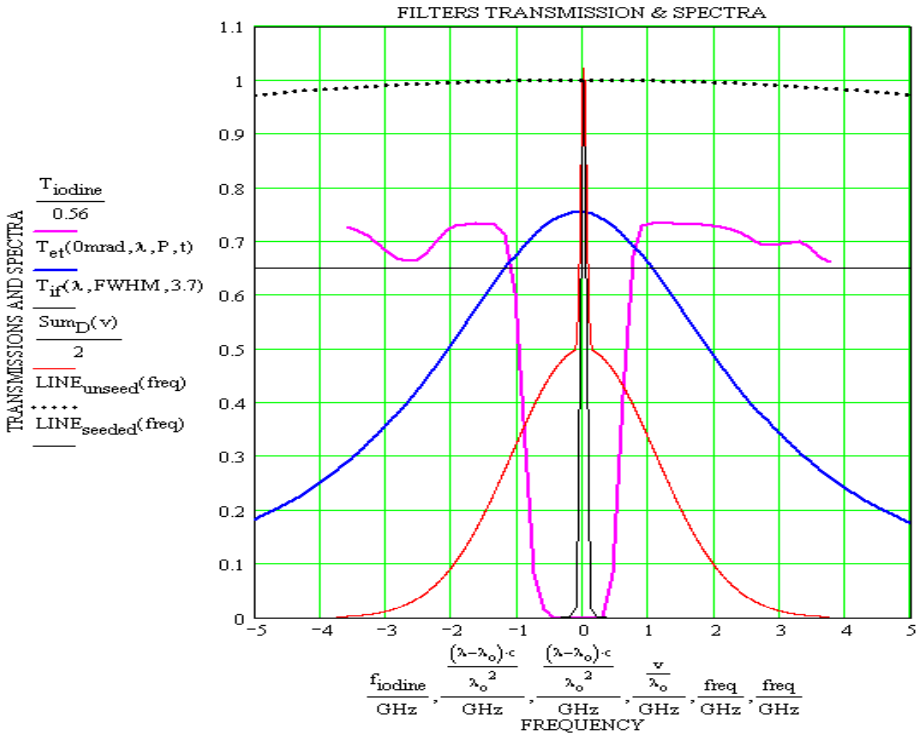

Figure 29. Filters' transmission and spectra. Detected signal spectrum (red curve), Transmission spectrum of etalon (blue curve), Iodine filter transmission spectrum (pink curve), Interference filter spectrum (black dotted curve).

Table 2: HSRL specifications [5]

Aperture	40 cm
Angular field-of-view	100 μrad
Average power	350 mW
Pulse repetition rate	4 kHz
Wavelength	532 nm
Laser bandwidth	<100 MHz
Detectors APD based SPCM	QE ~50%
Sky noise bandwidth	~ 8 GHz
Iodine filter bandwidth	1.8 GHz
Range resolution	7.5 m