Departament de Teoria
del Senyal i Comunicacions

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Doctoral Thesis

DSP-BASED CW LIDARS FOR CLOUDS AND AEROSOL

Author: Antoni Ardanuy Dellà

Thesis Advisor: Adolfo Comerón Tejero

Department of Signal Theory and Communications

Universitat Politècnica de Catalunya

CONTENTS

1. INTRODUCTION

Lidar is a word originated by the substitution of *radio* with *light* in *radar*. The operation principle of lidars is the same as of radar, and sometimes the name radar laser is used too. The difference is that in lidar the electromagnetic waves in the so-called radioelectric part of the spectrum are replaced by waves in or close to the optical range, specifically using laser light. The first idea of these systems was based on the same principle as radars and it continues being operative today. Using a short and high-powered pulse of light, and measuring the time of flight of this pulse echoed on the target (from the emitter to the target and returned to the receiver), and the received intensity from it, it is possible to determine the distance to the target and its reflectivity. Systems operating in this way are relatively easy to design, but also have some disadvantages, such as the high power per pulse that this kind of design needs (especially for atmospheric sounding) and the relative inefficiency of use of time. Actually, the whole amount of transmitted energy is concentrated in some nanoseconds, whereas the repetition period ranges are much larger, resulting in a low duty cycle. Moreover, high-power pulsed lasers tend to be bulky and heavy to achieve the necessary energy per pulse and the end-to-end energy efficiency is not very high. Laser diodes can reach efficiencies as high as 50% meanwhile in solid state lasers (no semiconductor) the efficiency can be as low as 1-2%.

Despite these drawbacks, lidars using pulsed lasers are widespread in atmospheric remote sensing (and in particular in aerosol profiling and cloud detection) because of the availability and technological maturity of pulsed lasers with high energy per pulse. Pulsed lidars are also used from spaceborne platforms to perform measurements of atmospheric aerosols and clouds — as the CALIOP instrument[1] on board of CALIPSO satellite or CATS (NASA)[2] on the International Space Station , and future EarthCARE (ESA)[3] —, to determine the height of the ice caps —ICESAT[4][5], and to measure wind velocities — AEOLUS (ESA)[6]. The idea of lidars in space has been developed from years ago[7][8] .

Another issue with the use of high peak power lasers, which limits their use in some environments, is the necessary care to avoid possible damage to the eyes (non eye-safe systems). Sometimes a solution is to expand the laser beam with an optical system to obtain a lower power density and to have a more eye-safe system.

A different option in a lidar system is to distribute the energy along time, that is, to send a sequence of pulses of lower power over a longer time. The amount of energy sent can be high but the peak power is many times lower than in the pulsed lasers. These lidars, with low peak power lasers, have two possible designs:

- High Pulse-Repetition-Frequency lidars. This is a version of pulsed lidar, but with low energy per pulse and using a high Pulse Repetition Frequency (PRF). These lidars still have a higher but still very low duty cycle (transmitted pulse duration of some ns with a repetition period in the order of hundreds of μs). The PRF can be as high as possible as permitted by the maximum non-ambiguous distance measurable. The advantage is that it is possible to reduce the laser peak energy (and consequently the size and weight) compared to standard pulsed lidars.
- Continuous Wave Lidars (CW lidars). Another option is to send energy almost continuously in time. This requires a modulated signal designed to recover the information of the target range. The energy is distributed along time (along the spectrum in transformed domain depending on the modulating signal), using a special coding system. Spreading on the time means a low peak power, tending to an eye-safe system because of the design itself.

This thesis focuses on the research on lidars using a CW laser modulated by a specific sequence as a possible alternative to some pulsed lidars. These CW lidar systems can be based on semiconductor lasers (laser diodes) that are small, efficient and relative simple to use when it comes to the design of the power supply and the control of the laser, leading to robust systems [9]. One key issue in these systems is the use an adequate signal coding to recover the desired information about the target (the atmosphere backscatter profile in our case). Ideally one is interested in modulating the laser using pseudo-random sequences with a similar number of ones and zeros (to maximize the average transmitted power) and with a delta-like autocorrelation function. Takeuchi et al.[10], [11] begun the first studies on the use of a continuous laser emission modulated with a selected sequences and they showed their potential practical use. The first proposals of the systems were based on the hardware available in those years (early 80es), with limitations to capture the signals and to calculate correlations between long sequences.

The overarching goal of this thesis is to explore the limits of such systems for atmospheric aerosol and cloud profiling using up-to-date laser diode transmitters, photoreceivers and digital signal processors (DSP), yet under practical restrictions of limited laser peak power and of internal-noise limited photoreceivers. This results in a study of adequate pseudo-random modulating sequences (essentially M-sequences and other related sequences) and in the design, construction and operation of a demonstration prototype with small size and weight, very low power consumption and a

configurable, real-time signal processing, thanks to the use of a DSP controlling the system and processing the signals, allowing flexibility in implementing different types of modulating sequences and signal recovering algorithms. Receiver is based in an Avalanche PhotoDiode (APD). Partial objectives are:

- Investigation of the features of some classes of pseudo-random sequences and their use in this kind of lidar systems, their advantages and drawbacks.
- Design and implementation of an operating prototype using in the transmitter a laser diode modulated by pseudo-random sequence, an APD-based receiver, and a DSP as processing and control core.
- Evaluation of the performance of the implemented system in real measurement conditions.

The thesis manuscript is organized as follows:

- A summary of previous published work on this kind of systems and obtained results. Theoretical background and determination of the ideal capabilities of the system under investigation. (chapter 2)
- Study of the modulating sequences and their properties and drawbacks (chapter 3).
- Development of a software tool allowing simulate the system operation before proceeding to the implementation (chapter 4).
- Design and implementation of a prototype following the findings and guidelines exposed in the previous chapters (chapter 5 and 6).
- Test of the implemented prototype in real atmospheric conditions to determine its capability to measure aerosol backscatter profiles and cloud range (chapter 7).

The architecture of the system has been granted a patent [reference] and the implementation and preliminary operation results have originated a publication in an indexed journal [reference]. Moreover, the system is operated on a regular basis and new measurements are continuously obtained.

2. PRINCIPLES OF A LIDAR SYSTEM

The principles of the lidar system are explained in this chapter, highlighting the differences between pulsed and pseudo-random sequence modulated CW lidar.

2.1. GENERAL PRINCIPLES OF LIDAR SYSTEMS AND THE BASIC LIDAR EQUATION

Let $P_0 s(t)$, where P_0 is the peak power and $s(t)$ a modulating signal $(0 \le s(t) \le 1)$, the power signal emitted from a light source like a laser and sent to the atmosphere. This signal is attenuated and scattered by the particles and gases and a part of the emitted power is returned to the origin from a distance z. The amount of the power returned to the emitter related to the total power sent depends on the backscatter coefficient $\beta(z)$ and the extinction coefficient $\alpha(z)$(in this work only mono-static lidars are studied). Then a generic lidar equation giving the received power at time t, $P_R(t)$ is:

$$P_R(t) = P_0 A_R \int_0^\infty s\left(t - \frac{2z}{c}\right) \frac{\beta(z)}{z^2} e^{-2\int_0^z \alpha(x)dx} dz, \qquad (2.1)$$

where c is the speed of light and A_R is the receiver surface.

This equation can be used for any modulating signal *s(t)* and specifically it can be reduced to the classical equation of the pulsed lidar or to other systems as the CW lidar. In this equation the effects of incomplete overlap between the transmitted beam and the receiver field of view are not considered, which would appear as a function of z comprised between 0 and 1 multiplying the received signal.

The pulsed lidar

If the modulating signal ***s(t)*** is a single pulse ***p(t),***

$$p(t)=\begin{cases}0 \text{ for } t<0 \text{ and } t>\tau \\ 1 \text{ for } 0<t<\tau\end{cases} \tag{2.2}$$

Eq. (2.1) becomes the classic equation of a pulsed lidar:

$$P_R(t)=P_0 A_R \int_{ct/2}^{c(t+t)/2} \frac{\beta(z)}{z^2} e^{-2\int_0^z \alpha(x)dx} dz\,, \tag{2.3}$$

which, assuming that the change of the optical coefficients of the atmosphere, $\alpha(z)$ and $\beta(z)$, is negligible in a range interval $\Delta z = c\tau/2$, and making the change of variable $z = ct/2$, can be cast in the usual way as power received from range z:

$$P_{received}(z)=P_0 \cdot A_R \cdot \frac{c\tau}{2z^2} \cdot \beta(z) \cdot e^{-2\int_0^z \alpha(x)dx} \tag{2.4}$$

The power returned from a distance z is the energy backscattered by a slice of the atmosphere after undergoing extinction in the round-trip path of the light to range z and back to the receiver. The volume of this slice is defined by the diameter of the light pulse —which depends on the aperture angle of the light source and the distance from the source to the considered point— and the range resolution of the system $c\tau/2$ (Fig. 2.1), being τ the pulse time duration. This figure also shows in a qualitative manner the cause of the overlap function mentioned before: there is no full overlap between receiver field of view and the illuminated scattering volumes until a certain range.

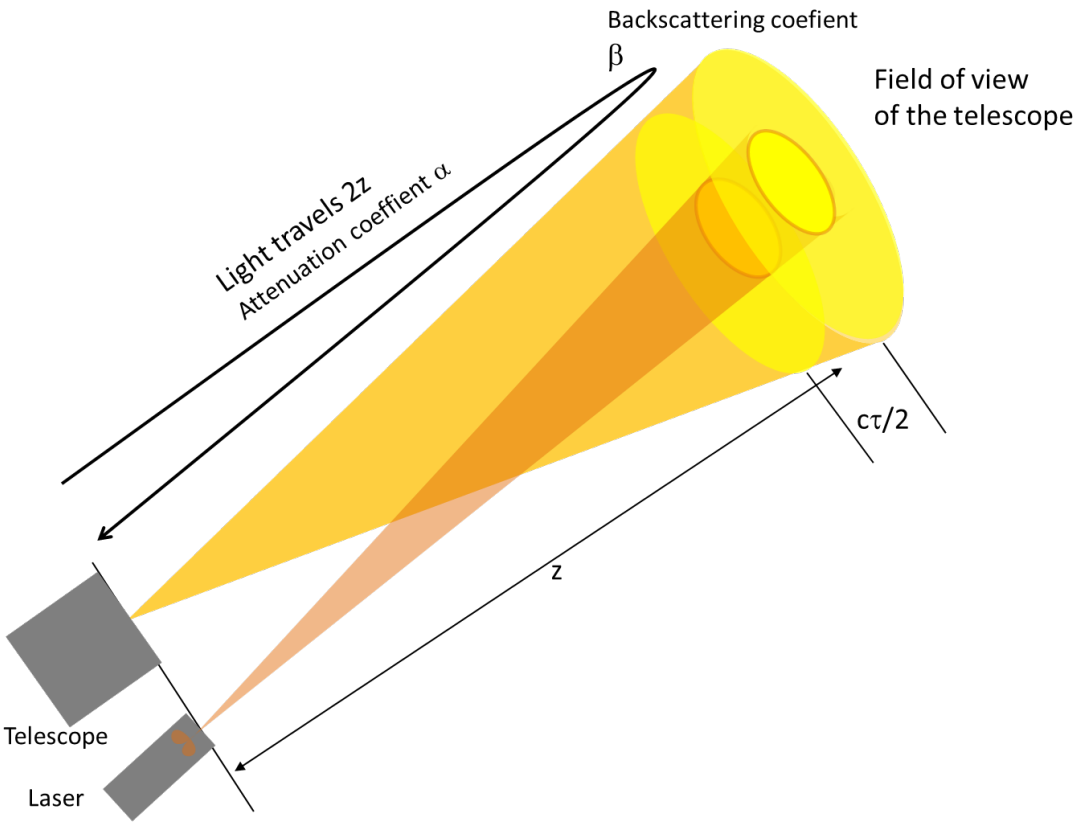


Fig. 2.1. Schematic view of the laser beam and the superposition to the field of view of the telescope.

The length of the light pulse fixes the limit of the range resolution of the system at the value $c\tau/2$. Higher duration for the same power means more energy and therefore more signal-to-noise ratio, but less spatial resolution at the same time.

An important characteristic of the lidar system is that the returned power varies as the inverse squared distance. The reason is that in a lidar system the cross-section of the target is the whole illuminated area of the atmosphere.

The final primary result of the lidar measurement is the retrieval of the profile:

$$h(z) = \frac{\beta(z) \cdot e^{-2\int_0^z \alpha(x)dx}}{z^2}, \tag{2.5}$$

And multiplying by z^2 the attenuated backscatter profile, defined as

$$Att_{bck}(z) = \beta(z) e^{-2\int_0^z \alpha(x)dx} = z^2 h(z) \tag{2.6}$$

is obtained. From another point of view, if a pulse is sent through the atmosphere, the lidar is measuring the impulse response *h*(*z*) of this path in the atmosphere, considering that the pulse is narrow enough to be treated as a delta function. Usually the laser pulses are very narrow, in the order of nanoseconds. A 100 ns pulse is like a 15-m thick bunch of photons traveling through the sounded path.

The lidar equation in a pseudo-random modulation CW lidar

In the case of a sequence-modulated lidar the equation to apply is the general case expressed by Eq. (2.1), which gives the time dependence of the received power .

From now on the use of distance or time variables in the case of signals will be applied indistinctly because one can be transformed in another using the light speed constant.

Let's assume that it is possible to find a modulating function s(t) that satisfies

$$R_{ss}(t_0,t_1)=R_{ss}(t_0-t_1)=\int_{-\infty}^{\infty}s(t-t_0)s(t-t_1)\,dt=K\delta(t_0-t_1) \tag{2.7}$$

where $\delta(t)$ is the Dirac's delta function and K has units of s^2. Applying this function to eq. (2.1) and doing some operations the next result is obtained:

$$R_{P_Rs}(t_0)=\int_{-\infty}^{\infty}P_R(t)s(t-t_0)\,dt=\frac{cP_0A_RK}{2}\cdot\frac{\beta\left(\frac{ct_0}{2}\right)}{\left(\frac{ct_0}{2}\right)^2}\cdot e^{-2\int_0^{ct_0/2}\alpha(x)dx} \tag{2.8}$$

This equation shows that the correlation of the received signal with the function s(t) is proportional to the attenuated backscatter from the different ranges.

Let us now assume a periodic modulating signal $s(t)$ of period T and the correlation property

$$R_{ss}(t_0)=\int_t^{t+T}s(t')s(t'-t_0)\,d\tau\approx\begin{cases}R_0 \text{ for } -\dfrac{\tau}{2}<t_o-mT<\dfrac{\tau}{2}\\ 0 \text{ for any other } t_0\end{cases}$$

or alternatively (2.9)

$$R_{ss}(t_0)=\int_t^{t+T}s(t')s(t'-t_0)\,dt'\approx R_0\Pi\left(\frac{t_0-mT}{\tau}\right)$$

where R_0 has time units and $\Pi\left(\frac{t}{\tau}\right)$ is the rectangular pulse function defined as

$$\Pi\left(\frac{t}{\tau}\right)=\begin{cases}1 \text{ for } -\tau/2<t<\tau/2\\ 0 \text{ elsewhere}\end{cases} \tag{2.10}$$

Then, with $P_R(t)$ given by Eq. (2.1), we have

$$R_{P_R s}(t_0)=\int_0^T P_R(t)s(t-t_0)dt = P_0 A_R \int_0^T \int_0^\infty s\left(t-\frac{2z}{c}\right)\frac{c\tau\beta(z)}{2z^2}e^{-2\int_0^z \alpha(x)dx}s(t-t_0)dz\,dt$$

(2.11)

which taking into account Eq. (2.9) yields

$$R_{P_R s}(t_0)=\int_0^T P_R(t)s(t-t_0)dt \approx P_0 A_R R_0 \frac{c\tau}{2}\sum_{m=0}^{\infty}\frac{\beta\left(c\frac{t_0+mT}{2}\right)}{\left(c\frac{t_0+mT}{2}\right)^2}\exp\left(-2\int_0^{c\frac{t_0+mT}{2}}\alpha(x)dx\right)$$

(2.12)

Assuming that $\beta(z)/z^2$ is negligible for $z>z_{\max}$, and $T>\frac{2z_{\max}}{c}$, Eq. (2.12) reduces to

$$R_{P_R s}(t_0)=\int_0^T P_R(t)s(t-t_0)dt \approx P_0 A_R R_0 \frac{c\tau}{2}\frac{\beta\left(c\frac{t_0}{2}\right)}{\left(c\frac{t_0}{2}\right)^2}\exp\left(-2\int_0^{c\frac{t_0}{2}}\alpha(x)dx\right) \tag{2.13}$$

In practice we will deal with discrete signals $s(n\tau)$, with the consideration that the sampling interval is equal to the bit duration τ and n denotes the sample number. To simplify the notation, we will use $s(n)$ instead of $s(n\tau)$, $s(n)\leftarrow s(n\tau)$, so that the time τ is implicit. Assuming that the period of the signal is $T=N\tau$, the correlation property given by Eq. (2.9) can be written

$$R_{ss}(n_0)=\sum_{m=1}^{N}s(n)s(n-n_0)=\begin{cases}R_0 \text{ if } n_0=mN\\ 0 \text{ for any other } n_0\end{cases}. \tag{2.14}$$

The discrete form of the correlation of $P_R(t)$ with $s(t)$ given by Eq. (2.12) becomes

$$R_{P_R s}(n_0) = \sum_{n=1}^{N} P_R(n)\, s(n - n_0) = \frac{P_0 A_R R_0}{\left(\frac{c\tau}{2}\right)^2} \frac{c\tau}{2} \frac{\beta(n_0)}{(n_0)^2} \exp\left(-2\sum_{m=1}^{n_0} \alpha(m)\right), \qquad (2.15),$$

where $\beta(n) \leftarrow \beta\left(c\frac{n\tau}{2}\right)$, $\alpha(n) \leftarrow \alpha\left(c\frac{n\tau}{2}\right)$, $P_R(n) \leftarrow P_R(n\tau)$, and it has been assumed that $\frac{\beta(n)}{\left(\frac{c\tau}{2}n\right)^2} = 0$ for $n > N$.

Actually, there exist sequences [10]–[12]based on pseudo-random binary data that satisfy the property of having a delta-like correlation as expressed in eq. (2.14). The most known application of this property is found in Spread Spectrum communications[13]. The next chapter is devoted to describe and study the application of these sequences in lidar systems.

The profile $h(z)$ defined in eq.(2.5) can be retrieved using the correlation operation of Eq. (2.15). And multiplying $h(z)$ by z^2 the attenuated backscatter profile defined in eq. (2.6) is obtained.

Therefore, if the correlation operation is applied to the received signal (an attenuated and delayed version of the modulating sequence), the impulse response of the channel is retrieved.

Digital Signal Processors (DSP) are optimized to perform mathematical operations related to signal processing tasks and this feature can be used to reduce the processing time. This time is potentially so short that it is possible to use long sequences —it will be shown in the chapter 3 dedicated to sequences that the use of long sequences is beneficial to the system performance. In fact, the correlation process is so fast that most of time the processor is dedicated to acquire the received signal. Therefore, these features lead to a very compact lidar.

As a summary, if the impulse response $h(z)$ of the atmosphere were a delta function, the received sequence would be strictly an attenuated and delayed version of the sequence transmitted. In this case, it can be easily shown that the correlation operation of the received signal with the sequence modulating the transmitted power returns the amplitude —related to the backscattering profile— and the delay of the signal —related to the distance to the target. If the impulse response of the atmosphere is a generic function $h(z)$ the result of correlating the two sequences (the transmitted one and the received one) is the function $h(z)$, from which the attenuated backscatter is retrieved.

In view of the above considerations, the design of the used modulating sequence is critical. A perfect retrieval of the attenuated backscattering profile forces the autocorrelation function of the sequence to be a delta function. If the autocorrelation function is not a delta function —and the result is wider than the thinnest structure of the profile being measured— the recovered profile is blurred. There is the possibility that the weakest signals returned could be masked when they are beside a strong signal. In an atmospheric lidar this would be the case of an aerosol layer and the presence of a cloud behind it. Figure 2.1 depicts this effect in an idealized scenario. In fig. 2.1 (a) the result of the correlation operation is shown for a target placed at a distance $cn_1t_s/2,$ with $n_1 = 41$ and the modulating sequence having a delta-like autocorrelation behavior; fig. 2.1 (b) shows the same assuming an additional stronger target is located at $cn_2t_s/2,$ with $n_2 = 44$. Fig. 2.1 (c) shows the weak target being masked by the strong one when the modulating sequence does not have a perfect delta-like autocorrelation property.

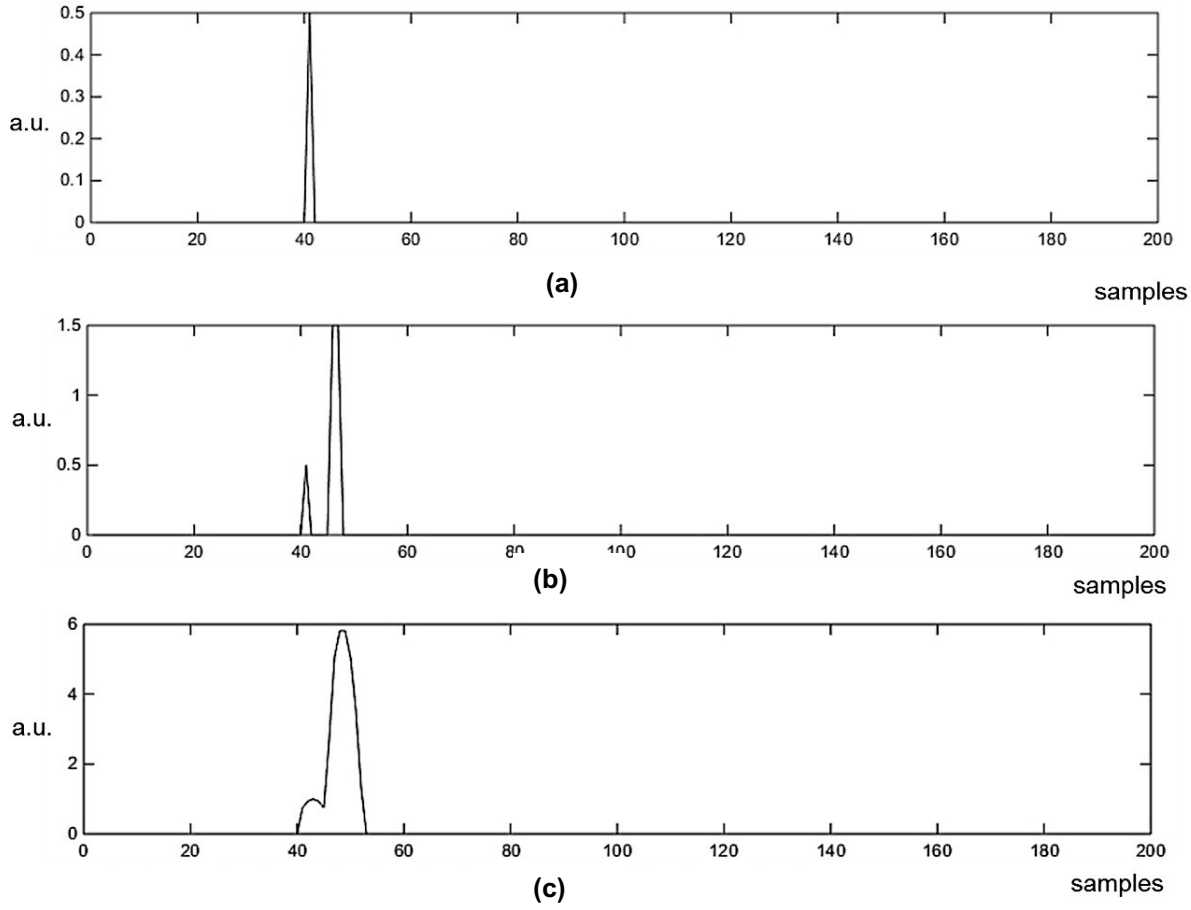


Fig. 2.1 The effect of not having a delta function correlation of the sequence used: a thin target is visible (a) or two closed targets (b) are visible if the sequence correlation function is a delta. If the correlation is wider the two targets can be confused in one only (c). The vertical scale of graphics is in arbitrary units (a.u.).

2.2. CW LIDAR WITH PSEUDORANDOM SEQUENCE

Focusing attention in the specific case of the CW lidar modulated with a Pseudorandom sequence, in the next sections there is a more specific general analysis of the expected behavior of this system. The deep analysis of the chosen sequences, its behavior and the advantages and disadvantages of each one will be explained in the next chapter.

Modulating sequence

The idea of using a CW laser modulated with a random sequence was proposed by some researchers [10]–[12] and based on this idea, they made prototypes to carry out experiments demonstrating the possible use of this type of modulation. This kind of sequences have been long used in communications, but also for other applications like the evaluation of impulse response in audio systems, response of neurosensory, on-chip testing of MEMS, linear-system identification, etc. [14]–[22]. Chapter 3 is widely dedicated to explore the sequences and their behavior. Another characteristic that makes attractive the use of this kind of sequences is that they are binary, composed by a set of 0s and 1s only, which opens up the use of digital systems to modulate the laser with a great simplification of the emitter system.

An important issue in the design of the lidar system is the choice of the modulating sequence s(t). To recover the atmospheric profile through the estimation of h(t), the autocorrelation of the sequence must approximate a delta function. Otherwise, the recovered profile is “blurred” and strong returns can mask weak ones, as would be the case of a cloud behind a light aerosol layer.

The use of pseudo-random M-sequences to modulate the laser transmitted power is suitable to address this issue, because of: a) the delta-like character of their correlation with the derived sequences as it will be discussed in chapter 3; b) the fact M-sequences have almost the same number of 1s and 0s (one 1 in excess over the number of 0s), which, for a limited transmitted peak power, maximizes the transmitted energy in a given time interval for a given range resolution.

Range resolution and maximum non-ambiguous distance

Assuming the modulation sequence is a M-sequence or related sequence (see chapter **0**), the correlation function is a delta-like function of width τ , the sequence chip duration. Therefore, the range resolution ΔR is defined exclusively by the value of τ and given by

$$\Delta R = \frac{c\tau}{2} \qquad (2.16).$$

The maximum unambiguous distance —defined as the maximum distance at which a target can be located without confusion in distance with another farer target (or in other words not information aliasing)— is related to the sequence length and pulse duration through

$$R_{ua} = \frac{c\,\tau\,N}{2} \qquad (2.17)$$

Therefore, the sequence length must be set to a minimum value of

$$N_{\min} = \frac{2R_{\max}}{c\,\tau} = \frac{R_{\max}}{\Delta R} \qquad (2.18)$$

where R_{max} is the maximum range from which non-negligible returns are expected.

Values of *N* above N_{min} provide a safety margin against unexpectedly high returns (for instance, from clouds) coming from ranges above the nominal one. This is achieved at the price of increase the processing time. It could be thought that using sequences longer than N_{min} would put an unnecessary (and eventually unaffordable) burden on the processing task, as the number of operations to perform the correlation of two sequences of length *N* is proportional to N^2. However, as no signal is expected from $R > R_{max}$, the correlation needs be computed only for N_{min} points, which makes the computation time proportional to *N* —*N x* N_{min} operations required. This indicates that, although equivalent from the point of view of signal-to-noise ratio, using longer sequences is better than mere accumulation of the processing result of shorter ones, because of the added immunity to echoes from ranges beyond the nominal one that would be indistinguishable from echoes within the nominal range. Another advantage of using long sequences is the increasing immunity in presence of non-compensated offset values.

Signal-to-noise ratio

The signal-to-noise ratio (SNR) in a photoreceiver operating in analogue mode is [23]–[25]

$$SNR = \frac{P_S}{\sqrt{P_q\left(P_S + P_b\right) + NEP^2 B}}$$

(2.19)

where P_s and P_b are respectively the optical power of the signal and of background radiation reaching the photodetector, *NEP* is the photoreceiver noise equivalent power, *B* is the photoreceiver electrical bandwidth and

$$P_q = \frac{2FhcB}{\eta\lambda}, \tag{2.20}$$

with *F* the excess noise factor of the photodetector, *h* the Planck's constant, *c* the speed of light, η the photodetector quantum efficiency, an λ the wavelength.

In absence of background radiation, the optical signal power at which the operation regime of the photoreceiver changes from signal-shot-noise limited SNR to NEP-limited SNR is [23]

$$P_{sk} = \frac{NEP^2 \eta \lambda}{2Fhc}. \tag{2.21}$$

Figure 2.2 shows the behavior of the SNR in eq. (2.19) as a function of P_S. The asymptotic curve shows the point P_{sk}, where the response of the SNR changes. When $P_S << P_{sk}$, the SNR is

$$SNR = \frac{P_S}{NEP\sqrt{B}} \tag{2.22}$$

and the system is dominated by additive photoreceiver noise. When $P_S >> P_{sk}$ the expression of the SNR is

$$SNR = \sqrt{\frac{P_S}{P_q}} \tag{2.23}$$

and the system is dominated by signal shot noise.

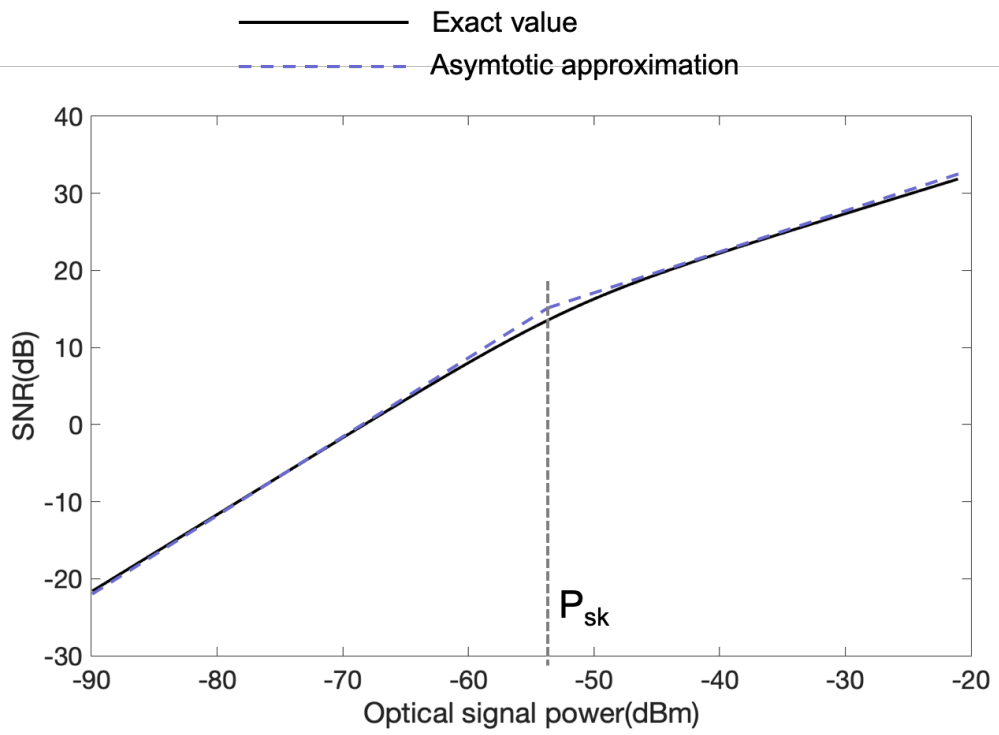


Fig. 2.2. SNR as a function of the received optical signal power according to* (2.19)*. In this graphic the values are NEP=1.2e-13WHz1/2, B=1.5e6Hz, F=3.45, η=0.78 and λ=785nm.

We can assess a typical value of P_{sk} by using the parameters of the APD-based photoreceiver we will deal with in chapter 5, section 5.5. For the used photoreceiver (see Table 5.2) NEP=1.2 10^{-13} $WHz^{1/2}$, and with the values for its APD, F=3.45 and η=0.78 at λ=785 nm, we find P_{sk}=6.45 nW.

For an atmosphere with typical optical coefficients and small or medium-sized telescopes, the optical power in the photoreceiver is always well below this value if the transmitter is a laser diode with low peak power. For example, assuming a peak laser power of 0.125 W, a telescope area of 0.03 m^2, $c\tau/2 = 72\,\mathrm{m}$, $\alpha = 2\times10^{-5}\,\mathrm{m^{-1}}$, $\beta = 2\times10^{-3}\,\mathrm{m^{-1}sr^{-1}}$ (a favorable situation with clear atmosphere and a cumulus cloud at 700m of altitude that means one of the most intense signals received). The value of the power received is 1nW, below P_{sk}.

Note that background radiation in the (2.19) would entail an apparent increase in *NEP*. Calling *NEP'* the effective noise equivalent power with background radiation, we have

$$NEP'^2 = NEP^2 + \frac{2FhcP_b}{\eta\lambda}. \qquad (2.24)$$

If P_s is the power received from a certain range and L independent returns from that range are accumulated, the resulting signal-to-noise ratio *SNR$_{ac}$ is*

$$SNR_{ac} = \sqrt{L}\, SNR. \qquad (2.25)$$

A system like this the correlation with a sequence containing N 1's is equivalent to cumulate N pulses from the point of view of noise. As it will be shown in chapter 4.3, in a lidar modulated by an M-sequence of length N, the signal-to noise ratio at the correlator output is equivalent to that of having accumulated $\frac{(N+1)^2}{4N}$ returns produced by transmitted pulses of energy $P_0\tau$, remembering that τ is the duration of a bit in the sequence. Obviously, if $N >> 1$, $\frac{(N+1)^2}{4N} \approx \frac{N}{4}$.

The equation (2.21) fixes a turning point (P_{sk}) about the behavior of the system. With optical power values below it (on the left side of this point in the graphic of fig. 2.2) the receiver is working limited by the system noise and therefore the increase of the SNR is lineal with the increase of the optical power. With values greater than P_{sk} (on the right side of this point the receiver is limited by the shot noise and therefore the increase of the SNR depends on the square root of the optical signal. At the same time the received optical signal depends on the energy per pulse (bit) and the PRF (number of bits per second). Therefore, to work in the limited shot noise region it is necessary to use a high peak power laser.

From a different point of view, when the power received is greater than $P_{sk,}$ if this power decreases the SNR decreases with the square root of it but when the power received is below P_{sk} then the SNR decreases lineally with the this received power. This can explain why when the received signal decreases under a limit (P_{sk}) the SNR decreases fast. Using a low power laser, the risk is to be below this limit when the backscattered signal is weak.

Power-range resolution trade-off

This is a short evaluation about the trade-offs of power and range resolution and the section 4.4 contains a more detailed evaluation.

The most important constraint in this system is the low energy per transmitted pulse (using the parameters of our system, with a laser of 125 mW, τ=520 ns the energy sent per pulse is 6.5 x 10^{-8} J.). For a system working in the SNR regime dominated by additive receiver-noise ($P_s << P_{SK}$), a transmitter merit factor, related to the capability of achieving a given SNR in given accumulation time, can be defined as $F_{me} = E\sqrt{PRF}$, where E is the energy per pulse and PRF is the effective pulse repetition frequency. In our system the effective pulse repetition frequency is $\frac{1}{2\tau}$, where τ is the bit duration. Considering that the energy per pulse and the spatial resolution are related by

$$E = P_0 \tau = \frac{2P_0 \Delta R}{c} \tag{2.26}$$

where $\Delta R = \frac{c\tau}{2}$ is the spatial resolution, P_0 the pulse power, and c the speed of light, the figure of merit can be written as

$$F_{me} = P_0 \sqrt{\frac{\Delta R}{c}} \tag{2.27}$$

This relationship shows that if the laser power is divided by a given value, to maintain the merit factor the resolution must be reduced by the square of this value.

This is a drawback in aerosol detection, especially for long-range detection, but the use of a low peak-power laser is nevertheless a practical option to increase the laser life, to reduce the cost of the laser component and the power supply, and for eye-safety reasons.

3. SEQUENCES IN THE MODULATION OF THE LIDAR

This chapter describes and analyses the pseudorandom sequences used in the modulation of the lidar system. We will rely on sequences developed for other purposes in communications systems.

Among the possible collections of pseudorandom sequences, the subset of sequences of interest for a lidar system would have to meet the next main features:

a) Correlation properties: the autocorrelation is a delta function or delta-like function.

b) Possibility to have sequences of the same length (quasi) orthogonal between them.

c) The number of 1's and 0's be the same more or less to maximize the transmitted energy with a fix mean power of the laser.

Property (b) would be useful in systems using several lasers (e.g. lasers emitting at different wavelengths in a system aimed at gas detection), where it could be possible to modulate the different lasers with different sequences (preferably of the same length) and use only one photoreceiver, the separation of the returns from the different lasers being made through the correlation of the photoreceiver output with the different modulating sequences.

The effect of possible non-linearity in the receiver chain will also be considered.

Usually in a pulsed lidar system the effect of non-linearity in the receiver chain is a local effect (where the signal is very intense), but when pseudorandom sequences are used instead of a pulse, some considerations have to be done because the effect is not negligible in these sequences as it will analyzed in this chapter (sections 0 and 3.3).

3.1. PSEUDO-RANDOM SEQUENCES: M-SEQUENCES AND DERIVED

Binary pseudo-random sequences satisfying properties similar to those of random processes are generated by means of special polynomials. Therefore, the real generated sequences are pseudo-random. In fact, the only necessary condition in the lidar system is that the autocorrelation behaves like a delta function (see sections 0 and 0).

The theory of pseudo-random sequences is well established and it is not necessary to discuss it here. It is possible to find a vast bibliography, of which [26]–[30] is only a small but representative sample, about pseudorandom sequences and their properties, because of the long-time since their development and the different applications found for them, for example in communications systems. Figure 3.1 shows a general and simple layout of a pseudorandom sequence generator based on a shift-register as a class of cyclic codes generated with shift-registers [31]. The different positions of the feedback links from the shift register to the adder changes the generated sequence maintaining the same length. The number of stages of the shift register fixes the sequence length. A greater number of stages creates a longer sequence and as consequence the set of the possible sequences of the same length increases.

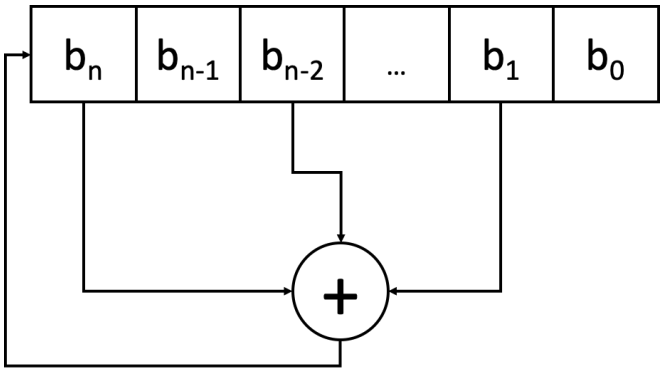


Fig. 3.1. Simple schema of a pseudorandom sequence generator based on a shift register and an addition operation. The different position of the feedback links changes the generated sequence . The number of stages changes the sequence length.

Definition of the M sequences[26]–[29]

The base sequences we focus our study on are the M sequences (or Maximum Length Sequences) [28], [29]. This name corresponds to a family of sequences that are generated by a shift-register of length L when the feedback connections represents a primitive polynomial. In this case the name of Maximum comes from the apparition of all the possible states in the shif-register except the state all zeros while the register is generating the sequence. The presence of all-zero values state would be the stop of the generator. An M-sequence is a periodic pseudorandom train of 0's and 1's of period $N=2^{L-1}$:

As a register of length **L** has $\mathbf{2^L}$ states, the remaining state is the 0, 0, 0…, that is, the trivial sequence that cannot be generated because if the register crosses this state it stops.

Some properties of M sequences, which make them interesting for an atmospheric lidar, are [26], [29]:

1. An L-bit register generates a set of M-sequences of period 2^{L-1} bits.
2. An M sequence of period N contains N/2 ones and N/2-1 zeros exactly. This is an advantage with respect to other families of orthogonal sequences (for example Optical Orthogonal Codes with more 0's than 1's)[32], as it maximizes the average transmitted power for a limited peak power.
3. If we define a run of length r as a subsequence of r consecutive 0's or 1's, then in an M-sequence:
    a. 1/2 of the runs are of length 1.
    b. 1/4 of the runs are of length 2.
    c. 1/8 of the runs are of length 3.
    d. …
    e. $1/2^k$ runs are of length k.
4. The modulo-2 addition of one sequence delayed $\mathbf{m_1}$ and the same sequence delayed $\mathbf{m_2}$ creates the same sequence delayed $\mathbf{m_3}$ where m3 can be any value < N.

An M-sequence with values (0,1) has a circular autocorrelation function as

$$r_{ss}(j) = \frac{1}{N}\sum_{i=0}^{N-1} s(i)\cdot s(i+j) = \begin{cases} (N+1)/2N & j=0 \\ (N+1)/4N & j\neq 0 \end{cases}, \tag{3.1}$$

as it is shown in the figure 3.2 . The correlation is periodic of period N

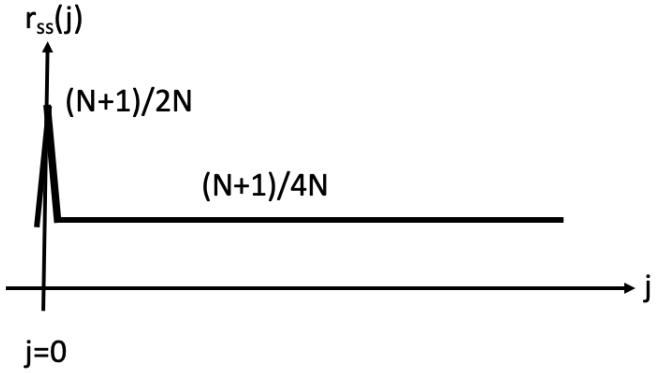


Fig. 3.2. Autocorrelation of the original function s(n) with values (0,1).

These sequences are very easy to generate but they have a value greater than 0 in the autocorrelation function outside the maximum at j=0. This is not strictly a Kronecker's delta function and this pedestal value can generate problems of masking when strong echoes are present mixed with other weak echoes in the same channel.

To solve this problem, some authors [10],[11] proposed a variation based in this family of M sequences. If **s** is the original M-sequence, the new sequence is:

$$s' = 2s - 1 \tag{3.2}$$

This new sequence is still a binary sequence but its values are (+1, -1) resulting of mapping the values (0,1) to (-1,1), Then the autocorrelation function is (fig. 3.3)

$$r_{s's'}(j) = \frac{1}{N}\sum_{i=0}^{N-1} s'(i)\cdot s'(i+j) = \begin{cases} 1 & j=0 \\ -1/N & j\neq 0 \end{cases} \tag{3.3}$$

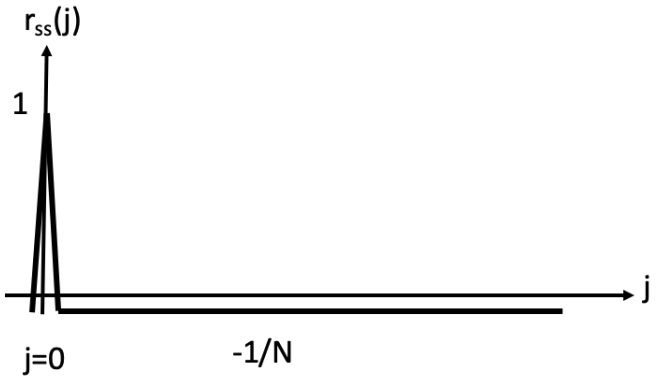


Fig. 3.3. Autocorrelation of the function s'(n).

The longer is N, the smaller is the residual term -1/N and this implies that N would have to be high (long sequence).

The residual term can be removed if instead of correlating **s'** with itself, it is correlated with the original sequence **s**. In that case the result is strictly a delta function. The cross-correlation function of **s** and **s'** is

$$r_{ss'}(j) = \frac{1}{N}\sum_{i=0}^{N-1} s(i)\cdot s'(i+j) = \begin{cases} \dfrac{N+1}{2N} & j=0 \\ 0 & j\neq 0 \end{cases} \tag{3.4}$$

resulting in a delta-like function (fig. 3.4).

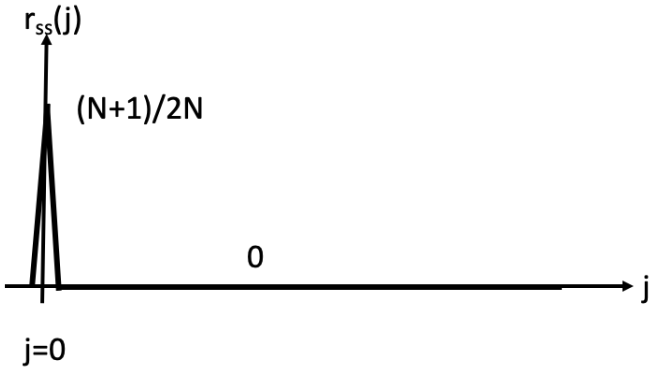


Fig. 3.4. Cross-correlation of s with s'. The residual value of the autocorrelation has disappeared), the value outside the peak is exactly 0.

The M-sequences have a different number of 0's and 1's —the difference is one unit for 1's. Because of this, if the signal present at the correlator input has an offset, whether

coming from a background radiation term not considered in Eq. (2.1) or from offsets in the photoreceiver electronics, the correlator output will produce the desired signal superimposed on a small constant value [12]. Although we will see later that this not a big issue in practice, to avoid this effect some authors [12], [33] have proposed variations around the M sequence with an even number of bits and the same number of 1's and 0's (or of 1's and -1's in their transposed versions). These new sequences, known as A1, A2 and AA1, are derived from the M-sequences.

The definition of these variants is based on the M-sequence and they are longer than the original sequence from which they are created. They have an even number of bits, with the exactly same number of 0's and 1's, or -1's and 1's in the transformed version of the original sequence **s**. The result of this parity is that offsets at the correlator input are automatically removed after applying the cross-correlation operation. An additional advantage of these sequences is the orthogonality between A1 and A2 sequences that can be used in applications of gases detection for example.

Definition of the A1 sequences [12]

Assume a sequence ***s'*** like that defined in Eq. **(3.2)**, then a new ***a'**** sequence of length 2N is created as

$$a_i'^{*} = (-1)^i \, s_i', \tag{3.5}$$

with $i = 0,1,\ldots,2N-1$, and *s'* assumed cyclic, i.e.,

$$\text{for } i>N-1,\ s_i' = s_{i-N}'. \tag{3.6}$$

The A1 sequence is created as

$$a_i^{A1} = a_i^{*} = \left(a_i'^{*} + 1\right)/2 \tag{3.7}$$

The cross-correlation function of **a*** and **a'*** is (see also fig. 3.5)

$$r_{a^*a'^*}(j)=\begin{cases} N & j=0 \pmod{2N} \\ 1 & j=2n-1 \pmod{N} \\ -1 & j=2n \pmod{N} \\ -N & j=N \pmod{2N} \end{cases} \qquad n=1,2,...,(N-1)/2 \tag{3.8}$$

where N is the length of the original M-sequence. Between the principal peaks appears a ripple of values (+1,-1) as shown in fig.3.5.

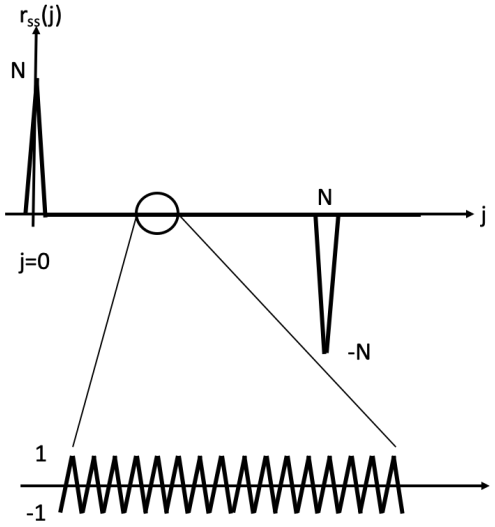


Fig. 3.5. Cross-correlation function of the A1 sequence with a zoom showing the ripple between principal peaks.

Definition of A2 sequence [12]

Taking again the s' sequence defined by Eq. **(3.2)** and assumed cyclic (Eq. (3.6)), a new **a'**** sequence of length 4N is defined by making

$$a_i'^{**}=\begin{cases} s_i' & (i=4m,\ 4m+1) \\ -s_i' & (i=4m+2,\ 4m+3) \end{cases} \qquad m=0,1...,N-1 \tag{3.9}$$

where N is the length of the original M-sequence. The A2 sequence is

$$a_i^{A2} = a_i^{**} = \left(a_i'^{**} + 1\right)/2 \tag{3.10}$$

The cross-correlation function of a** and a'** is (see also fig. 3.6)

$$r_{a^{**}a'^{**}}(j) = \begin{cases} N & j = 0 \pmod{4N} \\ 1 & j = 4n-2 \pmod{2N} \\ 0 & j = 2n-1 \pmod{2N} \\ -1 & j = 4n \pmod{2N} \\ -N & j = 2N \pmod{4N} \end{cases} \qquad n = 1,2,...,(N-1)/2 \tag{3.11}$$

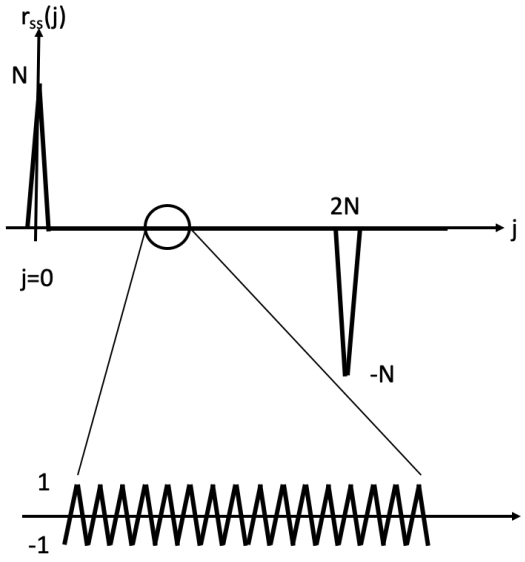


Fig. 3.6. Cross-correlation function of the A2 sequence with a zoom showing the ripple between principal peaks.

Definition of AA1 sequence [32][33][35]

The sequence AA1 is defined from A1 sequence as

$$a_i^{AA1} = a_i^{A1} + a_{i+1}^{A1}. \tag{3.12}$$

Its cross-correlation with the a'^{*} sequence is similar to that depicted in fig. 3.5, but with 0 values between the principal peaks.

$$r_{a^{AA1}a'*}(j)=\begin{cases} N+1 & j=-1 \\ N+1 & j=0 \\ 0 & j=1,2,...,N-2 \\ -(N+1) & j=N-1 \\ -(N+1) & j=N \\ 0 & j=N+1,...,2N-2 \end{cases} \qquad (all \bmod 2N) \tag{3.13}$$

The price to pay is that AA1 is a ternary sequence and therefore is more complicated to manage it in digital circuits. And additional price is that the correlation shows a peak of two samples instead of one in the previous sequences (fig. 3.7).

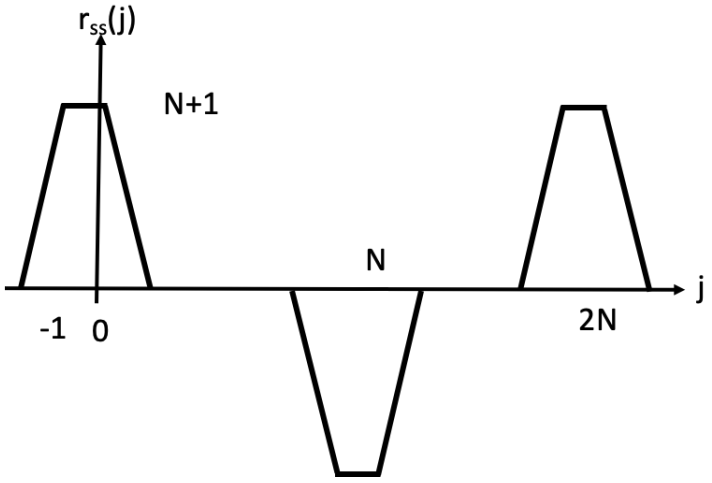


Fig. 3.7. Cross-correlation function of the AA1 sequence with a'*. The ripple disappears but the maximum of the correlation is two samples long.

This sequence is shown here as another possible option but it will not be used in this work, because it is focused in the use of digital circuits and binary sequences.

Comparison of the sequences

Using an M sequence as a source, the fig. 3.8 shows the first 10 bits of the M original sequence of 1023 bits length and the first 20 bits of the derived A1, A2, and AA1 sequences. Authors who propose these new sequences claim that these new modulations provide better Signal-to-Noise ratio performance [12][33][34][35]. Authors

concludes that as the number of 1's and 0's is equilibrated in them, the suppression on the background noise is better.

	Bit number																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
M	**0**	**1**	**1**	**0**	**0**	**1**	**0**	**0**	**0**	**1**										
A1	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**1**	**0**
A2	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
AA1	**0**	**1**	**2**	**1**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**0**	**1**	**2**	**1**	**0**	**0**	**1**	**1**	**1**

Fig. 3.8. The first 10 bits of the M original sequence and the first 20 bits of the derived A1, A2, and AA1 sequences

From the definition of sequence AA1 and the values in the fig. 3.8 it can be observed that the AA1 sequence is a ternary sequence (0, 1, 2) as it has been explained in section 0. Its use would increase the complexity of the electronics needed to modulate the laser. In practice it is easier to use binary values (0,1) in the sequences because all the electronic system can be implemented as digital chain working only with the two states. For this reason, the AA1 sequence is presented but it is not used in this work.

The theoretical behavior of the different sequences under different conditions can be evaluated through Matlab simulation.

Fig. 3.9 shows the cross-correlation of each sequence after being convolved with a delta impulse response of amplitude 1 and a 10-sample delay —the sequence H_{seq} represents this impulse response— with the corresponding modified sequence. Under noise- and background-free conditions the result of the correlation is correct in the sense that the delta target is retrieved.

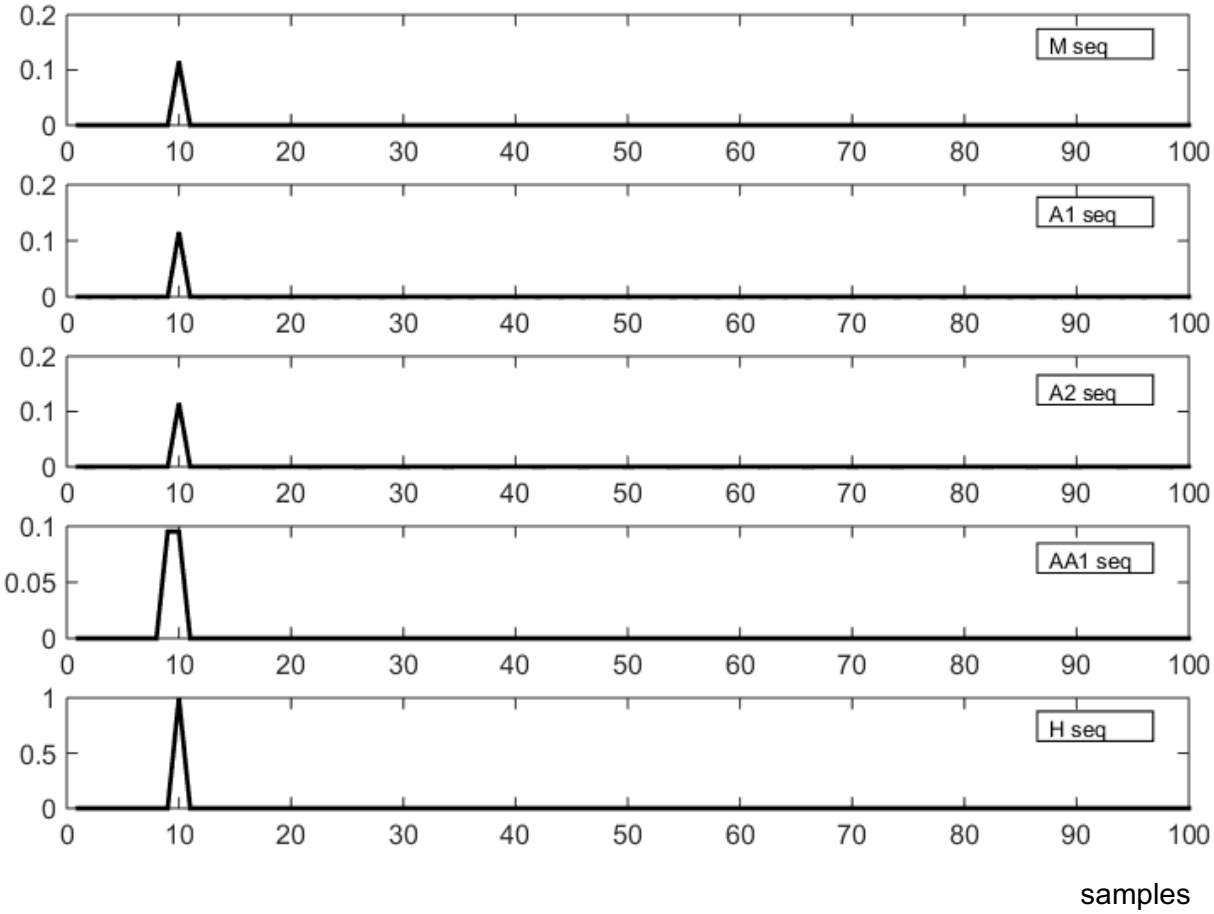


Fig. 3.9. Cross-correlation function between the return signal from an impulse target (H_{seq}) for the set of sequences studied in this work and the equivalent modified sequence. The base M-sequence is 1023 bits long and the derived A1,A2, AA1 are 2046,4092 and 2046 bits. The vertical axis units are arbitrary.

In figure 3.10 the range correction is applied —as usual in lidar systems— to the correlation output signal, but still in the situation of ideal conditions of noise- and background-free signals. Some sequences as A1 and A2 show their ripple behavior at far ranges (high number of sample).

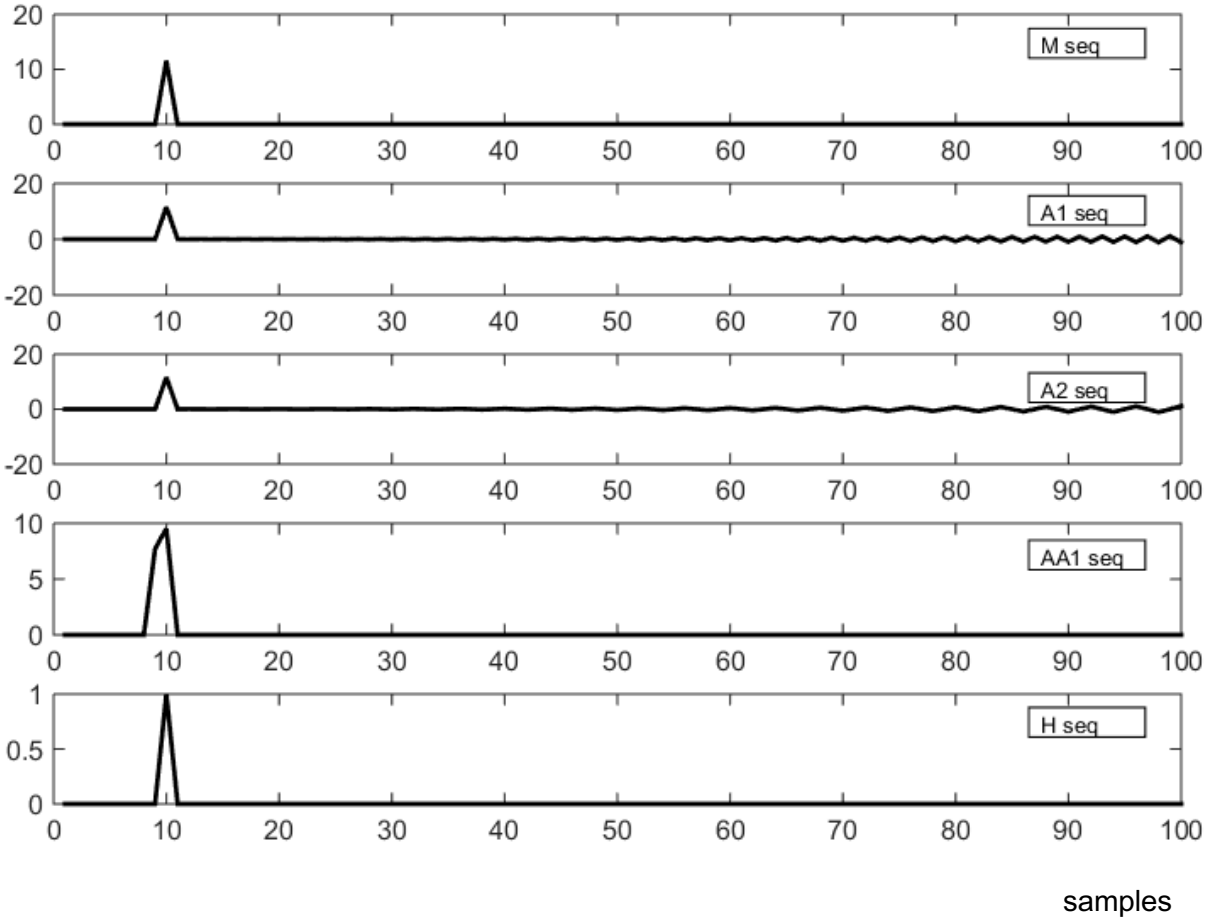


Fig. 3.10. Range-corrected cross-correlation function between the return signal from an impulse target (H_{seq}) for the set of sequences studied in this work and the equivalent modified sequence (decoding sequence). The vertical axis units are arbitrary adapted to the amplitude of each function. A1 and A2 sequences show a ripple that in the case of A1 is ±1 and in the case of A2 is ±2 (before applying the range correction).

An important issue to study is the behavior when these sequences are put at work in non-ideal conditions, i.e. in a system with noise, offsets, distortion and saturation, due to background radiation, detectors, amplifiers, A/D converters, etc. This is analyzed in the next section 0.

3.2. BEHAVIOR OF THE SEQUENCES IN NON-IDEAL CONDITIONS

The application of these sequences in a lidar system has to be considered in a real environment, where it is possible that they are affected by several adverse phenomena:

- Noise (in particular from the APD and/or the transimpedance amplifier).
- Background radiation from the sky, which affects in two ways: moving the offset of the signal (it can be corrected) and increasing the shot noise in the detector. This effect is especially important in daylight conditions.
- Offsets and noise in the amplifiers.
- Nonlinearities in the amplifiers and converters.
- Saturation in amplifiers and converters.

When sequences are working under these hampering effects, the final result after performing the correlation operation will depart from the ideal result, depending on the conditions. The behavior under noise or offsets should be good, because in a lidar system, and especially when using low power lasers, the signal-to-noise ratio is usually extremely.

The different cases are simulated through a simple equation:

$$y = a_1 s_r + a_2 s_r^{\,2}, \tag{3.14}$$

where s_r is the received sequence with the effects of noise and background/offset, a_1 is the coefficient of linear behavior of the receiver system, and $\mathbf{a_2}$ is the coefficient for quadratic behavior which tries to simulate a non-linear amplification of the system that can generate distortion. In turn

$$s_r = s + n + b, \tag{3.15}$$

with n representing the effect of random noise and b the background/offset. In the following studied situations, the coefficient a_1 is fixed to **1** and the other values a_2, n and b are changed depending of the situation to evaluate. Specially interesting is the case of nonlinearities.

Noise

In the equation (3.15) the noise sequence n is generated using a mean value of 0 and a standard deviation value $\sigma = 1$, and the s sequence has an amplitude of 1. The range-corrected result of the cross-correlation of the y sequence with the corresponding sequence is shown in figure 3.11. It is visible the peak of the correlation and the effect of noise in the correlation operation, amplified by the range correction. Other effects are the ripple of A1 and A2 sequences and the wider recovered impulsional response with AA1 sequence (2 samples instead 1 sample).

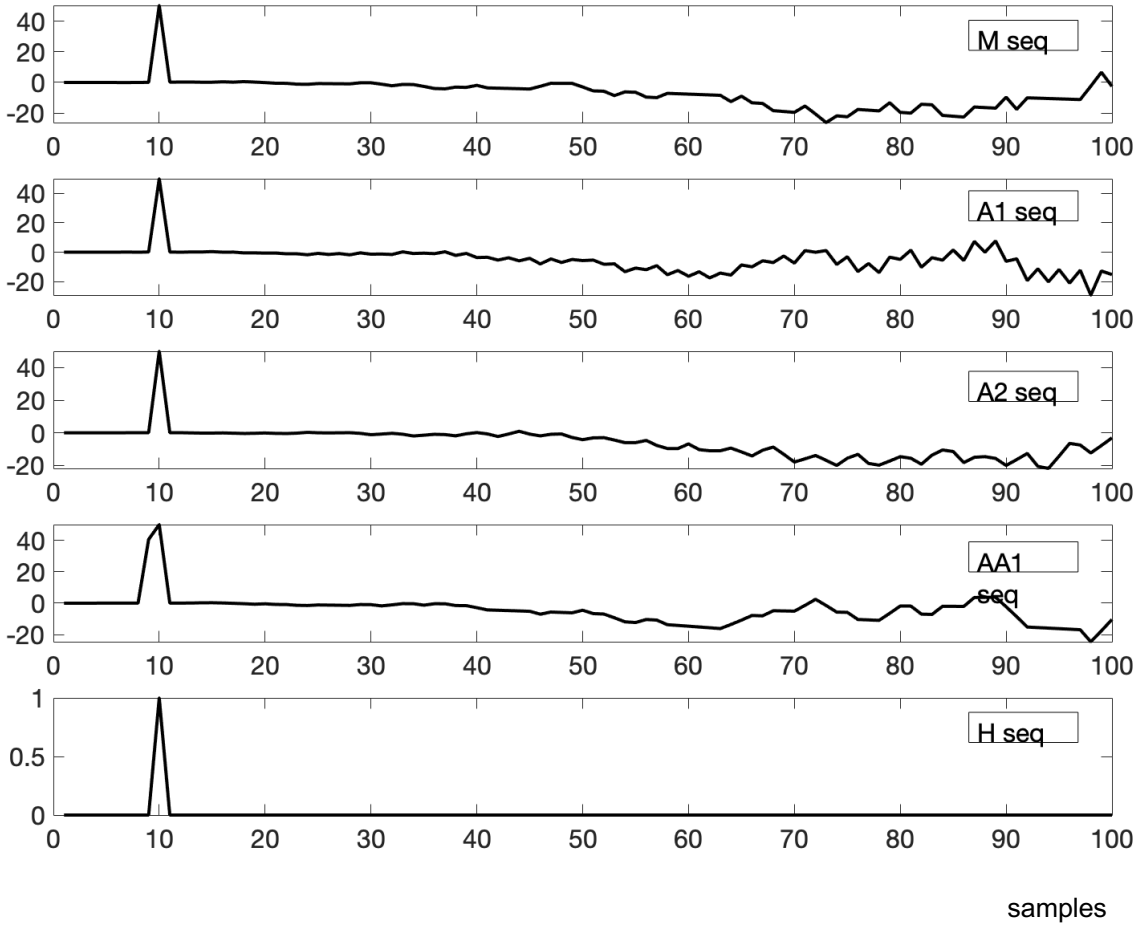


Fig. 3.11. Cross-correlation function in presence of noise with range correction. The effect of noise is visible. Sequence length 1023. Sequence amplitude=1. Noise standard deviation=1. The vertical axis units are arbitrary adapted to the amplitude of each function.

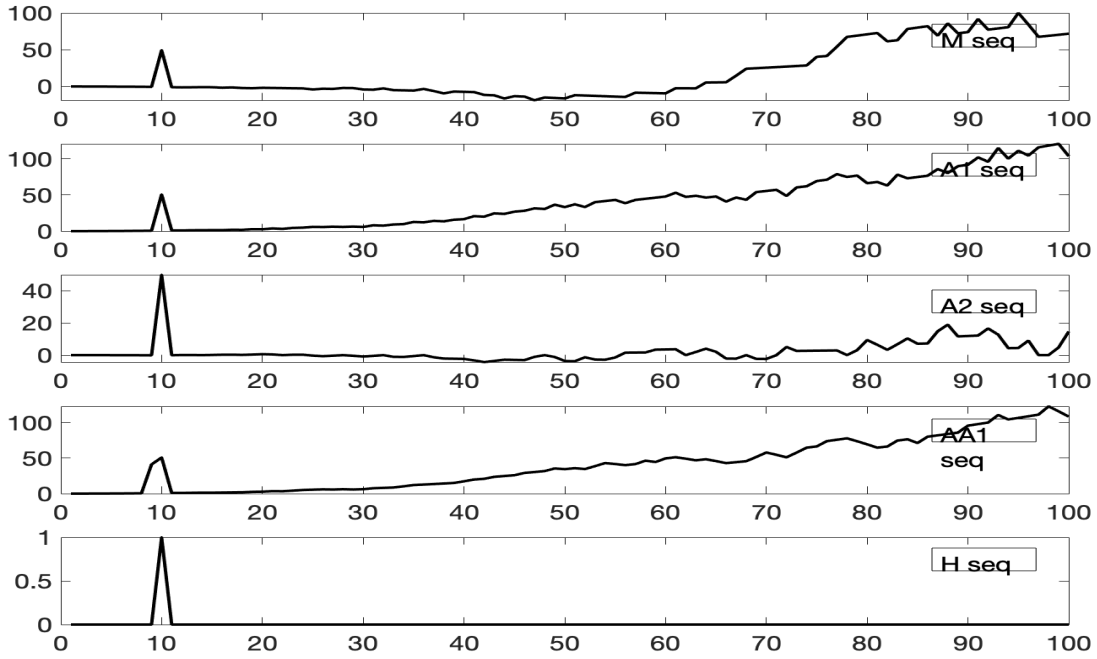
Offset from background radiation and electronics

Background radiation affects in two ways the photoreceiver: the generation of additional shot noise in the photodetector and the modification of the offset at the photoreceiver output. To reduce the background radiation effects an optical narrowband filter adapted to the laser wavelength is used.

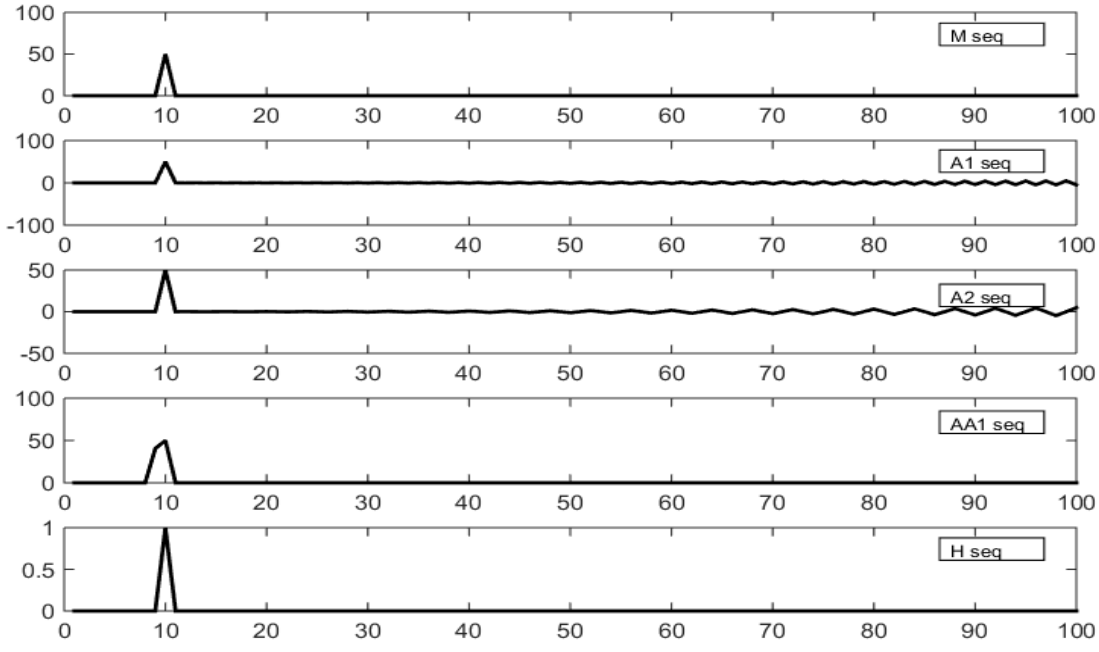
Background-radiation-induced shot noise is only an issue in daylight operation. During night time this problem is not relevant.

The effect of offset can be corrected within certain limits doing a correction in the amplifier circuit. If the photoreceiver module that contains the transimpedance amplifier and other amplification stages is saturated by the offset produced by the sky background radiation, correction is not possible. Saturation risk is especially important in the presence of sunlit clouds in the receiver system field of view. In this work a commercial photoreceiver module (Hamamatsu) is used and to have a good control of this offset a new design would be necessary to control all the amplifier stages.

The residual offset value depends on the pseudo-random sequence employed in the system. When it is a simple M sequence —with different numbers of 1 and -1 values in the sequence and in the decoding sequence— the offset at the correlator input results in an offset at the output however small. When it is an A1, A2 o AA1 sequence as the decoding sequence for them is balanced—with the same numbers of 1 and -1 values— an offset in the received signal at the input of the correlator is cancelled. But in fact, this issue van be solved in the correlation process. Calculating the mean value of the signal and subtracting before the correlation operation a good suppression of the offset can be obtained. Figure ***3.12***(a) shows the output correlation in presence of an offset value in the signal with the same amplitude than the sequence. Figure ***3.12***(b) shows the same output of correlation but subtracting the mean value of the signal before the operation.



(a)



(b)

Fig. 3.12. (a) Cross-correlation function with an offset present at the input of correlator. (b) The same but with subtracting the mean value of signal before correlation. Range corrected. Sequence length 1023. Sequence amplitude=1. Offset=1. The vertical axis units are arbitrary adapted to the amplitude of each function.

Non-linearities and saturation

Non-linearities

Using any of the previous set of sequences —M. A1, A2, AA1— with the condition that the channel response is just a time delay —a delayed delta function— the result obtained after the correlation is the expected result, that is, the impulse response of the channel —in this case, a delta function. In practice the physical response of the atmosphere, which is made up of aerosols, clouds, etc., cannot be a delta-function in most of the cases —a cloud in a clear atmosphere could be the most similar to a delta function.

When the system exhibits nonlinearities, the behavior of the system changes surprisingly. Some authors have theorized about the effects of non-linearities in the measure of the system impulse response [36]–[38].

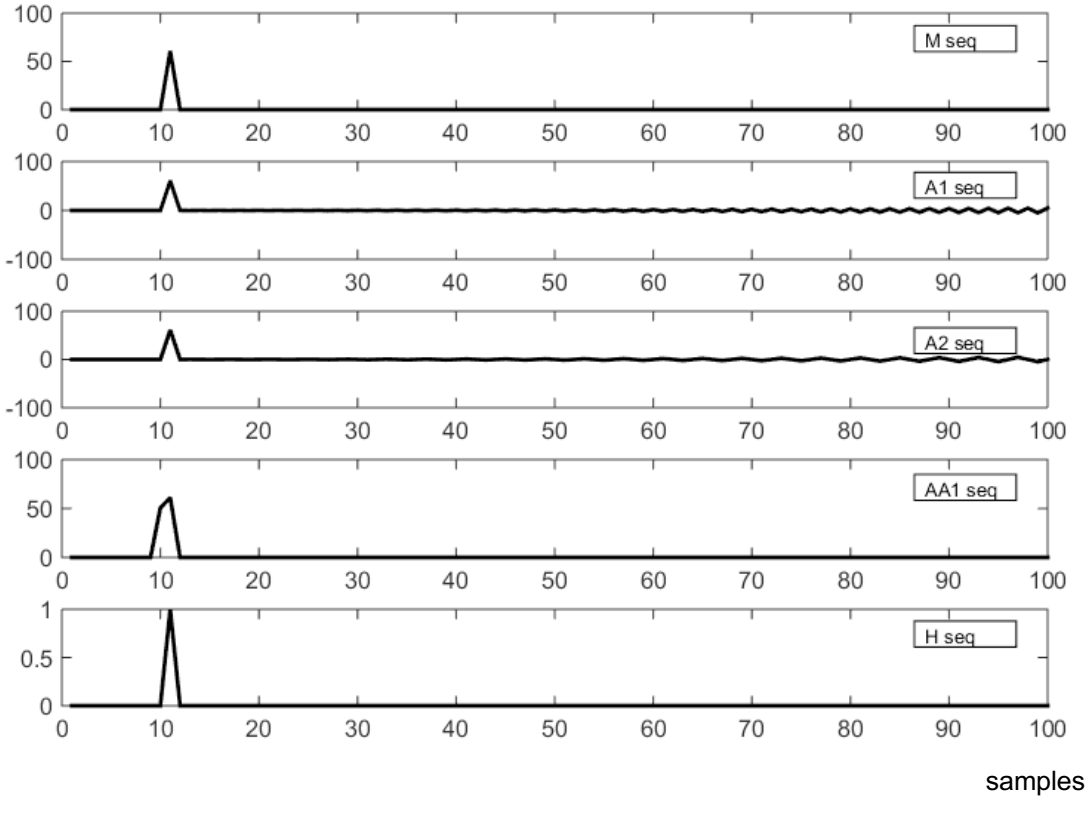


Fig. 3.13. Cross-correlation function with a small quadratic distortion (a2=0.005). Range corrected. Sequence length 1023. Sequence amplitude=1. Noise free. The vertical axis units are arbitrary.

Figures 3.13 and 3.14 are the output correlator with a distortion present in the received sequence when the response of the channel is a delta function (pure delay), using two

different values of the coefficient of the quadratic distortion in the equation (3.14) ($a_2 = 0.005$ and $a_2 = 0.1$ respectively). As it is possible observe in these figures the correlation output does not show the effect of the distortion (the ghost peak is not visible), only the result of a perfect correlation —the ripple in the sequences A1, A2, AA1, and the 2 samples of the correlation result in the sequence AA1.

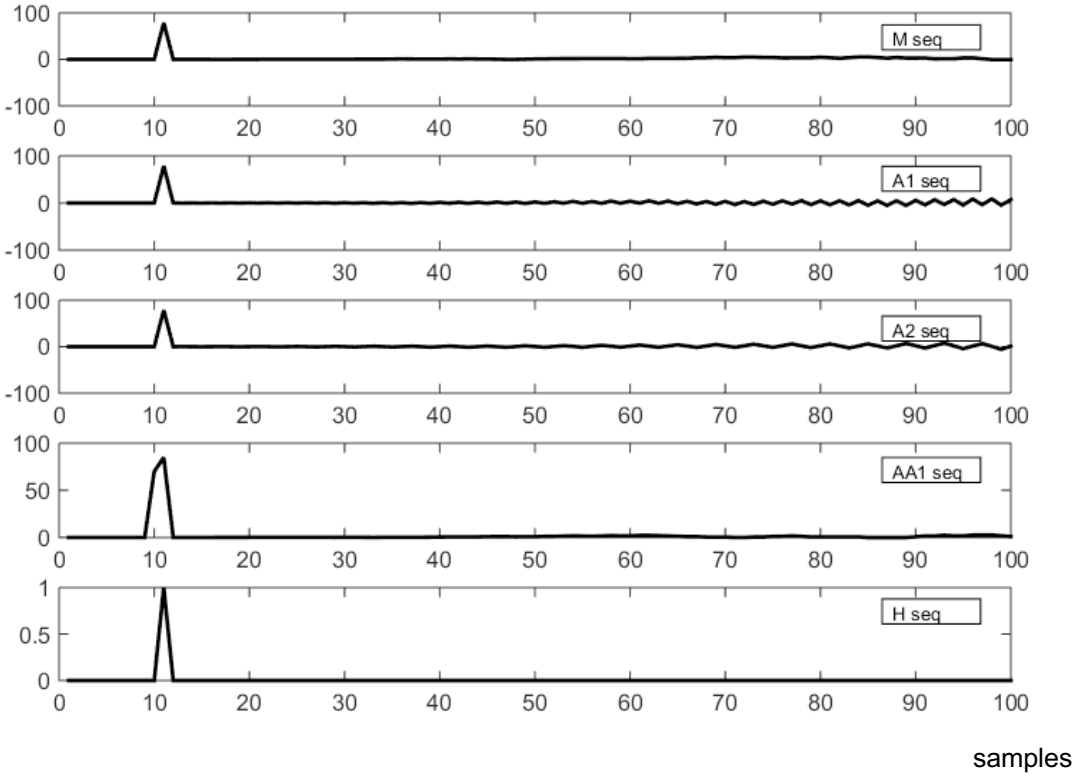


Fig. 3.14. Cross-correlation function with a small quadratic distortion (a2=0.1). Range corrected. Base M-sequence length 1023. Sequence amplitude=1. Noise free. The vertical axis units are arbitrary.

Figures 3.15 and 3.16 are obtained with the same non-linearity parameters than the fig. 3.13 and 3.14, but with a channel impulse response wider than precedent situation —2 samples wide instead of one, that is, two delta functions. A new effect is visible in figure 3.15. A secondary, ghost peak, appears in the cross-correlation function for the M-sequence, exactly at 76 samples in for quadratic coefficients $\mathbf{a_2}$ as small as 0.005. This effect depends on the sequence, because it is visible in the **M** sequence but not in the **A1**, **A2** or **AA1** sequences of the figure. In fact, it depends on the **M** sequence chosen, because each sequence of the same length has a different position of the ghost peak and sequences of different length also change the position of this peak. The same could

be said of sequences **A1**, **A2**, **AA1**, because they depend on the **M** sequence, but it would be necessary to test where the ghost peaks could appear because no rule exists.

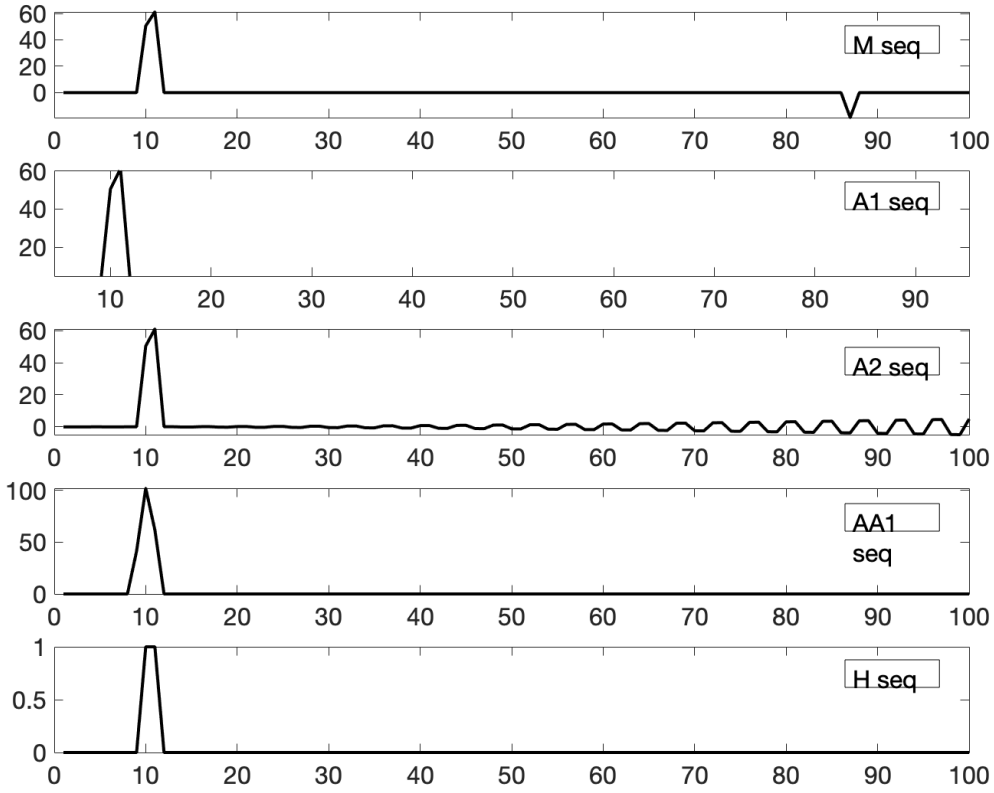


Fig. 3.15. Cross-correlation function with a small quadratic distortion (a2=0.005). Range corrected. Sequence length 1023. Sequence amplitude=1. No noise. The vertical axis units are arbitrary adapted to the amplitude of each function.

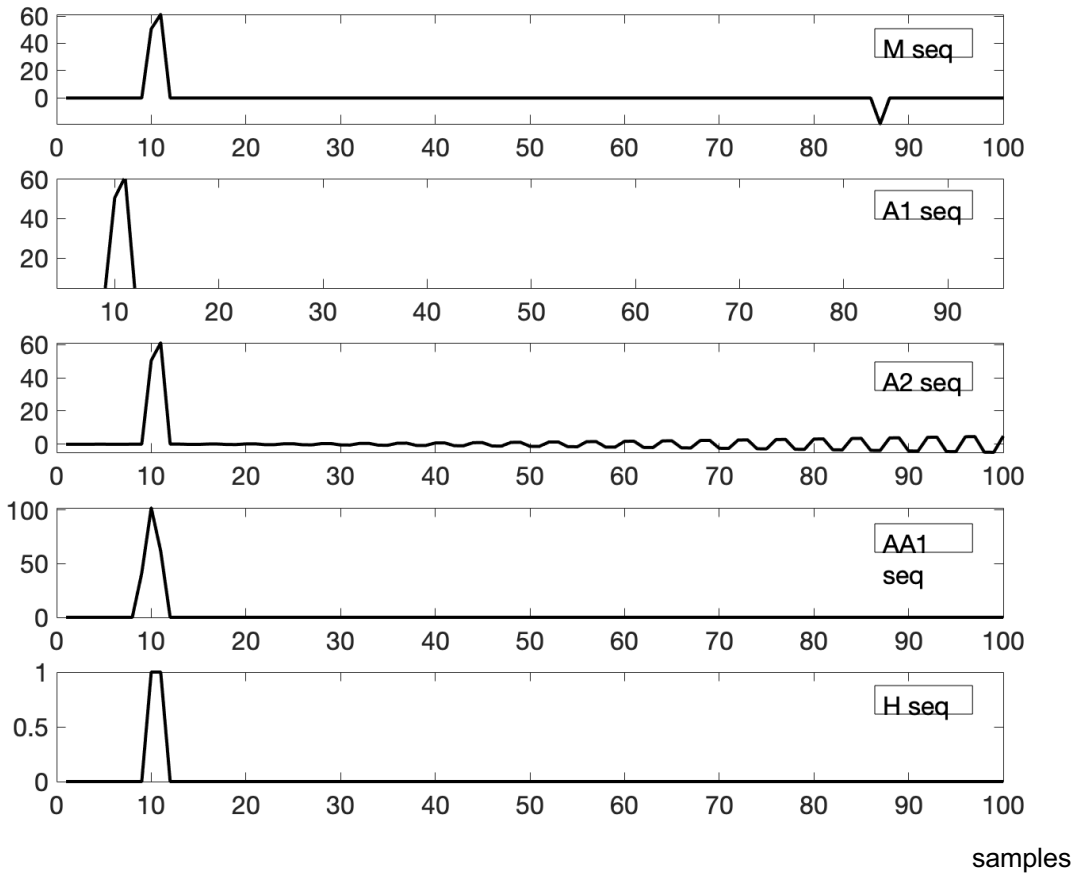


Fig. 3.16. Cross-correlation function with a small quadratic distortion (a2=0.01). Range corrected representation. Sequence length 1023. Sequence amplitude=1. No noise. . The vertical axis units are arbitrary adapted to the amplitude of each function.

This phenomenon is present on the M-sequence and in the other derivate sequences from this base sequence. The random sample where it appears depend on the sequence, the impulsional response, etc. but when these elements are fixed, the sample where it is present is the same, as it shows the figures respect to M-sequence. Testing different parameters of the distortions and an enough number of samples of the result, this ghost peak also appears in the other sequences. As example the figure 3.17 with the A1-sequence and a cubic distortion (a2=0.03). The false echo appears 950 samples approximately from the real one.

This false echo has been observed in the implemented lidar system. It is interesting to note that the secondary peak appears exactly at same sample number as in the simulation, i.e., at 76 samples after the echo if the same M sequence is used. Figure **3.18** shows the output of the real lidar, with this effect visible, using an M-sequence of 1023 length.

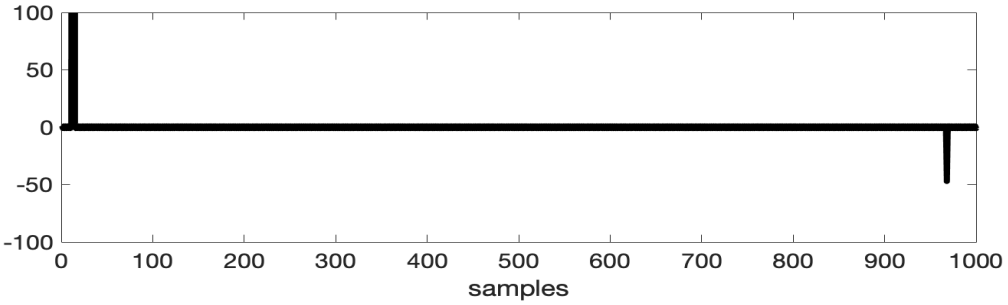


Fig. 3.17. Cross-correlation function of A1-sequence of 2046 samples long with a cubic distortion (a2=0.03) in the receiver chain transfer function. Range corrected representation. Sequence amplitude=1. No noise. The vertical axis units are arbitrary.

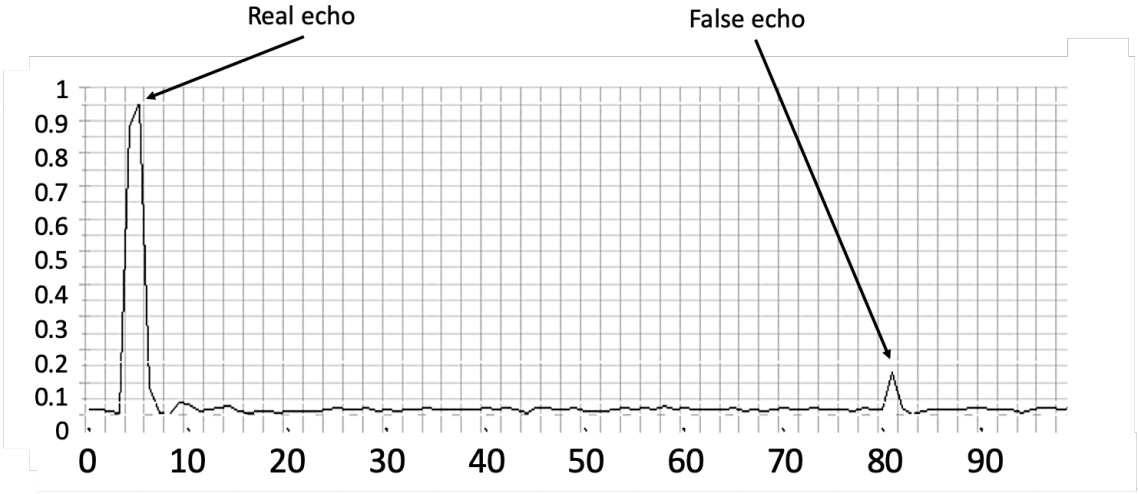


Fig. 3.18. Cross-correlation function at the output of the lidar pointed on a near solid target using an M-sequence. Not range corrected. The peak near the sample number 80 is the ghost or false echo, in fact is delayed 76 samples from the origin. The effect is due to the saturation in the amplifiers that in fact is a distortion. The vertical axis units are arbitrary.

The reason that explains this phenomenon is the behavior of the sequences when non-linearity is present in the system, further explained in the section 3.3 of this chapter.

Therefore, in real systems it is necessary to ensure the limitation of nonlinearities in the receiving chain to prevent the appearance of spurious echo even for very low distortion levels. To visualize this effect, a very simple non-linear function, a sigmoid, defined as

$$f(x) = \frac{1}{1+e^{-\gamma x}} - 0.5 \tag{3.16}$$

is used to model the relationship between the output and input of the photoreceiver electronics (fig. 3.19).

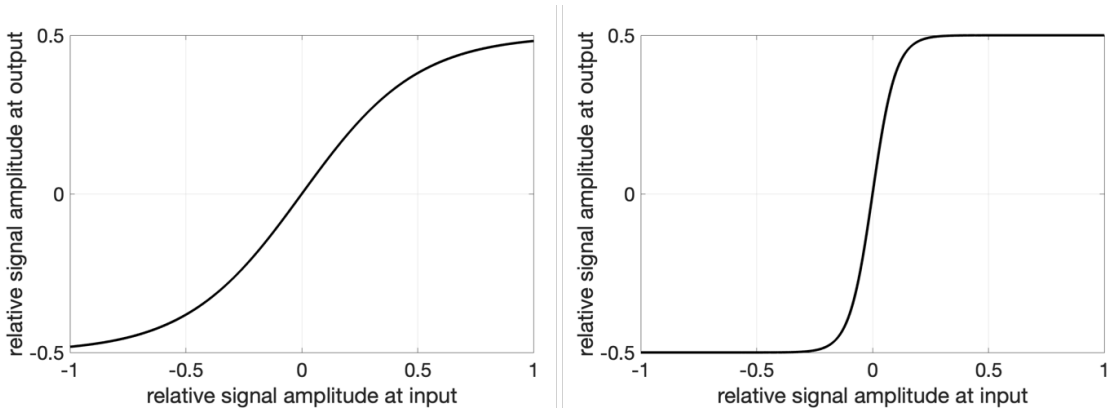


Fig. 3.19. Sigmoid functions with γ = 0.2 (left) and γ =1 in Eq. (3.16).

Figures 3.20 and 3.21 show the effect of the sigmoid-model distortion depending on the sharpness of the non-linearity. In the sequence M it appears the ghost peak after the correlation but in the A1, A2 and AA1 is not observed, but it exists. It is the same case as the figure 3.17, the peak occurs in a distant sample and it is not visible showing only 100 samples. Figure 3.21 corresponds to a sharper sigmoid (γ = 1) than fig. 3.20 (γ = 0.2).

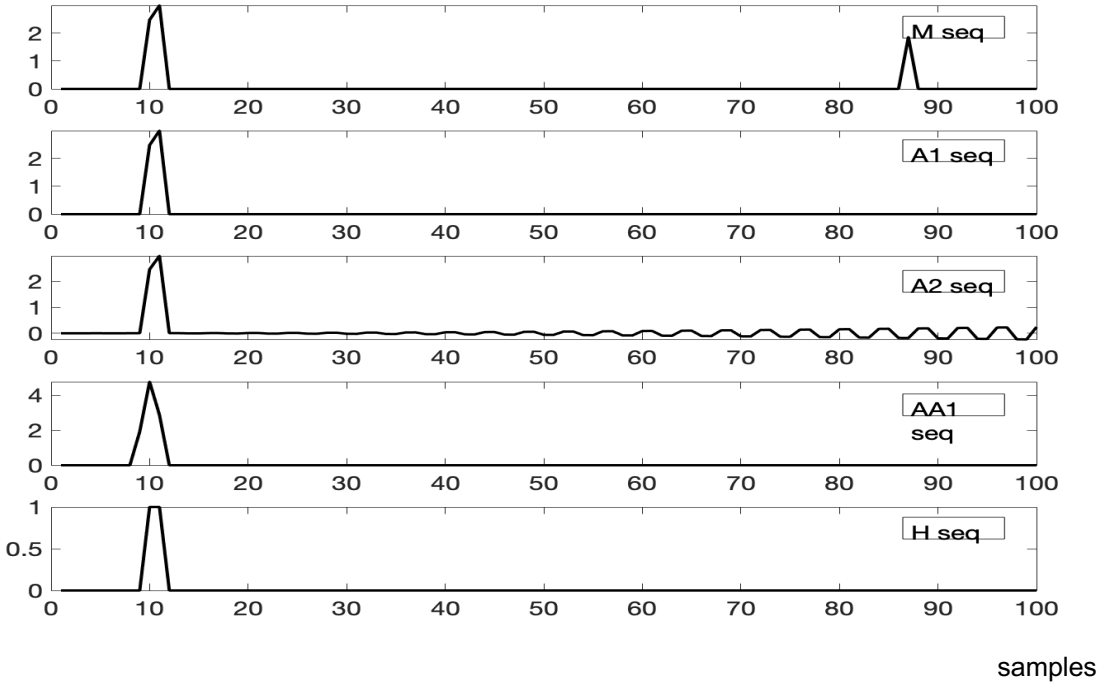


Fig. 3.20. Cross-correlation function with a sigmoid transfer function distortion. γ= 0.2. Range corrected. The vertical axis units are arbitrary adapted to the amplitude of each function.

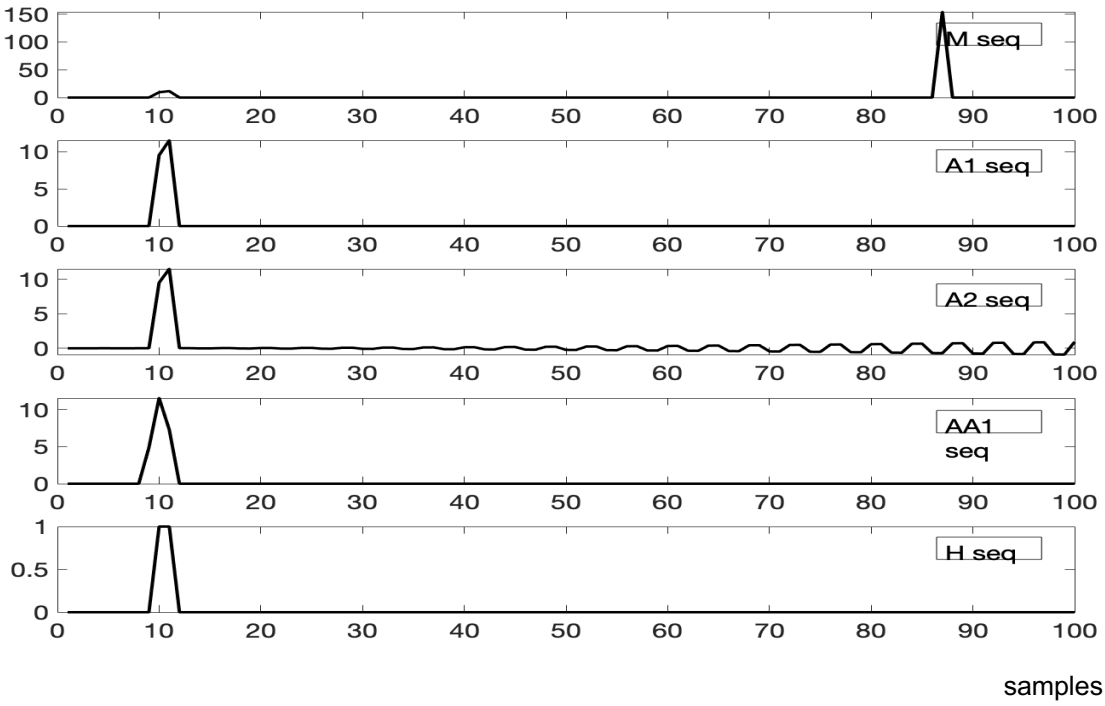


Fig. 3.21. Cross- function with a sigmoid transfer function distortion sharper than the 3.20. Gamma= 1. Range corrected . The vertical axis units are arbitrary adapted to the amplitude of each function.

The position at which the ghost peak appears is independent of the non-linearity sharpness used in the simulation, and it coincides with the behavior observed in the measurements done with the instrument developed. This means that the position of the ghost peak is due to the structure of the sequence and not to the kind of nonlinearity. Some bibliography can be found about this phenomenon [27], [36]–[38].

Saturation

The saturation of the amplifier in fact is a nonlinear effect and it also affects to the cross-correlation function result in the same way than the precedent case. Figure 3.22 shows an example of an amplifier with unit gain whose the output would be clamped outside the range (-1,+1) —relative amplitudes.

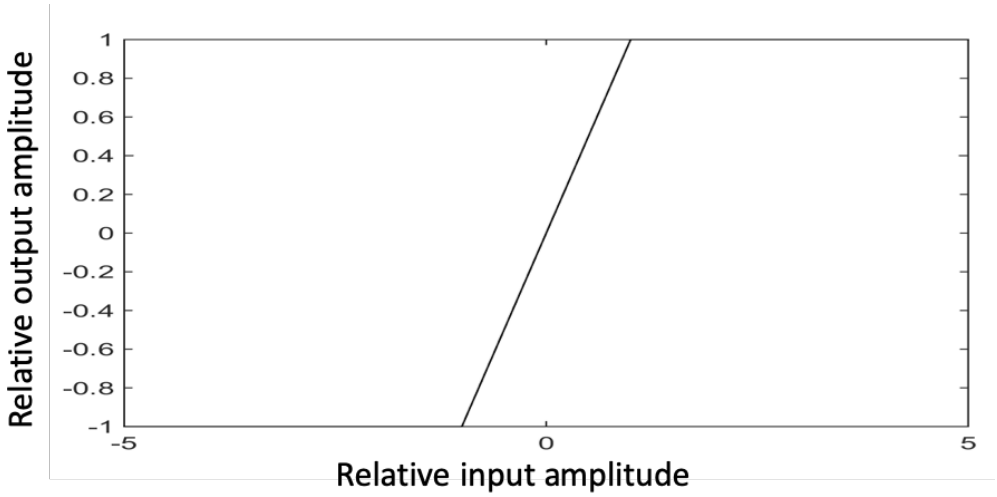
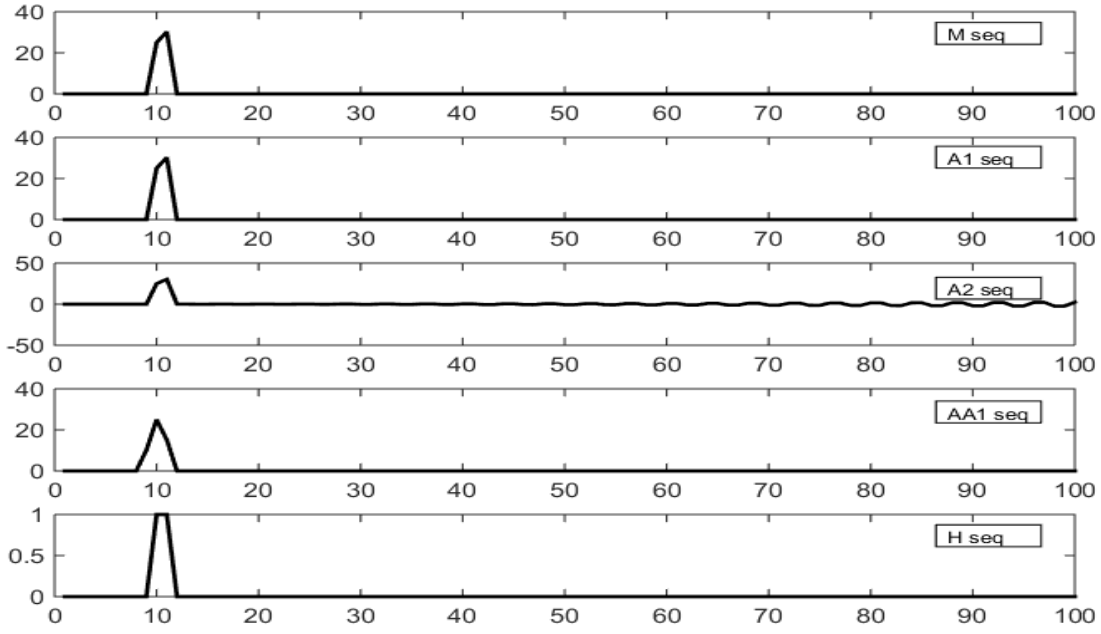
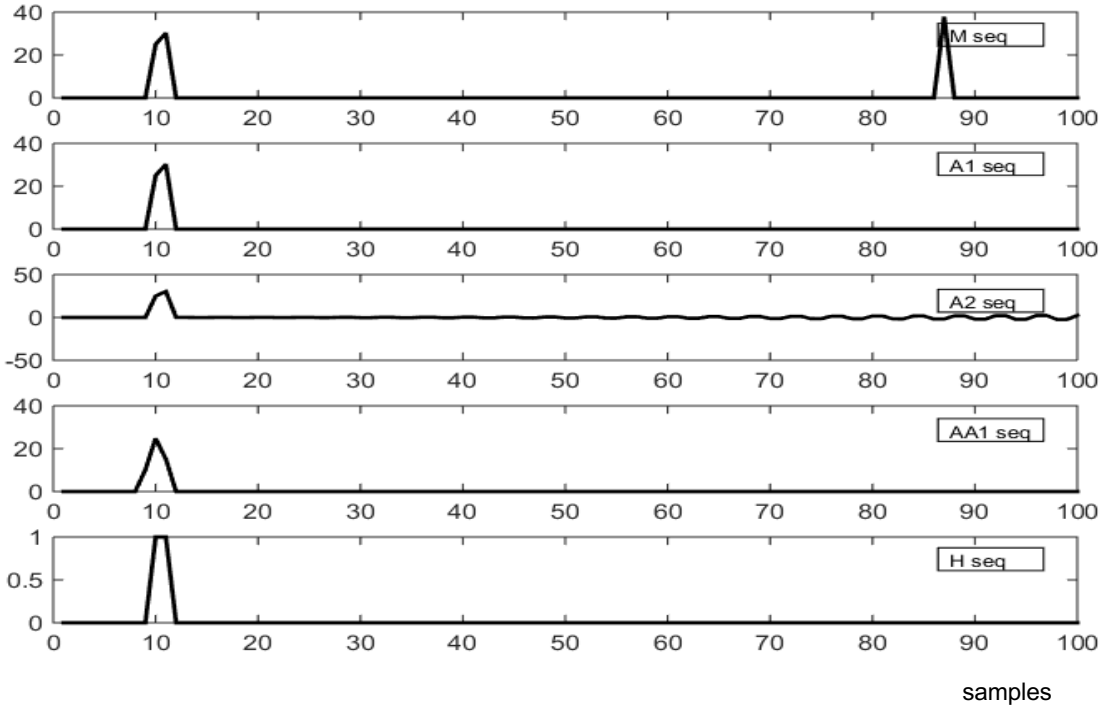


Fig. 3.22. Normalized function of a linear amplifier with saturation outside the central interval (-1,1) to use in the simulation.

The effect of this saturation is shown in figure 3.23, which shows the correlation when the amplitude of the sequence is in the limit to saturation (a=1 in fig,3.23a) and the correlation when the amplitude is only a bit higher (a=1.02 in fig. 3.23.b). For a>1 the effect of ghost secondary peak appears suddenly. In fact, this situation is similar to distortion in analogic world with the apparition of intermodulation products.



(a)



(b)

Fig. 3.23. Cross-correlation function with a saturation transfer function. Range corrected representation. (a) Shows the correlation when the amplitude of the sequence is a=1. (b) Shows the correlation when the amplitude is a=1.02. The effect appears suddenly.

Conclusion

Summarizing, we have found that nonlinearities in the receiver chain can produce false echoes in the lidar signal. Looking at the different kinds of non-linearities a conclusion can be obtained: The sample where the false echo appears does depends on the sequence, the impulsional response of the system —the atmosphere and it is variable— and the transfer function of the receiver chain. Fixed these parameters, each sequence has the false peak in a specific point (or specific points). As the impulsional response of the atmosphere is variable, the sample where it appears is not determined.

3.3. REASONS FOR THE BEHAVIOR OF THE PSEUDO-RANDOM SEQUENCE UNDER NON-LINEAR CONDITIONS

When the returned signal from the atmosphere is a single delta the system has a correct response to this delta independently of if the sequence received saturates or not the receiver system, because only one sequence is present in the receiver chain. But usually the echo returned from the atmosphere is longer than a delta function —a cloud or an aerosol plume are extended targets and the echo can be expanded in a train of deltas. This is the same as to have more than a sequence returning from the atmosphere with different delays and amplitudes in the receiver chain (Eq. (2.1)). It has been shown in the previous section that when these sequences pass through a system that is not exactly lineal some non-desired effects appear, like secondary or ghost echoes.

Some authors [14]–[16] have described this effect and given reasons for it, but in general the literature is short. Here we present another explanation for this behavior in the lidar system and quantify the expected amplitude of the ghost peak.

In fact, this behavior is due to one of the properties of M-sequences, namely the property 4 in the list of properties in the section 3.1 of this chapter. According to this property, the addition modulo 2 of one sequence and a delayed version of itself creates another delayed version of the sequence where the new delay is apparently random. The new delay is shorter than the length of the sequence. This new delay is not easy to calculate due to the random properties of the sequences and it can appear at a minor distance than the maximum range of interest or not.

Fig. 3.23 shows clearly this effect. The saturation can be in the electronic parts —amplifier, A/D converter, etc.— or in the software. When the signal has a value under the saturation limit of an amplifier, converter, etc., the system runs correctly but if the signal crosses only a little above the saturation threshold the effect appears immediately. In the case of the saturation due to the software, the same effect can be observed if there is a limitation of the maximum values of the variables used in the software for signal processing due to a bad selection. It is necessary to take care to avoid that the additions or multiplications contained in the software do not reach the limit values of the type of variables employed —integers, long integers, float, etc. This is possible when different types of variables are mixed.

Summarizing, the values of the signal in all the points of the receiver chain until the correlation process (amplifiers, A/D converters, variables used in the software, etc.) has to be below the saturation levels of these parts of the system.

Behavior of the system under non-linear transfer function

Let's assume that at the output of the photodetector there is a signal like

$$S = as_1 + bs_2 + c \tag{3.17}$$

where s_1 and s_2 are two shifted with respect to each other, but otherwise identical, M-sequences, a and b are their amplitudes and c is an offset value, and let's assume too that the whole receiver chain until the correlator can be modelled as a function $f(S)$ expressed as a Taylor expansion around a S_0 value

$$S' = f(S_0) + f'(S_0)(S - S_0) + \frac{1}{2} f''(S_o)(S - S_0)^2 + ... \ , \tag{3.18}$$

with non-linear behavior appearing through the terms order higher than 1, that is, S^2, S^3, etc. Taking the first quadratic term S^2 and expanding:

$$S^2 = (as_1 + bs_2 + c)^2 = a^2 s_1^2 + b^2 s_2^2 + 2abs_1 s_2 + 2acs_1 + 2bcs_2 + c^2 . \tag{3.19}$$

As the sequences only take '1' or '0' values, $s_1^2 = s_1$, $s_2^2 = s_2$, and the equation (3.19) can be written

$$S^2 = (as_1 + bs_2 + c)^2 = a(a + 2c)s_1 + b(b + 2c)s_2 + 2abs_1 s_2 + c^2 \tag{3.20}$$

The first and second terms are the sequences with the corresponding amplitudes and the fourth term is a residual offset value. The third term contains the product of the two sequences, that can be considered a logical product because the sequences are made of '1' and '0'. To analyze this term, consider the development of logical or-exclusive operation shown in the appendix A1. The result is that

$$s_1 s_2 = \frac{1}{2}(s_1 + s_2 - s_3) \ , \tag{3.21}$$

where $s_3 = s_1 \oplus s_2$ and $\oplus$ denotes the "or exclusive" operation.

Therefore the third member of the equation (3.20) is

$$2abs_1s_2 = ab\left(s_1 + s_2 - s_3\right) = abs_1 + abs_2 - abs_3 \tag{3.22}$$

Finally the equation (3.20) results in

$$\begin{aligned} S^2 &= \left(as_1 + bs_2 + c\right)^2 = \\ &= a\left(a + 2c\right)s_1 + b\left(b + 2c\right)s_2 + abs_1 + abs_2 - abs_3 + c^2 = \\ &= \left(a^2 + ab + 2ac\right)s_1 + \left(b^2 + ab + 2ac\right)s_2 - abs_3 + c^2, \end{aligned} \tag{3.23}$$

where the term s_3, corresponding to a sequence of the same class as s_1 and s_2, but with another shift, gives rise to a ghost peak in a pseudo-random . Hence, a quadratic behavior on the system is associated to negative peaks, as it was shown numerically in the top panel of fig. 3.15.

Similar developments can be carried out for non-linear terms of higher order.

If the system has a value of offset c=0 then the equation (3.23) reduces to

$$S^2 = \left(a^2 + ab\right)s_1 + \left(b^2 + ab\right)s_2 - abs_3 \ . \tag{3.24}$$

Assume that the system has a transfer function like

$$y = \alpha_0 x + \alpha_1 x^2 \, . \tag{3.25}$$

If $x = as_1 + bs_2$, and using the above reasoning, the output will be

$$y = \left(\alpha_0 a + \alpha_1 a^2 + \alpha_1 ab\right)s_1 + \left(\alpha_0 b + \alpha_1 b^2 + \alpha_1 ab\right)s_2 - \alpha_1 abs_3 \tag{3.26}$$

If $a = b = \alpha_0 = 1$ then the equation (3.26) reduces to

$$y = \left(1 + 2\alpha_1\right)s_1 + \left(1 + 2\alpha_1\right)s_2 - \alpha_1 s_3 \, . \tag{3.27}$$

For example, with $\alpha_1 = 0.1$, the system simulation shows that the output peak corresponding to s_3 has the predicted value of -0.1 (figure 3.24).

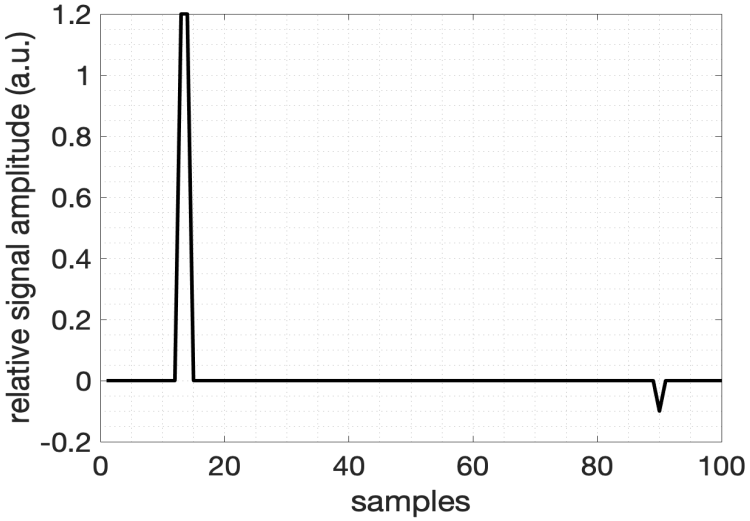


Fig. 3.24. Simulation of a non-lineal behavior when a=1, b=1, c=0 and α=0.1.

Of course, when there are more than two sequences the cross-products increase and it is possible observe "ghost" peaks in any position.

Behavior of the system under saturation

The system under saturation is a particular case of non-linear function, but it can be analyzed separately. Suppose again that the output of the photodetector is a signal like

$$S = as_1 + bs_2 + c \tag{3.28}$$

The effect of the saturation can be modelled as a block with a threshold level d_{th} such that if $S < d_{th}$ the output will be S but if $S > d_{th}$ the output will be d_{th}. This can be modelled by the function

$$f_s(S) = \begin{cases} S & if\ S < d_{th} \\ d_{th} & if\ S > d_{th} \end{cases} \tag{3.29}$$

assuming that $c < d_{th}$. Consider different cases:

a) *The threshold is only crossed when* $s_1 = 1$ *independently of the* s_2 *value.*

In this case

$$f_s(S)=\begin{cases} bs_2+c \quad \text{if } s_1=0 \\ d_{th} \quad if\ s_1=1 \end{cases} \tag{3.30}$$

Algebraically this is

$$f\left(as_1+bs_2+c\right)=s_1 d_{th}+\left(1-s_1\right)\left(as_1+bs_2+c\right)\cdot \tag{3.31}$$

Developing the right member and taking account that, because s_1 and s_2 only take 0 and 1 values, hence $s_1^2=s_1$, $s_2^2=s_2$, we have

$$f\left(as_1+bs_2+c\right)=c+\left(d_{th}-c\right)s_1+bs_2-bs_1s_2 \tag{3.32}$$

Using the equation **(3.21)** in this previous equation the result is

$$f\left(as_1+bs_2+c\right)=c+\left(d_{th}-c-\frac{1}{2}b\right)s_1+\frac{1}{2}bs_2+\frac{1}{2}bs_3 \tag{3.33}$$

Therefore, the sequence s_3 appears again. With $b>0$ the correlator output contains positive ghost peak as the simulation confirms.

b) *The threshold is only crossed when* $s_2=1$ *independently of the* s_1 *value.*

This case is totally symmetric to the case (a), which allows writing the result directly

$$f\left(as_1+bs_2+c\right)=c+\left(d_{th}-c-\frac{1}{2}a\right)s_2+\frac{1}{2}as_1+\frac{1}{2}as_3 \tag{3.34}$$

c) *The threshold is only crossed when* $s_1=1$ and $s_2=1$ *simultaneously*

$$f_s(S)=\begin{cases} as_1+bs_2+c \quad \text{if } s_1=0\ or\ s_2=0 \\ d_{th} \quad if\ s_1=1\ and\ s_2=1 \end{cases} \tag{3.35}$$

As an algebraic expression this equation can be written as

$$f\left(as_1+bs_2+c\right)=s_1s_2d_{th}+\left(1-s_1s_2\right)\left(as_1+bs_2+c\right) \tag{3.36}$$

DEVELOPING THE EXPRESSION, YIELDS

$$f\left(as_1+bs_2+c\right)=\left(d_{th}-a-b-c\right)s_1s_2+as_1+bs_2+c \tag{3.37}$$

And using the equation (3.21) we find

$$f\left(as_1+bs_2+c\right)=\frac{1}{2}\left(d_{th}+a-b-c\right)s_1+\frac{1}{2}\left(d_{th}-a+b-c\right)s_2-\\-\frac{1}{2}\left(d_{th}-a-b-c\right)s_3+c \tag{3.38}$$

As *a*, and *b* are positive values, if d_{th} is also positive for it to be crossed when $s_1=s_2=1$, it is necessary that $a+b+c>d_{th}$ and therefore s_3 the is multiplied by a positive coefficient. It leads consequently to a positive ghost peak as it can be seen in the simulation (figure 3.25). In this example, with $a=0.6,\ b=0.6,\ c=0,\ d_{th}=1$, the result for the peaks are 0.5, 0.5 and 0.1 as the equation predicts.

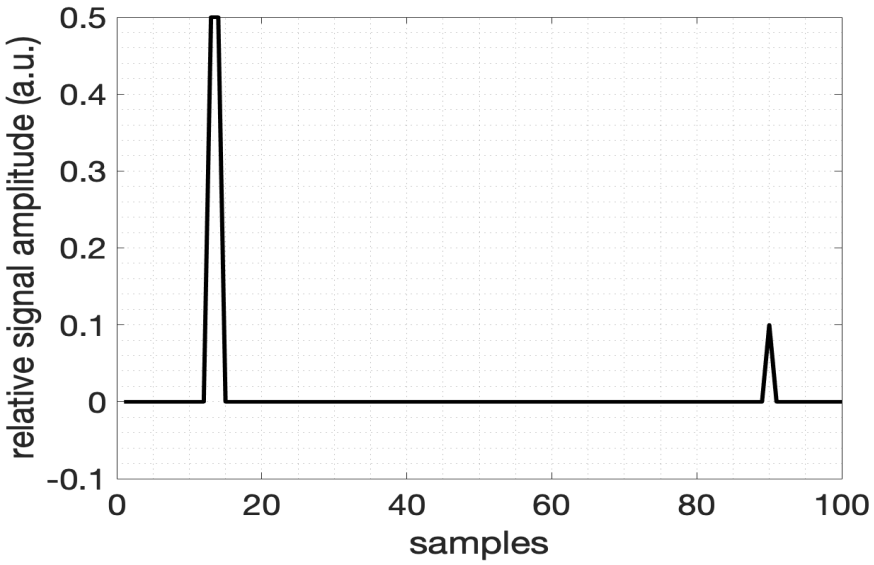


Fig. 3.25. Simulation of a saturation behavior when a=0.6, b=0.6 and c=0.

As a conclusion, it is of the utmost importance to ensure the linearity of the receiving chain and to avoid overflows in the correlation operations, as non-linear effects not only affect locally the retrieved atmospheric profile, but can propagate to the whole retrieved range.

4. MODELLING AND CHARACTERIZATION OF THE SYSTEM

4.1. MODELLING OF THE SYSTEM

To evaluate the behavior of the proposed CW lidar, it is possible to use a classical model based on a communications channel. In fact, a lidar can be thought of as a special kind of communications system, with the particularity that the transmitted and received signals are known, being the channel the unknown variable. The lidar purpose is precisely to evaluate the channel. In this context the proposed model is the one shown in figure 4.1, where, anticipating the digital signal processing to be carried out on the signal and as it has been shown in chapters 2 and 3, the time functions are expressed as sample sequences. As it has been discussed in the chapter 2 (section 0) in this system the receiver noise dominates over the signal shot noise and therefore the noise considered is the receiver noise and the shot noise due to the background radiation.

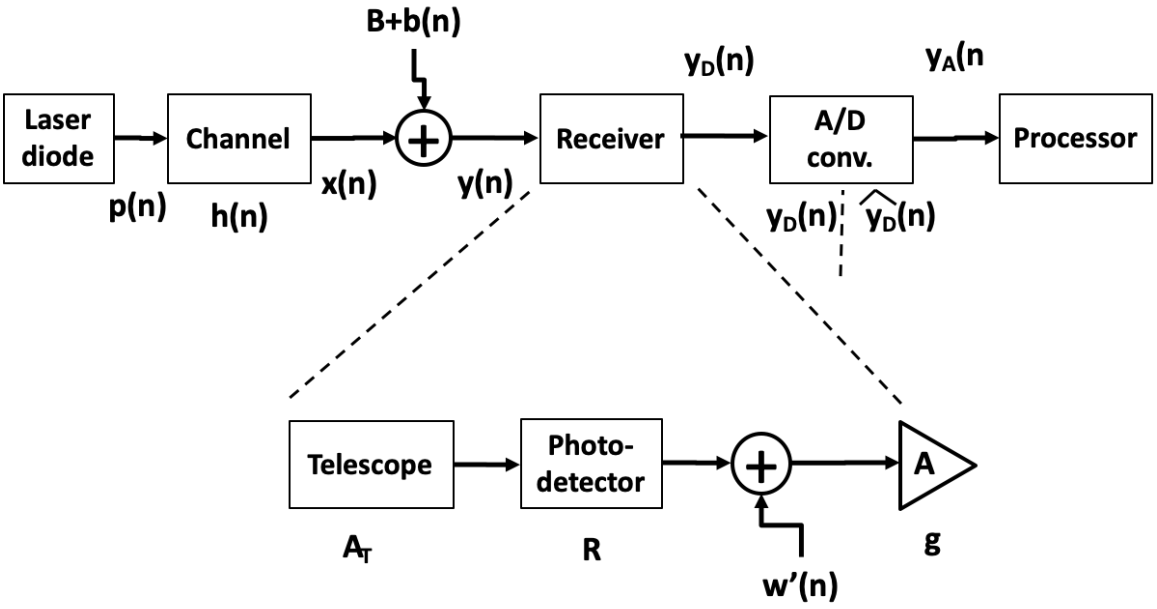


Fig 4.1. Equivalent block chain of the lidar system seen as a communications system. Symbols are explained in the text and their units are indicated between parentheses.

The meaning of the symbols in the figure is

- p(n): Laser output (W) with $p(n) = P_0 s(n)$ and *s(n)* is the sequence modulating the transmitted signal
- h(n): channel impulse response
- x(n): signal at the channel output, i.e. at the input of the receiver
- B: level of background radiation (offset in the photoreceiver)
- b(n): background-radiation induced equivalent noise (Wm^{-2})
- y(n): signal and noise at the input of the telescope (Wm^{-2})
- A_T: Telescope area (m^2)
- R: Photoreceiver responsivity (VW^{-1})
- w'(n): Receiver noise
- g: gain of the amplifier chain matching the output amplitude range of the photodetector to the input range of the A/D converter (matching gain)
- $y_D(n)$: signal at the input of A/D converter
- $\hat{y}_D(n)$: estimated signal $y_D(n)$ at the output of A/D converter
- $y_A(n)$: signal at the output of the A/D converter, $y_A(n) = \hat{y}_D(n) + n_A(n)$, with $n_A(n)$ the A/D converter noise).

In this model, the signal (Wm^{-2}) reaching the receiver front-end (the input of the telescope) is

$$y(n) = p(n) * h(n) + b(n) \tag{4.1}$$

where **h(n)** (eq. (4.2)) is the impulse response of the channel (the atmosphere in this case), **p(n)** is the power transmitted, and **b(n)** is the background radiation induced equivalent noise. The background radiation considered (irradiance Wm^{-2}) is the amount captured by the field of view of the telescope, calculate, considering the field of view of the telescope small, as the radiance of the sky background ($Wm^{-2}sr^{-2}$) multiplied by the field of view of the telescope in sr. The continuous time version of the impulse response is shown in eq. (2.5) and discretizing in time the result is

$$h(n) = \frac{\beta\left(\frac{c\tau n}{2}\right) \cdot e^{-2\sum_{k=0}^{n}\alpha(k)}}{\left(\frac{c\tau n}{2}\right)^2} \tag{4.2}$$

Calling τ the bit duration. The output of the photoreceiver is

$$y_D(n) = \Big(A_T R\big(p(n) * h(n) + b(n) \big) + w'(n) \Big) g \tag{4.3}$$

Where the parameters are defined in figure 4.1. If all the noise is included in one term **w**, then

$$y_D(n) = g\, A_T R\big(p(n) * h(n) \big) + g w(n) \qquad \text{where} \quad w(n) = A_T R\, b(n) + w(n)' \tag{4.4}$$

4.2. CHOICE OF THE RECEIVER GAIN

If n_D is the noise at the input (with σ_D^2) and n_A (with σ_A^2) is the noise introduced by the converter at maximum amplitude A, derived from the A/D converter signal-to-noise ratio specification —usually it is specified at maximum input amplitude ***A*** ($SNR_{AD} = A^2 / \sigma_A^2$) but if not then it will be normalized— the total noise n_T (fig.4.2) is

$$n_T = n_D + n_A \quad \rightarrow \quad \sigma_T^2 = \sigma_D^2 + \sigma_A^2 \tag{4.5}$$

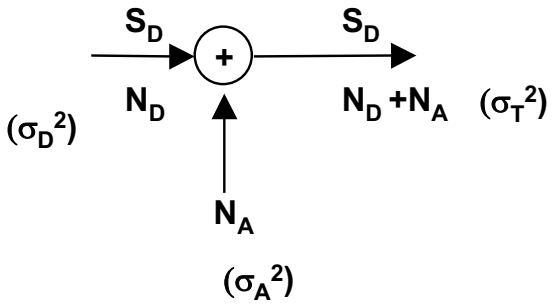


Fig 4.2. Simplified model of the noise added by the A/D converter. The variance of the different noise components is indicated in parenthesis.

If the signal at the input of the A/D converter is sampled over a sufficient number of bits, and we consider that the A/D converter gain is 1 (the possible gain of the A/D converter is absorbed in the matching gain *g*), then $\hat{y}_D \approx y_D$ .When a very low signal is received

from the atmosphere, the input y_D of the A/D converter is largely dominated by noise. In this case, for having a good digital processing and averaging of the noise, it is very important to obtain a good digitized signal, that means that it is necessary to spread the noise over all the A/D converter range. In this situation at the noise introduced by the A/D converter is negligible and therefore the total noise at the output is the approximately the same as the input.

$$\sigma_T^2 \approx \sigma_D^2 \tag{4.6}$$

In other words, if the range of input signal to the A/D converter matches the complete range of input amplitude, the output signal-to-noise ratio is approximately the same as the input one. Also, if the noise fluctuation uses the full range of the ADC then the digitization error is negligibly respect to the noise. [11]

If we define

$$a = \frac{A}{\sigma_D} = \frac{A}{\sqrt{N_D}} \tag{4.7}$$

A low value of ***a*** will means that the rms value of the voltage at the A/D converter input is so high that there is a non-negligible probability of saturating the converter. The minimum value of ***a*** can be set by limiting the saturation probability under the (reasonable) assumption that the A/D converter input behaves essentially as a zero-mean Gaussian noise. This probability is given by

$$p = erfc\left(\frac{a}{2\sqrt{2}}\right) \tag{4.8}$$

where

$$erfc(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt \tag{4.9}$$

is the complementary error function. The photoreceiver gain should be set to provide the traded-off ***a*** value. By choosing the saturation probability as 0.001 a value ***a*** = 6 is found. Table 4.1 gives the probability of converter saturation for several values of ***a***. It is seen that a good value of *a* is 6 and a value of 8 is excellent.

a	1-p
2	0.85
4	0.98
6	0.999
8	0.9997

Table 4.1. Probability of not saturating the A/D converter for several values of the parameter a, under the assumption that the converter input behaves essentially as a Gaussian noise.

4.3. PROCESSING GAIN

Assuming that *a* has been chosen according to the considerations of the precedent section 4.2 to have $\hat{y}_D \approx y$ with a good degree of approximation, the processing leading to the recovery of *h(t)* is accomplished by convolving $\hat{y}_D$ with a suitable signal related to the modulating signal *s(n)*.

Remembering the chapter 3 about the use of pseudo-random M-sequences to modulate the laser that results in a delta-like function (Eq. (3.4)), the processor in the figure 4.1 will perform the correlation of the A/D converter output with the reference sequence.

Assuming, taking into account equations (3.4) and (4.4) and considering

$$z_p(k) = \frac{1}{N}\sum_{n=0}^{N-1} y_D(n)\cdot s'(n+k) \tag{4.10}$$

one finds

$$z_p(k) \approx \frac{N+1}{2N} g\, A_T R P_0 \frac{c\tau}{2} h(k) + \frac{1}{N} g \sum_{n=0}^{N-1} w(n) s'(n+k) \tag{4.11}$$

Clearly the right side of equation (4.11) shows that the output of the processor for a shift ***k*** contains an estimate of the atmospheric “impulse response”, in the term,

$$S_p(k) = \frac{N+1}{2N} g\, A_T R P_0 \frac{c\tau}{2} h(k) \tag{4.12}$$

affected by noise contained in the term

$$N_p(k) = \frac{1}{N} g \sum_{n=0}^{N-1} w(n) s'(n+k) \tag{4.13}$$

Because the values of s'(n) are either 1 or -1, assuming the samples of **w(n)** independent random variables with variance σ_D^2, the variance of $N_p(k)$ in Eq. (4.13) is $\frac{g^2\sigma_D^2}{N}$, the standard deviation $\frac{g\,\sigma_D}{\sqrt{N}}$ and the signal-to-noise ratio in the estimation of **h(k)** is

$$SNR_p(k) = \frac{(N+1)\sqrt{N}\,A_T R P_0}{2N\sigma_D} \frac{c\tau}{2} h(k) \simeq \frac{\sqrt{N}\,A_T R P_0}{2\sigma_D} \frac{c\tau}{2} h(k) = \frac{\sqrt{N}\,A_T P_0 c\tau}{4NEP\,B^{1/2}} h(k) \tag{4.14}$$

with $\sigma_D / R = NEP\,B^{1/2}$ and considering N>>1.

If a number M of the cyclic sequences is processed and the results accumulated, the resulting SNR will be the same as if a sequence of length MN was used, namely

$$SNR_p(k) = \frac{\sqrt{MN}\,A_T P_0}{2NEP\,B^{1/2}} \frac{c\tau}{2} h(k) \tag{4.15}$$

Using the definition of *h(z)* in the eq. (2.5) in discrete version, the SNR at the output of the chain (the output of the processing system) is:

$$SNR_p(k) = \frac{\sqrt{MN}\,A_T P_0}{2NEP\,B^{1/2}} \frac{c\tau}{2z^2(k)} \beta(k)\, e^{-2\sum_{j=1}^{k} \alpha(j) z(k)} \tag{4.16}$$

In the real implementation it is necessary to consider the losses of the optical elements. Calling **T** the transmittance combined of all transmittances and reflectivities of the different optical elements the SNR can be calculated as

$$SNR_p(k) = \frac{\sqrt{MN}\,A_T P_0 T}{2NEP\,B^{1/2}} \frac{c\tau}{2z^2(k)} \beta(k)\, e^{-2\sum_{j=1}^{k} \alpha(j) z(k)} \tag{4.17}$$

4.4. TRADE-OFFS OF POWER AND RANGE RESOLUTION

From Eq. (4.14) is apparent that B should be as low as possible to increase the signal-to-noise ratio. But, to preserve the pulse shape, hence the power efficiency and the range resolution, B must satisfy

$$B \geq \frac{1}{2\tau}. \tag{4.18}$$

Case a): matched receiver

Let's assume that we work in a "matched" condition, whereby B takes the limit value $B = \frac{1}{2\tau}$. In that case, Eq. (4.14) can be written as

$$SNR_p(k) = \frac{\sqrt{N} A_T P_0 c \tau^{3/2}}{2\sqrt{2} NEP} h(k). \tag{4.19}$$

Calling $T = N\tau$ the sequence duration, Eq. (4.19) can be recast as

$$SNR_p(k) = \frac{\sqrt{T} A_T P_0 c \tau}{2\sqrt{2} NEP} h(k), \tag{4.20}$$

which shows that, for a given pulse energy $P_0\tau$ the signal-to-noise ratio is proportional to the square root of the sequence duration. Because the range resolution ΔR is given by $\Delta R = \frac{c\tau}{2}$, Eq. (4.20) can be written in term of system range resolution as

$$SNR_p(k) = \frac{\sqrt{T} A_T P_0 \Delta R}{\sqrt{2} NEP} h(k), \tag{4.21}$$

showing that transmitted power can be traded for range resolution in a linear way to keep the same signa-to-noise ratio.

Case b): sequence bandwidth narrower than fixed receiver bandwidth

In this case we assume a fixed receiver bandwidth B_0 and $\tau > \frac{1}{2B_0}$. Equation Eq. (4.14) can then be written as

$$SNR_p(k) = \frac{\sqrt{T}\, A_T P_0 c \tau^{1/2}}{4 NEP\, B_0^{1/2}} h(k), \tag{4.22}$$

or, taking into account the relationship between bit duration and range resolution,

$$SNR_p(k) = \frac{\sqrt{T}\, A_T P_0 \sqrt{2c}\, \sqrt{\Delta R}}{4 NEP\, B_0^{1/2}} h(k). \tag{4.23}$$

In this case a decrease of a given factor in P_0 must be compensated by an increase of the square of that factor in ΔR to keep the same signal-to-noise ratio.

By defining

$$SNR_i = \frac{SNR_p(k)}{h(k)} \tag{4.24}$$

(units m3), which can be considered as a system factor of merit with respect to the signal-to-noise ratio, the behavior of the normalized dimensionless factor of merit, defined as

$$SNR_{in} = \frac{SNR_i 4\sqrt{2}\, NEP\, B_0}{\sqrt{T}\, A_T P_o c}, \tag{4.25}$$

with respect to the normalized bit duration $2\tau B_0$ and normalized range resolution $\frac{4}{c} B_0 \Delta R$ is shown in fig. 4.3.

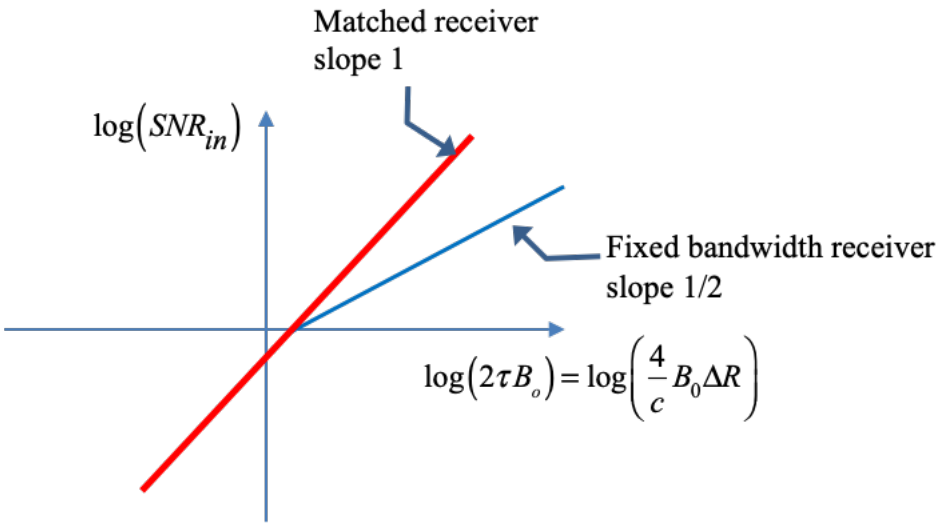


Fig. 4.3. Normalized factor of merit as a function of the normalized bit duration and normalized rage resolution for both types (matched and fixed bandwidth) types of receiver.

Note that for the matched receiver, the value of B_0 is immaterial, as it will vanish in the denormalization of SNR_{in}, whereas in the fixed bandwidth receiver τ cannot be shorter than $\frac{1}{2B_0}$ (and ΔR smaller than $\frac{c}{4B_0}$), that is

$$\Delta R > \frac{c}{4B_0} \tag{4.26}$$

We note as well that, everything else fixed, the signal-to-noise ratio is proportional to the square root of the sequence duration $T = N\tau$.

4.5. SIMULATION RESULTS

An estimation of the lidar behavior can be obtained using the formulation of the precedent section 4.3. This will approximate the maximum distance at which each class of target can be observed, like clouds, aerosols, etc. To evaluate the lidar system, it is implemented as a model in MATLAB environment. The model can compare the simulation with the results of formulation calculation. Modeling the system can test what happens when the lidar is working in a real implementation. Some features are the evaluation of the presence of the A/D converter, with different resolutions in bits, the loss of some bit due to a synchronization error, saturation in the receiver chain, non-linearities, etc.

For the calculations and simulations presented in this section, the lidar system and atmospheric parameters of table 4.2 have been assumed.

Parameters	
Laser diode peak power	0.125W
Wavelength	785 nm
Receiver Noise (NEP)	$1.2\ 10^{-13}$ W.Hz$^{-1/2}$
Responsivity of the photodetector	1.02×10^{7} V.W^{-1}
SNR A/D converter	88 dB
a-parameter (eq.	6
Number of bits of the A/D converter	16
Telescope effective area	0.017 m^2
Bit duration	675 ns
Extinction coefficient (α_1) (Clear atmosphere)	2×10^{-5} m^{-1}
Backscattering coefficient (β_1) (Clear atmosphere)	5×10^{-7} m^{-1} sr^{-1}
Photoreceiver electrical bandwidth	1.5 MHz
Range resolution	100 m
Narrowband filter transmittance	0.55
Mirror reflectivity	0.9
Window transmittance	0.85

Table 4.2. Assumed system parameters and atmosphere conditions for end-to-end performance simulation of the pseudorandom lidar system.

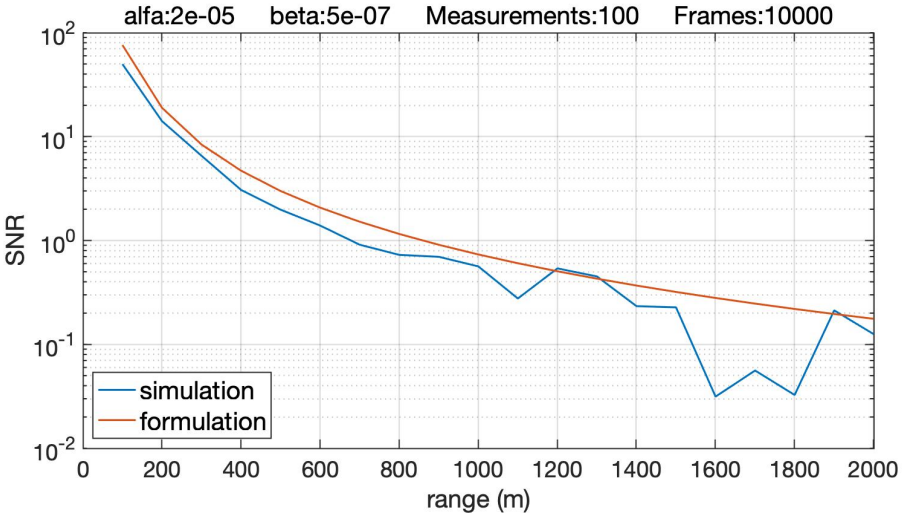
In the next figures (4.4, 4.5,) the SNR simulation results for some situations of the atmosphere (clear atmosphere, clouds, aerosols) are presented and compared with the results given by the theoretical formulation (Eq. (4.16)) are shown. The sequence used in the simulation is an M-sequence of 1023 bits length. Figures 4.6 and 4.7 shows the effect of reduction in the A/D converter of the number of bits and the effect of the saturation in the amplification chain. This last figure shows that the effect of saturation in the system is equivalent to a reduction of the number of bits of the A/D converter.

The most important drawback is that the amount of noise comes from the transimpedance amplifier and a noise reduction is difficult. In the simulations shown in this section no background radiation is added, supposing night observation conditions.

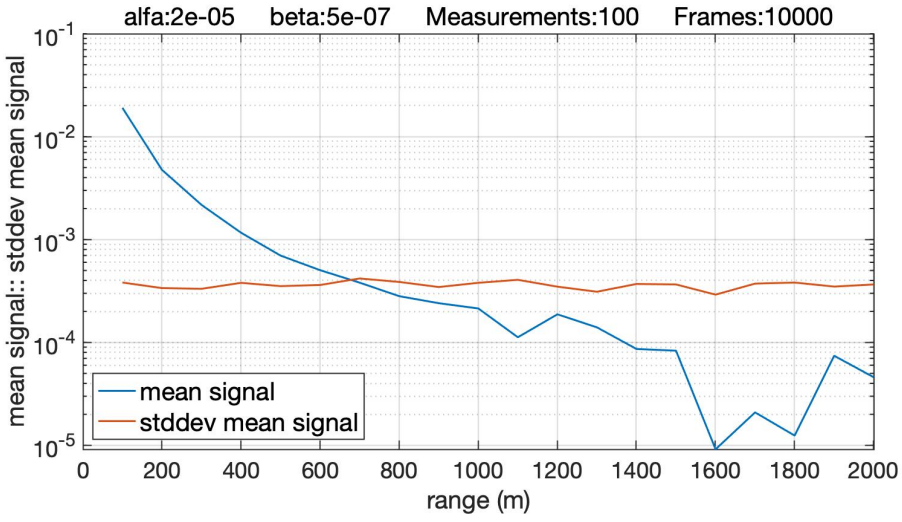
To obtain the simulated value for each distance the simulation is executed a number of times (100 times) with a number of frames each time (10000 frames), considering that a frame is equivalent to a sequence in the figures shown. The signal mean value and its

standard deviation are calculated too. The error in the estimation of the SNR value is proportional to $1/\sqrt{N}$ where N is the number of executions multiplied by the number of frames to calculate the SNR.

There is a small difference between the theoretical results given by the formulation of Eqs. (4.14) to (4.17) and the obtained result by simulation. This is because the simulator follows as closely as possible the hardware and software implementation of the real system. In practice a guard time is included in the processor design of the receiver system (see section **5.8** for details), which introduces more noise than in the ideal situation (no guard time).

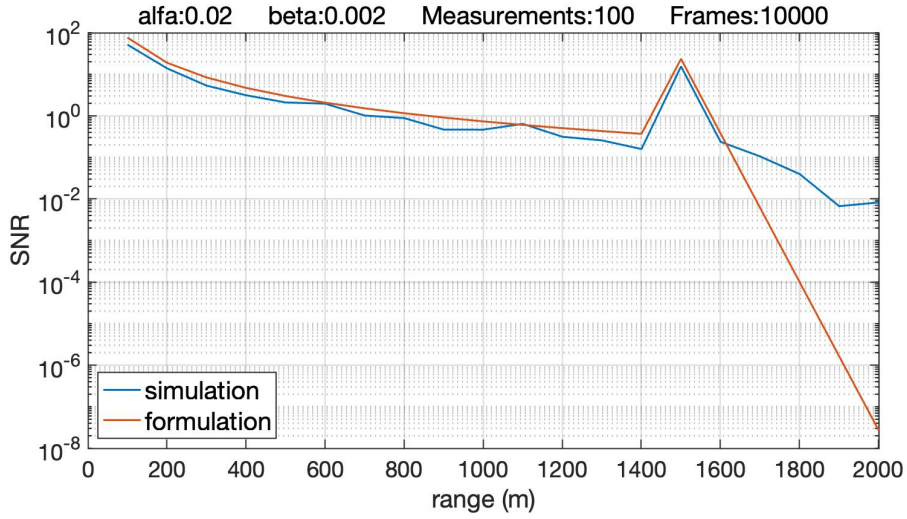


(a)

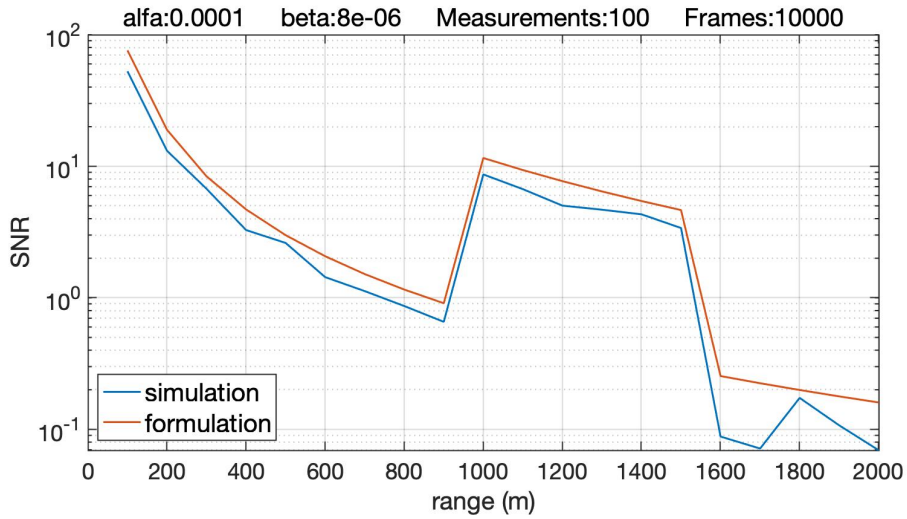


(b)

Fig 4.4. SNR simulation for clear atmosphere and the comparison with the results of the formulation developed in section 4.3. (a) Simulation and formulation agree. (b) Signal mean and variance, at 6000 m the SNR is 1 when the two values coincide. 100 simulations of the result of the accumulation of the returns from 10000 frames, infinite bits in A/D converter, sequence length of 1023 bits.

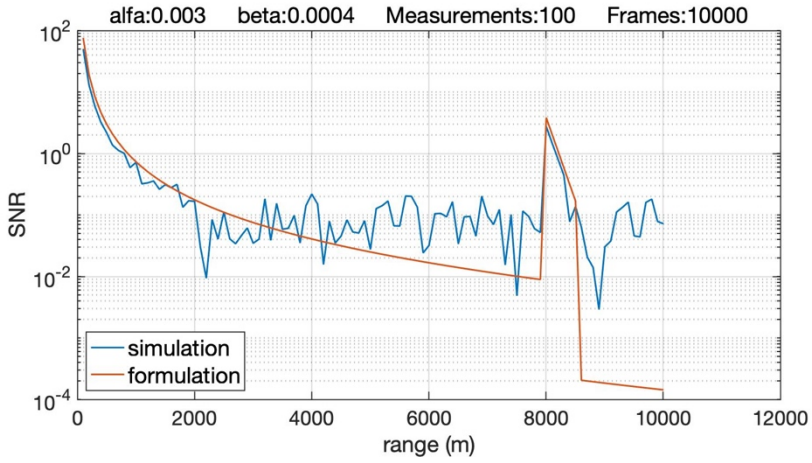


(a)

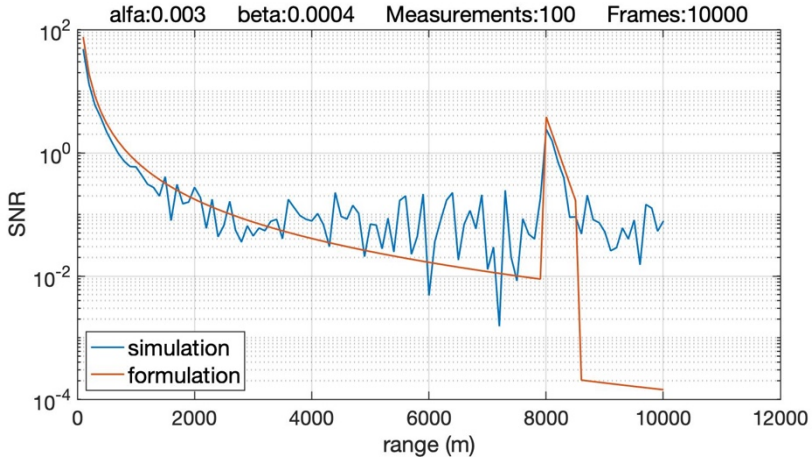


(b)

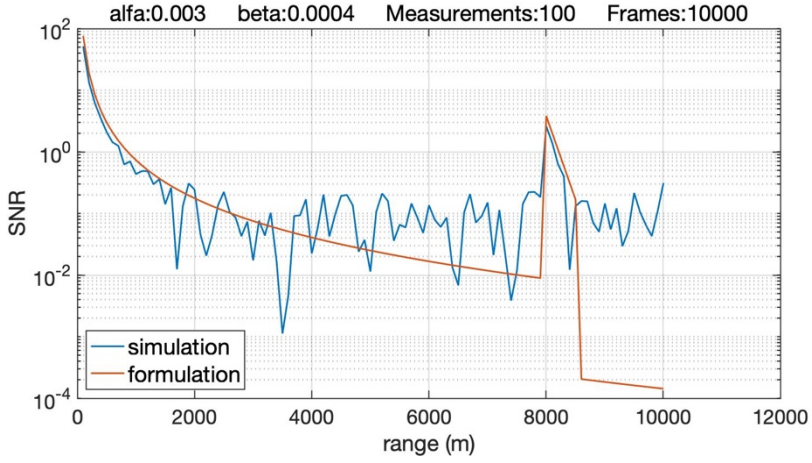
Fig 4.5. Simulation and comparison with the formulation developed. (a) cumulus. (b). Haze at low altitude. The simulation uses 100 iterations and 10000 sequences accumulated in each iteration. Infinite bits in A/D converter, sequence length of 1023 bits.



(a)



(b)



(c)

Fig 4.6. Effect of reducing the number of bits in the A/D converter maintaining without change the rest of parameters. (a) SNR with infinite bits in the A/D converter. (b) SNR with 16 bits in the A/D converter. (c) SNR with 8 bits in the A/D converter. 10000 sequences accumulated respectively in each iteration and 100 iterations. Infinite bits in A/D converter, sequence length of 1023 bits.

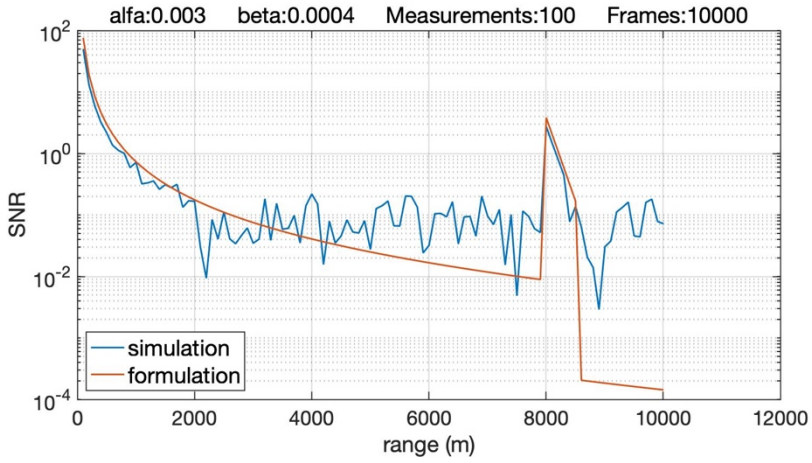


Fig. 4.7. Simulation and comparison with the formulation developed. Effect of the saturation in the amplifier chain which is very similar to the reduction in the number of bits in the A/D converter, comparing with the figure 4.6 (c). Thin and high cloud with 10000 sequences accumulated respectively in each iteration. The simulation uses 100 iterations. Infinite bits in A/D converter, sequence length of 1023 bits.

The results of calculations and simulations in the figures shows that with the parameters of the lidar defined in the table 4.2, the usual targets in reasonable conditions to reach a good signal to noise ratio are:

- Presence of clouds, until altitudes of 8-10 km depending on the cloud thickness, type and of day or night observation. During daytime the observation of clouds is difficult due to the background radiation that saturates the photoreceiver with the parameters of filter and telescope field of view.
- Haze, aerosols, at low altitudes, until 2-3 km depending on the backscattering coefficient and of day or night observation. In this case the highest sensitivity is obtained during nighttime but in the real context the aerosols due to local conditions (are present during the day.
- The number of bits of the A/D converter influences the SNR at the output of the system. 16 bits is a good value but 8 bits can be tolerable depending on the kind of measurements needed. In general clouds shows more robust measurements.

The simulator can evaluate other cases as for example errors of synchronism in the transmission-reception process and losses of bits, the effect of background radiation, the non-linearities of the amplifiers, etc.,

5. SYSTEM DESIGN

5.1. INTRODUCTION

To demonstrate the principles and developments laid out in the preceding chapters, we have have designed, built and operated a system based on those principles. The design of the system is discussed in this chapter.

The designed system is based on a Digital Signal Processor (DSP) and complemented by circuitry where it is necessary, especially for interfacing the different electronic devices.

The whole system is divided into different functional blocks. Some of these blocks are implemented as hardware and other blocks are inside the software loaded into the DSP. The design could be referred to as a “software defined lidar”, which means that some functions of the lidar can be changed by software (e.g. by default it is configured as a CW lidar modulated with a given pseudorandom sequence, but it can be used with different types of sequence (see chapter 3), or even as a pulsed lidar loading another software, for example). A summary of the different block functions is detailed next.

5.2. BLOCK LAYOUT

Figure 5.1 shows the basic block layout of the lidar system with the interactions between them. After this section 5.2 the next sections will be explained with more detail the components of each block.

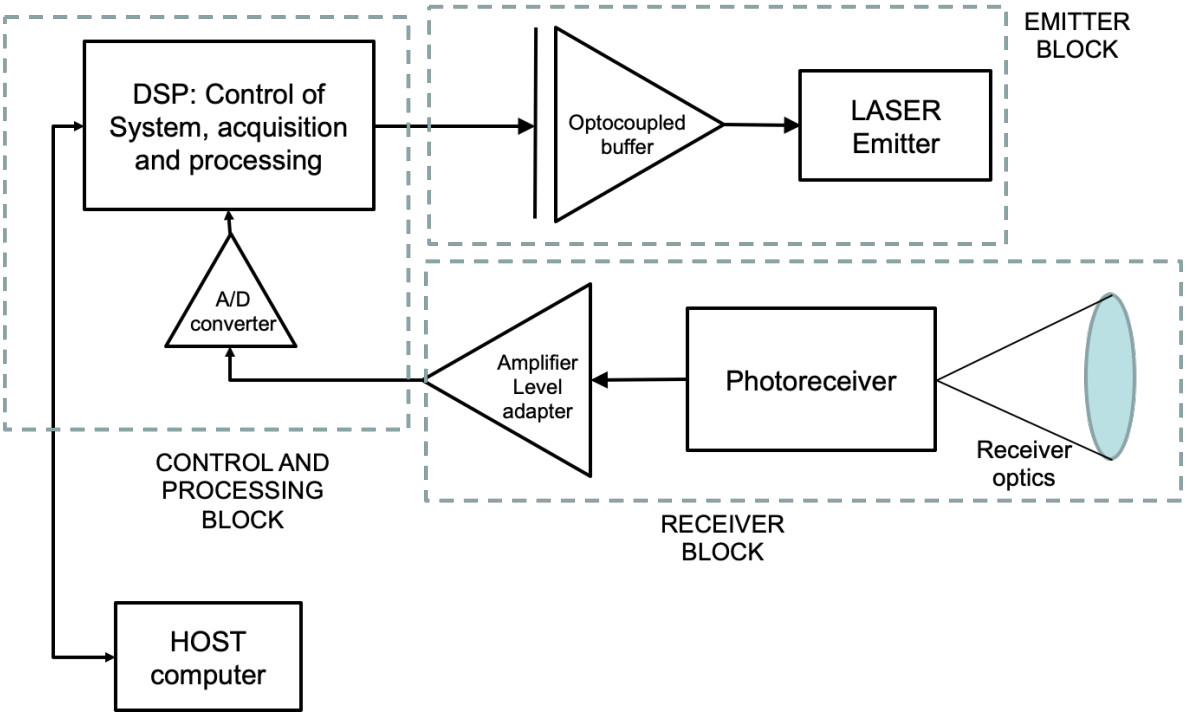


Fig 5.1. General block layout of the complete lidar system.

Block 1: System control

This is the core block of the system, built around a DSP. This block controls the whole system, it generates the sequence that modulates the laser and it processes the data coming from the photoreceiver. It generates the necessary synchronization between the emitted sequence and the received data. The process of sending and receiving data synchronously can be done until a maximum frequency, which is limited by the rate at which the DSP can control the laser, the frequency at which the laser can be modulated and the speed of the A/D converter (reading data from the photoreceiver). Because of the high calculation rates, the execution time of the correlation routine (or a possible implementation of an inversion routine, which is out of the scope of this work) is only a small fraction of the processor time (in the order of milliseconds).

The DSP control works in real time. The processor has to send and receive the signals at the same time. The synchronization is critical, because a delay in only one sample of the transmitted or received sequence would generate a sequence misalignment that would increase the noise and would not allow detecting the signal from the target.

The execution of the correlation algorithm is very fast in a DSP and therefore it is possible to show the received data in real time if needed. The execution of the correlation algorithm is very fast in a DSP and therefore it is possible to show the received data in real time if needed, for example for the lidar optical alignment.

In the measurement mode it is necessary to integrate for longer times to have more energy received, from some seconds to some minutes. The amount of sequences sent and received each second is in the order of thousands (with a sequence length a bit longer than 0.5 ms, nearly 2000 sequences are sent and received per second). The integration process has to be executed to have a minimum SNR at the output of processed data.

Block 2: Laser emitter

This block includes the laser and the driver to interface with the DSP (to accommodate the card output signal to the requirements of the laser unit). In the implemented system, the laser modulator needs 5 V and 100 mA, which is equivalent to 50 Ohm input impedance. The output pins of the DSPs cards are not able to provide this current level and they works with 3.3 V logical levels. Therefore, it is necessary to insert an additional step to drive the laser modulator, which is challenging given the required high frequencies and high currents.

The frequency of the signal modulating the laser is in the range of the MHz and the rise and fall times of the pulses have to be very short to maintain the squared shape as much as possible. If the signal is smoothed the practical effect is a reduction of the transmitted energy (area under the signal). The implication of these requirements is that a DC-coupled wide-band amplifier is needed with some MHz bandwidth and large amplitude signals at the output (5V of amplitude to modulate the laser module).

Block 3: Receiver system

The receiver block includes the photoreceiver — a module composed of an avalanche photodiode followed by a transimpedance amplifier — an additional amplifier and the A/D converter. As it has been explained in chapter 4, the main function of the additional amplifier is to adapt levels from the APD module to the A/D converter in order to use the maximum number of A/D converter bits, providing some extra gain if it is possible, and therefore minimizing the effect of the A/D converter noise. Working with APD detectors a very compact and small receiver circuit can be designed. An operational amplifier with more than 1 GHz of bandwidth and low noise is the core of this amplifier. It is necessary to take care in the design process because this kind of operational amplifiers are designed to work at important bandwidth-gain products, which could lead to instabilities. In the detailed design of this block some issues related to this amplifier will be explained.

The A/D converter is a high-speed one. In fact, one of practical limiting factors of the hardware system operation is the maximum rate at which the A/D converter can work. In the design of the version 1, a16 bits A/D converter has been employed. In the version 2 the A/D converted has 12 bits because it is integrated in the DSP chip.

In the receiver very weak signals have to be detected at high-speed rate. Therefore, it is necessary to have an important bandwidth with the lowest possible noise. One of the most difficult problems to solve in this sensitive circuit is the presence of ground loops that can introduce a feedback with levels equal o higher that the received signal from the atmosphere.

Block 4: The processing system

The function of this block is to calculate the correlation between sequences. As section 6.1. explains, two different DSP have been tested in two different implementation versions. Due to high computing power of the present-day processors this time is small compared to the integration time. To calculate a circular correlation of two N-bit sequences, the number of operations would be N^2 bits if the direct calculation is implemented. But in addition, it is not necessary to calculate the complete correlation, only for the maximum expected delay corresponding to the farthest possible expected target. For example, with 2048 bits, and assuming that the maximum delay would be 200 samples —200 samples at 70 m for each sample is 14 km—, the number of operations is 2048x200=409600. This is a small number related to the flops of the processor. The TMS320C6713 DSP has a minimum calculation power of 1000 MFlops, which means the calculation time as short as 400 μs.

Other system blocks

To display the system output, that is, the data received and processed, to save it, etc., it is necessary to use a computer as interface. The configuration of the lidar operation parameters, loading of software and commands has to be done from this external computer.

A mechanical design is necessary with two functions:

- To keep the receiving telescope and the laser aligned. During the operation no differential movements are allowed, for example those caused by the wind or by temperature changes.
- To achieve the necessary and precise movement to align the laser with the field of view of the receiver (telescope + photodetector), usually in the range of only some arc minutes.

5.3. SYSTEM CONTROL

As it has been explained, the DSP is the core that controls the whole lidar system. It generates the timing in the system, it prepares the data based on a pseudorandom sequence, it adjusts the offset of the receiver, it controls the A/D conversion, it processes the received data and it communicates with the host. The employed DSPs are a single processor configuration, but there is the possibility of adopting multiprocessor or multicore strategies as a plan of increasing capabilities. In the configuration with one processor only, it is necessary to work absolutely in a sequential manner and to design software processes accurately to maximize the time devoted to operate the laser and to properly synchronizing the emission and reception of signals. In fact, a great part of system time is dedicated to transmit the laser signal and to receive it. These two operations must be done at the same time exactly, as this is critical to have a perfect sampling in the time axis of received bits

Control and Timing of the system

System control and timing ensures that all processes run in the correct order. Basically, the order of the processes is:

1. Read the sequence to be transmitted by the laser. In the practice, as the sequence is fixed during the measure, it is stored in the software.
2. Adjustment of the offset and the gain of the receiver system if it is necessary.
3. Sending of the data to the laser emitter.
4. Controlling the A/D conversion and reading the received data.
5. Accumulating the received data.

6. Calculating the correlation between the received signal and the original sequence. An optional implementation is to insert the inversion algorithm to recover the backscattering profile inside the software of the DSP.
7. Sending of correlator output to display and storage unit, a PC and an SBC in this implemented laboratory prototypes.

It is necessary to repeat the steps 3 to 5 until the number of received sequences achieves the programmed number (established previously in the parameters of the system before starting the laser emission).

Data

The modulating sequence does not change during the integration time. Therefore, to save time in the processor the sequence is generated only once and saved for subsequent use. This procedure can be applied to the necessary number of sequences that could be used. Obviously, the sequence is independent of the rate used to transmit. Changing the length or the kind of the employed sequence only changes the distribution of 0's and 1's. The most practical method in the implementation is to generate a library of sequences for later use. In the experiments a length from 1023 to 4095 bits for the binary sequence are employed.

Data processing

Data processing is done in three parts.

The first part consists into accumulating the received signal. This first processing step is necessary to gather enough energy to increase the signal to noise ratio before the main processing. It is important to consider that the signal-to-noise ratio increases as the square root of the number of integrated sequences (see section 4.3). There are practical limits to data accumulation. The limit is the time resolution of the observation because the atmosphere is not static and therefore there is trade-off between the integration time and the tracking of atmosphere dynamics.

The second part of the processing consists in calculating the correlation of the received signal with the original sequence as it is detailed in the theoretical part of this thesis (see chapters 3 and 4).

The profile obtained from the correlation process is the profile of power returned from the atmosphere, without any kind of correction.

The third step is correcting the power by range multiplying by the squared range z^2, obtaining the attenuated backscattering as it has been defined in the equation (2.6). This processing procedure and the data display is done in an external computer, but it could be integrated in the control computer of the system or in the DSP, but the use of external analysis increases the flexibility in the process.

One more optional step is necessary to retrieve the optical coefficients (extinction and backscatter coefficients) of the aerosols or clouds in the atmosphere by applying an inversion algorithm, using standard algorithms for do it. This last step is not included in this thesis, which focus on the lidar system.

5.4. LIDAR EMITTER

The lidar emitter consists of three parts:

- DSP signal generation
- Driver circuit
- Laser module

Sequence generation

The DSP generates the sequence that is sent to the driver, which adapts the levels of the DSP logic circuits to the voltage and intensity levels necessary to modulate the laser.

The data sequence for modulating the laser is software-generated and sent out from the DSP to the laser using the different options of hardware. For practical reasons related to save memory, in the software coding the sequence is packed in groups of words. A sequence of 2047 bits, for example, is packed in 128 packets of 16-bit words, therefore the last bit is filled with a 0 value. Two guard words of 0's (32 zero bits) are added at the

beginning of the sequence or a word of 16 bits, depending the lidar version, to account for possible delays in the complete emission-reception chain, as shown in figure 5.2.

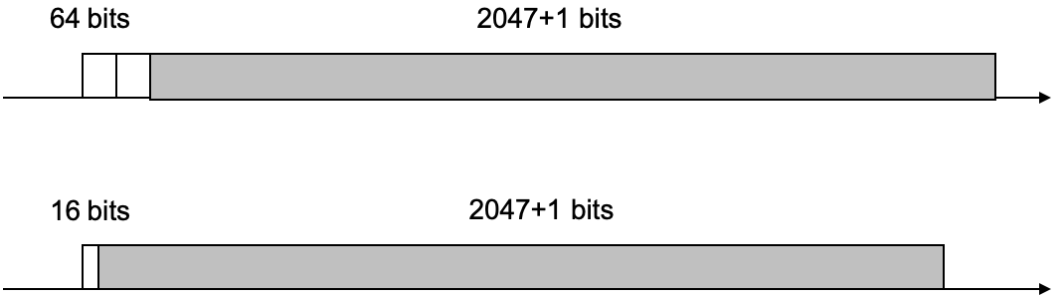


Fig 5.2. The structure of the transmitted frame. A sequence of 2047 bits is completed with one zero at the end to have an integer number of 32- or 16-bits words.

Driver circuit

It is necessary to translate the signal output from the DSP to the needed voltage and current levels of the laser modules. To provide this function a driver circuit has been designed to modulate the lasers.

The driver module must operate in the 3-4 MHz bandwidth range at least (in practice because of the squared pulses, a wider bandwidth would be necessary to obtain pulses as sharp as possible), with high currents. This large bandwidth with these currents implies a difficult design.

If the energy sent by a pulse is calculated, it is possible to distinguish two extreme situations. If Vmax is the maximum value of the modulating signal, the transmitted energy for a squared pulse is proportional to $Vmax^2$. If the pulse is shaped by a squared sine the transmitted energy is proportional, with the same proportionality constant, to $Vmax^2/2$. This shows that depending on the bandwidth of the emitter system as much as 50% of the maximum power is lost (figure 5.3). The real situation will be between the two cases. The sharper the pulse, the higher the energy sent to the atmosphere and received by the system.

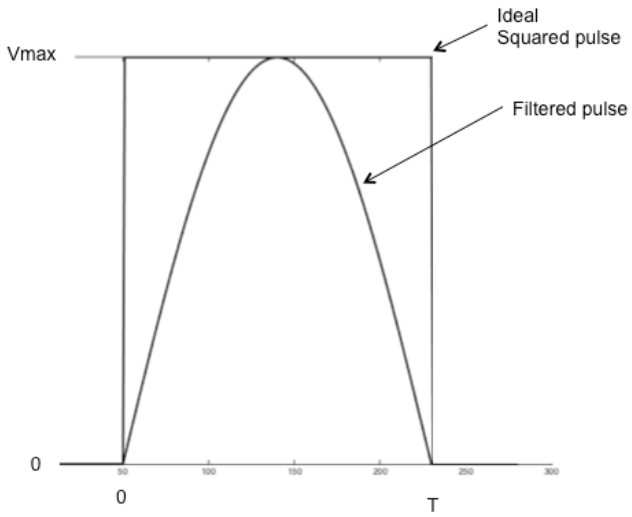


Fig. 5.3. Pulse shape. The energy of the squared pulse is $(Vmax)^2T$ (Area of the pulse-shape under the function). The energy of the filtered pulse, assuming that the result is a sine function (the first harmonic) is $(Vmax)^2T/2$. Therefore, the ratio of the energies is 1/2.

Laser modules

Two different laser modules have been used in this work. One of them is a 10 mW red visible laser and the other is an IR 125 mW laser. The first was used to test and to align the system, and the second as operative laser for atmospheric measurements. Chapter 6, dedicated to the prototype, presents the specifications of both laser modules.

5.5. LIDAR RECEIVER

The lidar receiver is made of five parts:

- Telescope
- APD and first stage of conversion and amplification
- Second amplifier and filtering (optional)
- Analogue-to-Digital converter
- DSP processing

After the DSP processes the data, the result is sent to an external system, like a general-purpose computer.

The telescope

The receiver telescope presents am equilibrium between two factors: the diameter (size and weight) and the simplicity and portability.

A larger diameter of the telescope will offer more signal collected to the APD module. Ideally a big telescope would be useful. But as the diameter increases, the cost, the volume and the weight increase too and decreases the portability. The telescope diameter depends on the desired sensitivity of the instrument and the available space and the portability itself. In fact, it is a balance between these two factors. Two telescopes have been tested: A 15 cm diameter Newtonian and a 20 cm diameter mirror with the APD in the focus without any secondary mirror to eliminate power losses.

APD and first stage of the receiver

The optical receiver used is an APD module Hamamatsu C5460SPL5343, coupled to the telescope. This module contains the APD itself, the transimpedance amplifier, and a first amplification stage. All is made in a single block integrating the High Voltage source for the APD, the whole being powered by a single 5 V power line.

To simplify the photodiode is placed in focal plane of the telescope.

Second amplifier

The APD module output is amplified before reaching the analogue-to-digital converter.

A one-stage amplifier with adjustable gain and a controlled offset input amplifies the signal. The configuration is in addition mode, that is, one input is the signal received from the APD, and the other input is a DC level to compensate the offset due to the sky background radiation. This DC level is generated by the DSP and an interface circuit.

As it has been explained in the section 0, when the amplifier has a large bandwidth the transmitted (and received) pulses are sharper, and therefore more energy is available. In fact, the bandwidth of the APD module in this prototype (1.5 MHz) sets the practical limit in the bandwidth.

Offset compensation

The offset-adjustment process is done before each integration period, but it takes a short interval with respect to the measurement time. Figure 5.4 shows the complete sequence of adjusting the offset, measuring and processing the received signal. If fact, the measuring interval lasts much more time than the other two processes. Although this thesis does not focus on the system optimization, it is noteworthy that processing and offset adjustment take a virtually negligible time compared to the measurement (sending the sequence and receiving its delayed echoes) time.

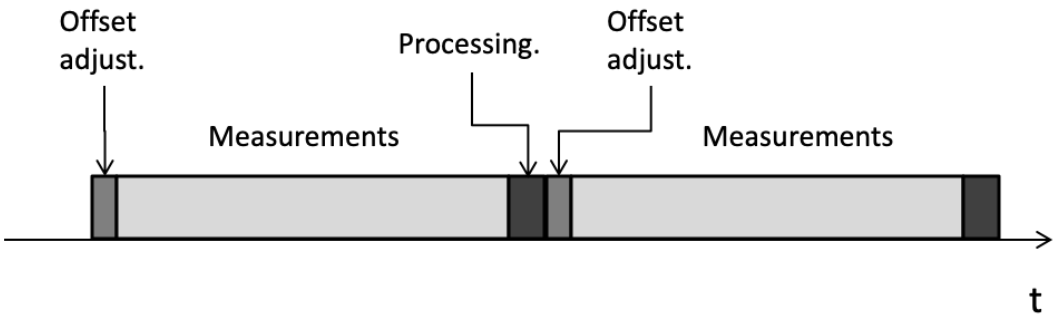


Fig 5.4. The sequence of adjusting the offset in the amplifier, measuring and processing. Each measurement can contain tens of thousands of sequences.

Meanwhile the first version of the lidar uses this second external amplifier after the APD photoreceiver to compensate the offset, the second version of the lidar do not use this second amplifier with the objective of testing a simpler system. The light of the sky can saturate the amplifier inside the APD module and it is not accessible, therefore the offset to avoid saturation in this first stage cannot be applied. A completely new design of the module with APD and transimpedance amplifier would has to be done, but in this thesis, it is out of the purpose.

Amplification and saturation

As it has been explained, the amplification is performed by a single stage amplifier, made with an operational amplifier in non-inverting configuration and adjustable gain. The gain can be changed by means of a fixed resistor, but it would be possible to design a circuit for controlling this gain from the software running on the DSP.

The first experiments showed an effect of ghost echoes appearing after processing the received signal that did not correspond to real targets (fig. 3.18). The investigation about the origin of these ghost echoes led to discover an important issue when a lidar system according to this design is working in large-signal mode, which can occur when the echoes come from a near target. The explanation has been developed in chapter 3.

This effect puts a limit in the use of this lidar design for very close solid targets unless large dynamic-range electronics are employed to avoid saturation and non-linear effects.

A/D conversion

In this type of lidar, the emission and the reception of the signals are done at the same time, therefore the A/D conversion is done at the same rate as the emission. Every sample sent implies another sample read from the A/D converter in the period fixed. This means that the DSP has to control these two processes accurately.

5.6. POWER SUPPLY AND MECHANICS

Power supplies

Several power supplies are used for different parts of the circuit, because of the different voltages needed. The implemented prototype is a laboratory one and no complete integration of elements has been sought. The presence of different power supplies as different elements creates an additional problem due to the possible presence of glaring or hidden ground loops. A careful control of the connected ground lines is necessary to avoid problems, especially in the receiver blocks, and a careful work has been carried out to locate and remove ground loops.

Mechanical support

Mechanical design require precision. The support of the telescope is fixed and it is not necessary a sophisticated mount. The most important consideration in the mechanical

part is to maintain the position of the laser fixed with respect to the optical axis of the telescope, while allowing the necessary and fine adjustment.

For the experiments the lidar is pointing to the zenith (90º elevation) but it is possible to do experiments at other altitude angles. In this case the mechanical design has to be robust to avoid flexures of the block telescope-laser when it is pointing to another altitude angle. As this requires a good mechanical design to test the lidar, the zenith angle is preferred.

5.7. ALIGNMENT PROCEDURE

With the characteristics of the optical system and photodetector used in the first version of the lidar the calculations of the field of view are done:

- Telescope diameter: 150 mm
- Telescope focal length: 750 mm
- Diameter of photodetector: 3 mm

$$FoV = \frac{d_{phot}}{FL_{tel}} , \qquad (5.1)$$

where *FoV* is the resultant field of view, d_{phot} is the diameter of the photodiode (APD) and FL_{tel} is focal length of the telescope.

The result is a field of View of 4 mrad (13.8 arc minutes).

Using the characteristics of the optical system of the second version of the lidar the calculations are the next

- Mirror diameter: 200 mm
- Mirror focal length: 400 mm
- Diameter of photodetector: 3 mm

As in the first optical configuration, the field of view is calculated and the result is a field of view of 7.5 mrad (25.8 arc minutes).

An alignment procedure is designed in order to be able of put the laser beam inside the field of view of the receiver. A precision mechanic mount is used to steer the laser beam.

The sensitivity of this element is 3 arc seconds (14.5 μrad) for adjusting the position angle, according to the manufacturer, which is enough for this application.

The figure 5.5 shows the geometry involved in this optical configuration and the different situations in the alignment of laser beam with respect to the field of view of the receiver (assuming the divergence of the former much smaller than the latter). The first and second situations are the only possible to obtain a correct measurement. The fourth can detect the backscatter in short distances but obviously it is not useful to measure correctly in the full range of the system.

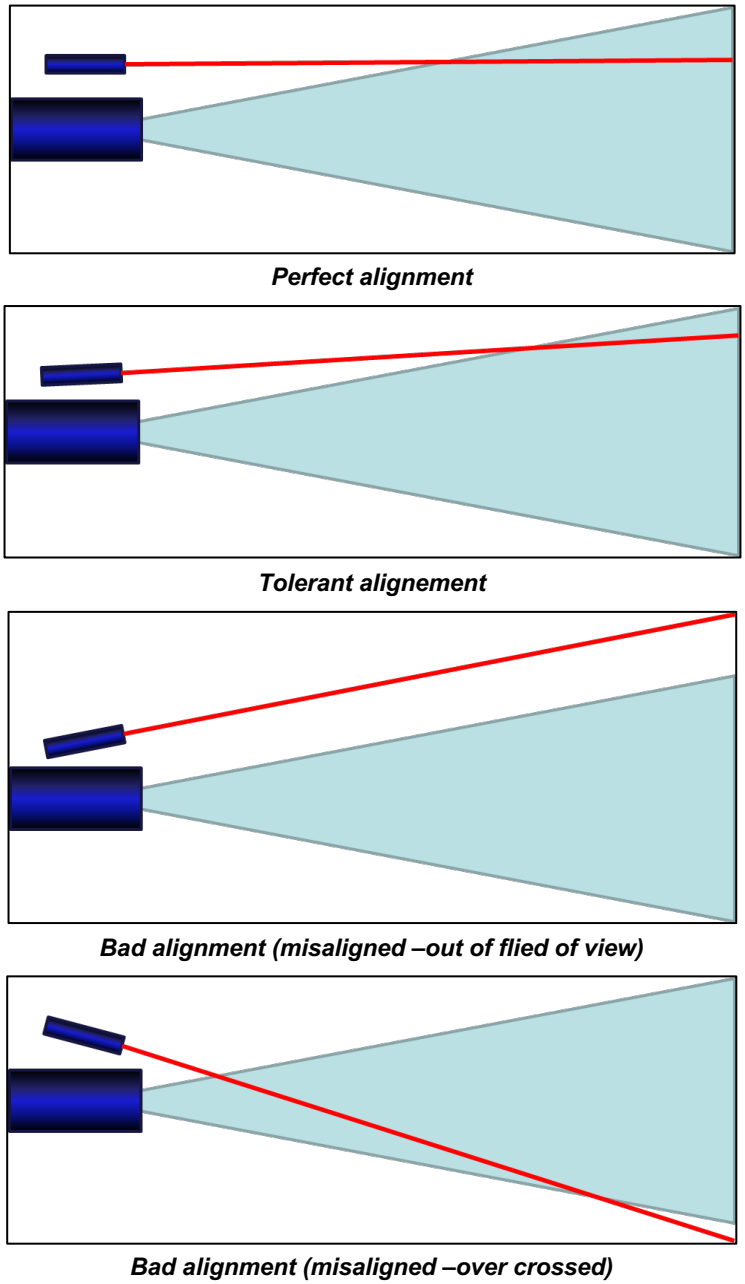


Fig. 5.5. Different situations in the alignment of laser beam with respect to the field of view of the receiver telescope.

5.8. PROCESSING, MEASUREMENT AND DISPLAY OF THE RECEIVED SIGNALS

The use of a Digital Signal Processor (DSP) in the lidar offers the advantage of a flexible processing system. In fact, the flexibility of controlling the system by software opens up a wide range of possibilities, as it is listed in the last section of this document about the future work. Only the correlation function has been explored in this work.

The received signal

The received signal is an addition of the train of sequences received from each cell of resolution delayed in the time (fig. 5.6). The number of cycles can be 1 to N+1, depending of the software implementation. In the evaluation of the lidar only 1 cycle has been used in each measure.

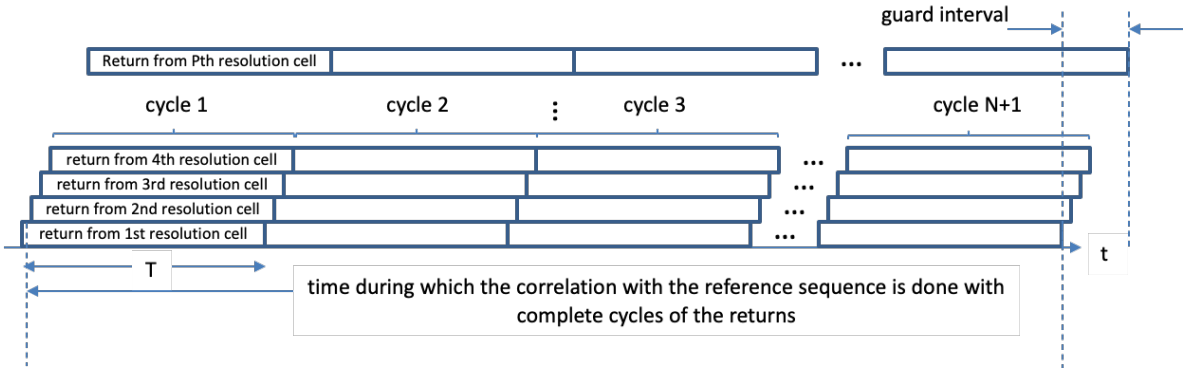


Fig 5.6. Addition of the received sequences of duration T in each cycle of measure.

The processing of the signal

The lidar demonstrated in this thesis is basically created around the use of pseudorandom sequences and the correlation function to detect the range and

backscattering of the target, be it an aerosol layer, a cloud or any kind of target in a general case. In fact, this is a particular case of the general deconvolution problem.

The correlation of the return signal with a reference sequence can be performed very fast in a DSP because these processors are designed to execute very fast the multiplication and addition operations, which are usually core operations in signal processing (correlation, filtering, FFT calculations, etc.).

For example, according to manufacturer's data [39], a Texas Instruments TMS320C62x will calculate $\mathbf{n_r}$ points of an autocorrelation of $\mathbf{n_x}$ points sequence using

$$n = n_r \frac{n_x}{2} + 31 + \left(\frac{n_r}{2} - 1 \right) , \tag{5.2}$$

where ***n*** is the necessary number of cycles of the processor. Assuming a 4.4 ns cycle time (225 MHz clock speed), and 512 calculated points (25.6 km range according to a theorical 50-m resolution assumption), the number of cycles would be 262430 and the processing time would be 1.2 ms for a 2^{10} -bit sequence.

Therefore, the processing time is much less than the acquisition and measuring time, which depends on the desired integration time chosen, typically from some seconds to some minutes. The DSP processing time is not a burden when the correlation is used to calculate the response of the atmosphere. This fast processing aspect can be interesting in solid target measurements to achieve real time operation in other kinds of application.

In the actual implementation the processing time will be longer because to achieve the maximum speed it is necessary a good optimization for the designed code.

Calculation of the correlation in the system

The properties of the correlation discussed in chapter 3 are well defined for the circular correlation, which represents one period of an infinite correlation.

The practical implementation of the circular correlation in a system like this has a particularity that must be considered. If the sent sequence and the reference sequence are of the same length the result is a correct accommodation to the sequences and circular correlation of the two sequences as shown in figure 5.7 (a). However, in the real system it is necessary to wait for a time until the sequence returns completely from the maximum observation range; this is called the guard time in fig. 5.7 (b). As a result, the perfect length accommodation of the sequences disappears and additional noise is

introduced in the system. The reduction of this noise depends on the amount of integrated time.

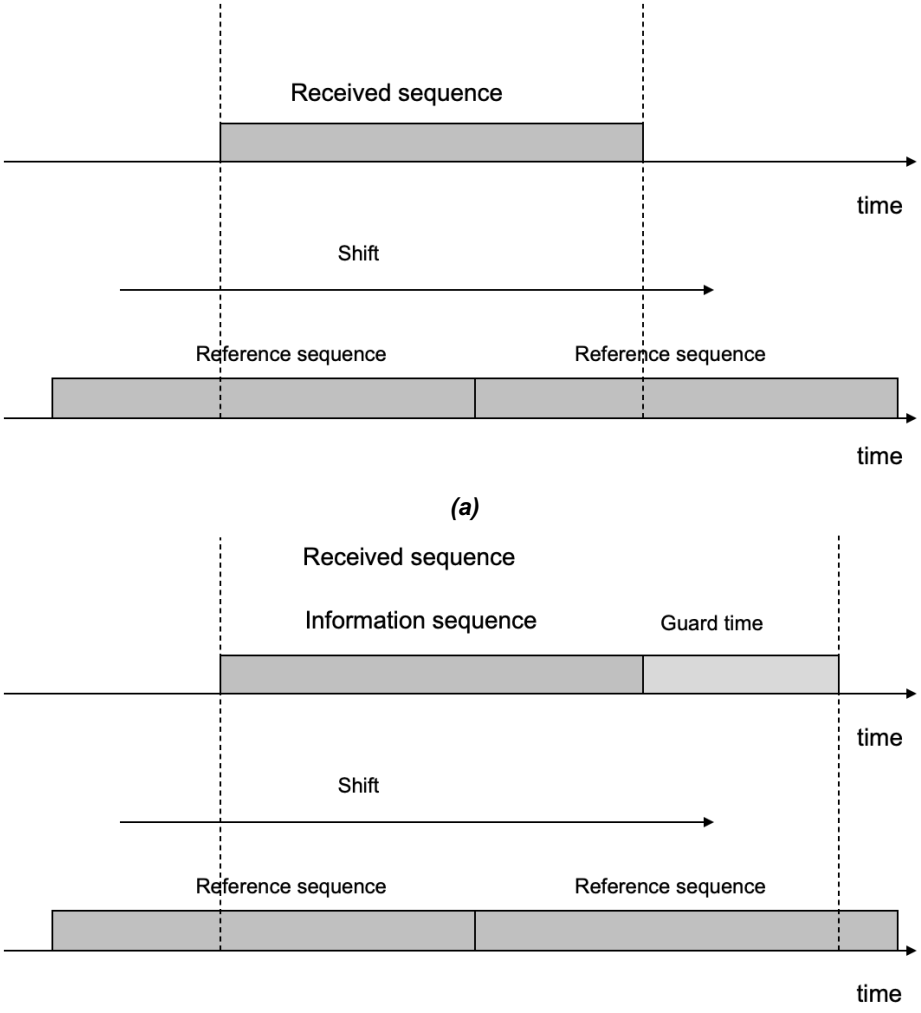


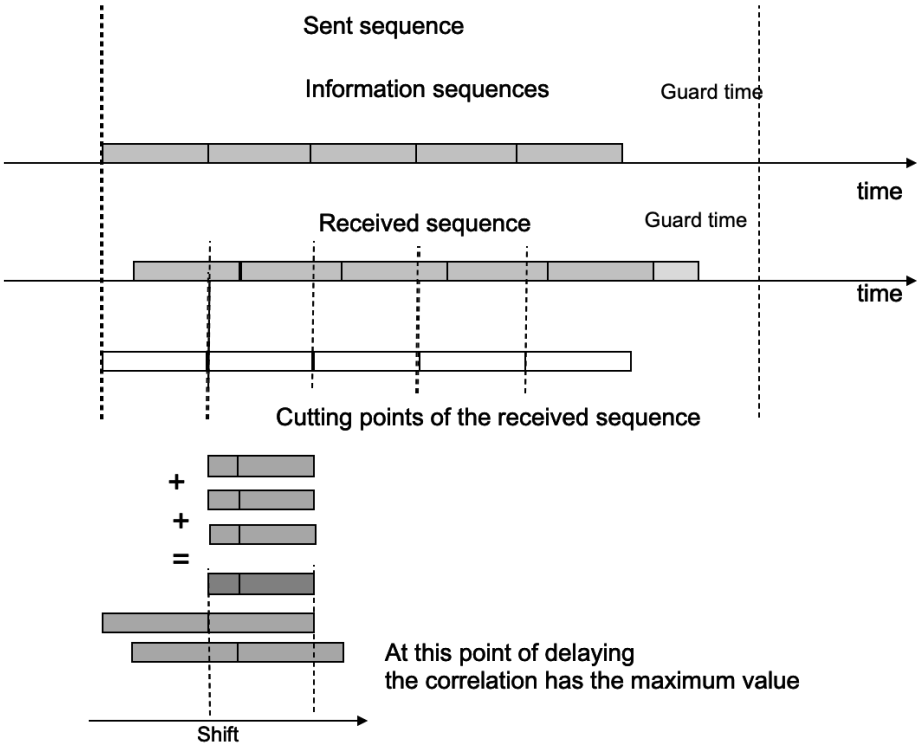
Fig. 5.7. The circular correlation when the two sequences have the same length (a) and when they do not (b)

The problem in the received signal is the incomplete received sequences, that when the correlation is done the result is different from the zero desired and therefore the result value is non predictable.

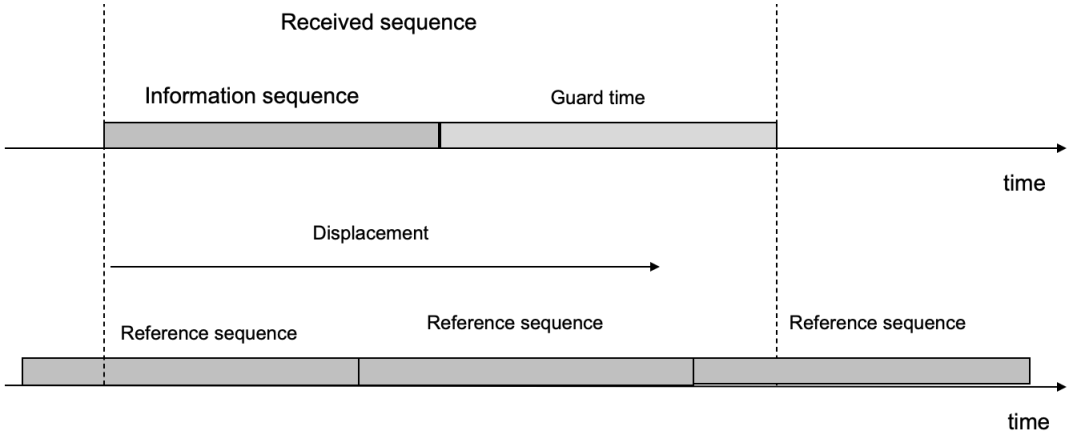
A first possible and simple solution is to adjust the guard time to have the exactly same length as the reference sequence. This solution reduces the efficiency of the system — because the 50% of time nothing is emitted— but solves the problem of the unbalanced sequence due to guard time (Fig. 5.8 (b)). This solution of using a guard time of the same length than the sequence shows that when the reference sequences are delayed by the

same time as the received sequences the correlation is done between sequences of the same length.

A second and more efficient solution is to send a set of sequences without any guard time, and at reception to cut exactly at the length of the sequence and to add them, discarding the first and the last segment because they are not complete. Figure 5.8(a) shows this schema of processing. The received concatenated sequence is cut into pieces of the same length than the modulating sequence starting at the beginning. After accumulating these sequences, the result is correlated with the reference sequence in the same way as using the process of receiving, accumulating and adding the sequences one by one.



(a)



(b)

Fig 5.8. The two proposed ways to deal with guard times discussed in the text. (a) Guard time at the end of chained sequence cycles. (b) Guard time of the same duration as the sequence.

The larger the number of continuous sequences transmitted the more efficient is the system, because the proportion of discarded sequences to the total number of transmitted ones is lower.

$$Efficency = \frac{M_{TOT} - M_{DISCH}}{M_{TOT}} \tag{5.3}$$

For example, a train of 10 sequences would have an efficiency of (10-2)/10 =80%. But a train of 100 sequences would increase the efficiency up to (100-2)/100=98%.

The implementation of the second solution requires very well-designed software to get a perfect synchronization without losing bits and to use correctly the memory of the processor and no to limit the acquisition of the received and acquired sequence.

At present, sequences are sent and received one by one, which implies a reduction of the efficiency. If a sequence of 1024 bits is used and a guard time of 1024 more bits is applied, only 50% of the time is used to send energy to the atmosphere.

At this point there is a compromise between the length of the guard time and the amount of extra noise introduced in the receiver system. A length of 0 bits would introduce no extra noise but using a guard time of the same length of the sequence duplicates the amount of noise. This is clearly visible in the simulation (see section 4.5) because it uses the same implementation that the operating lidar, that is, the guard time has the same length than the sequence. This explains the difference of a between the theorical SNR (the maximum achievable) and the real (simulated) SNR.

Effects of truncation in the mathematical operations

When data is processed the effects of the limited values of variables in a computer (due to the number of bits in number types) must be considered. As it was discussed in chapter 3, the effect of saturation in amplifiers is detrimental because it causes the apparition of non-desired “ghost” echoes in the processor output. This same effect can occur if the data type variables in the designed software are not selected correctly. If these variables reach the limit, causing overflows because of the accumulation of sequences, the result could be the same that the saturation effect in the amplifiers.

If the A/D converter provides integer words of 16 bits —65536 values— when the system is integrating thousands or tens of thousands of sequences there is the risk that the variables intervening in this process cause overflow. For example, adding 10,000 sequences the maximum number could be 10,000 x 65,535=655,350,000, which needs at least 30 bits to be represented. Therefore a 32-bit variable would be enough. But if the integration is done with 200,000 sequences, the maximum value can be greater than $13x10^9$ and 34 bits would be needed. Then, a 32-bit variable is not enough and at least a 40-bit variable would be needed. For this reason, the type of numeric variables has to be selected correctly. In this design floating point variables are preferred because the limit is high and the processor is oriented to floating point operations. If there is enough memory in the processor, the precision of the necessary operations of accumulation and correlation is increased if variables of double types are chosen instead of float types.

The largest the number of bits for each kind of variable used, the more precise is the calculation, which prevents unexpected effects that could be unintendedly created in the processing part of the lidar chain [17].

6. PROTOTYPE OF THE LIDAR SYSTEM

To prove the proposed features and capabilities of the DSP-based lidar working in real conditions, a prototype has been built. The different parts, explained in the previous chapter, have to be evaluated, which is done in the next sections. The figure 6.1 shows the second lidar version in which the external amplifier to the photoreceiver is not present.

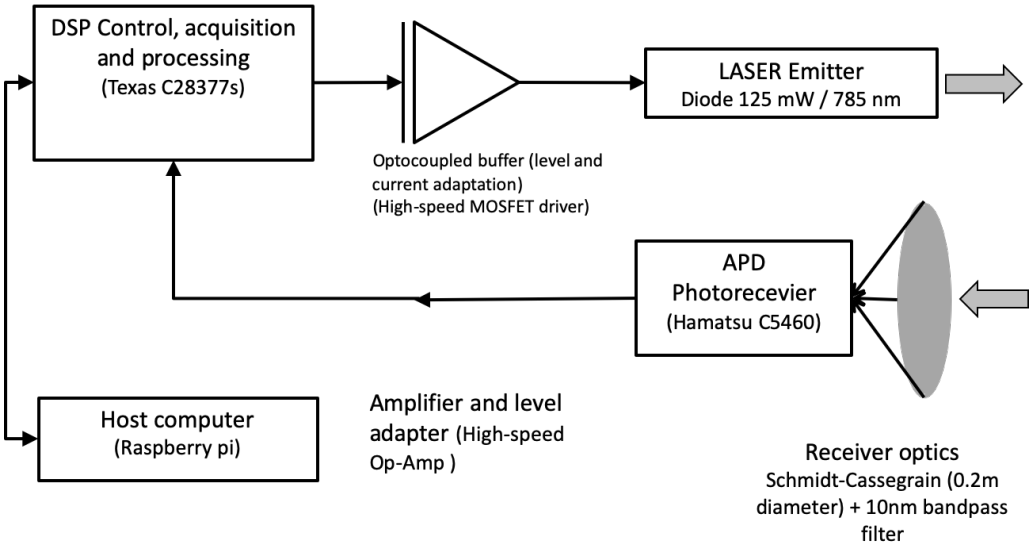


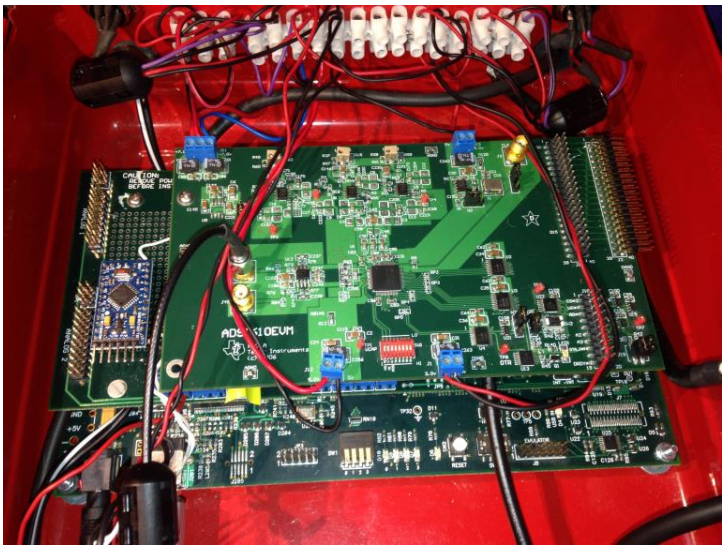
Fig. 6.1. Elements of the second version of the LIDAR.

6.1. DSP AND CONTROL

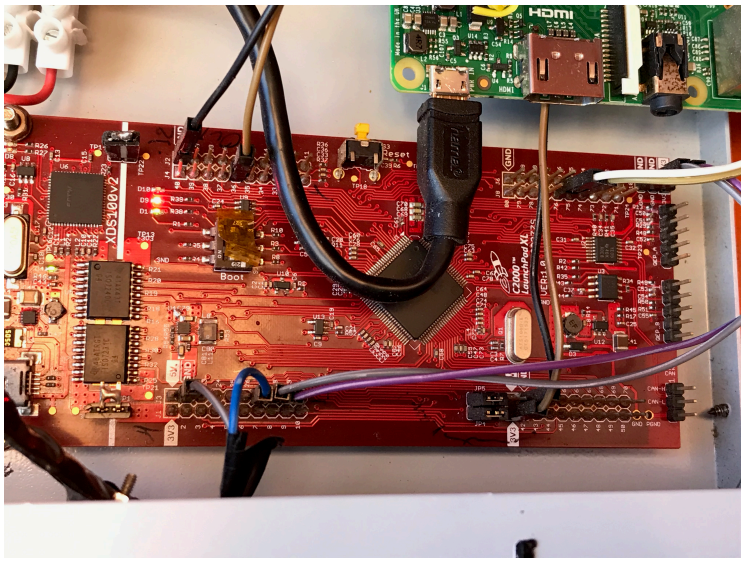
Two versions of the lidar have been developed (fig. 6.2). They are summarized at the end of this chapter. The first version used the DSP Texas Instruments (TI) TMS320C6713 [39] mounted in a DSP Starter Kit (SDK) made by Spectrum Digital [40]. The second version uses the TI DSP Delfino 2837xS in a launchpad configuration [41], [42], a low cost board to develop applications with this processor.

The two versions of the lidar use different computer styles to interface it (fig. 6.3). The first maintained the communication trough a PC that both read the received data and wrote the software and configuration in the DSP. The newest version 2 uses a simple Raspberry Pi single-board computer (SBC) that controls and communicates with the

DSP board sending the commands with the operation to do and receiving the output data and sending it to a server. The software and configuration are loaded with a PC (but this operation is occasional, at the beginning of the system configuration and subsequently only if it is necessary to update it). The final processing in both versions is done in the MATLAB environment.

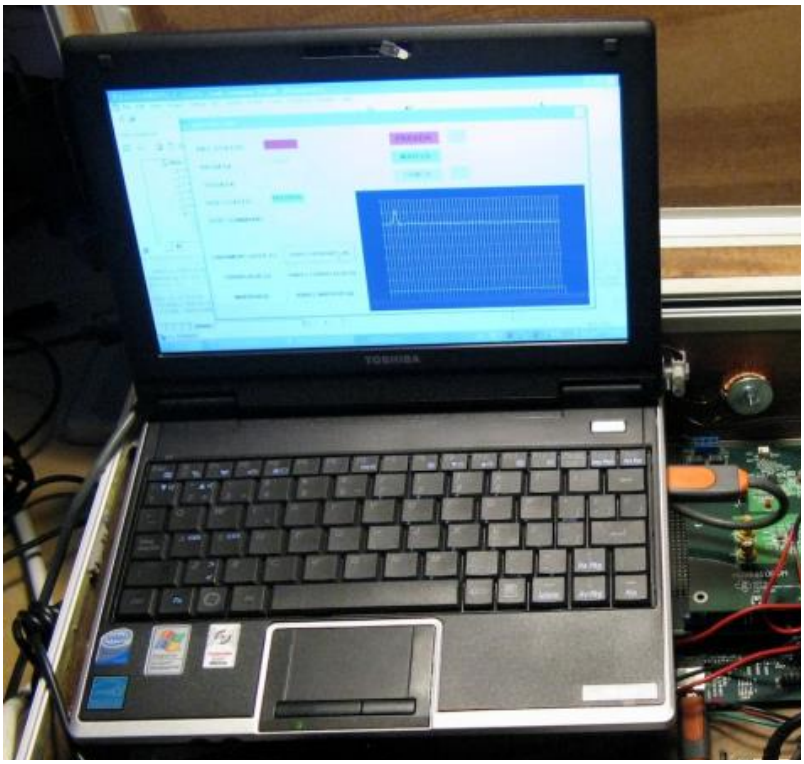


(a)

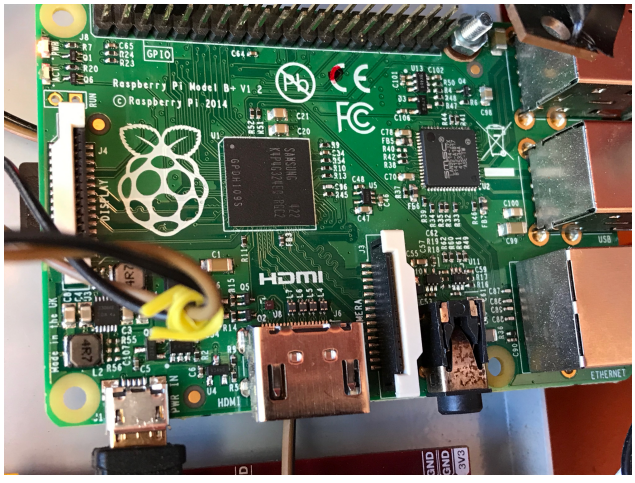


(b)

Fig. 6.2. (a) View of the DSK 6713 board (bottom board) with the A/D converter board on top. Between the two boards there is the 5-6K expansion board. The small board on the left of the image is the Arduino open source board that adjust the gain of the amplifier in the receiver under the control of the DSP processor. (b) The second version of the lidar uses a smaller DSP card and associated electronics. In this card there is the complete system: DSP, A/D converter integrated in the DSP, amplifier, power, etc.



(a)



(b)

Fig. 6.3. (a) The first version with a small PC connected to the USB port of the DSP board to program and control the lidar. (b) The second version uses a Raspberry Pi SBC to control and receive data from the lidar.

6.2. LIDAR EMITTER

The data sequence for modulating the laser is sent from the DSP to the laser using a serial port, through the McBSP —Multichannel Buffered Serial Port— in the C6713 and through the port SPI —Serial Peripheral Interface— in the C2837x, the DSP clock controlling with precision the bit time of this signal.

As it has been explained in the section 0., for practical reasons related to memory savings, in the software codification the sequence is packed in groups of 16 bits. A sequence of 2047 bits, for example, is packed in 128 packets of 16 bits, therefore the last bit is filled with a 0 value.

In the version 1, using the configuration of the serial transmitter, each bit of the frame is 675-ns long (spatial resolution of 101 m). In fact, it is possible to reduce more the bit duration, but for the experiments shown in chapter 7 (Results) this bit time has been chosen. Moreover, the optical module available in the receiver has a 1.5 MHz bandwidth (-3dB), therefore this is a practical limit. Some tests show that the period could be reduced but having a lower response of the photodetector amplifier. In the version 2 of the lidar it has been reduced to 72 m of resolution reducing a bit the response of the photodetector (near 2 MHz).

Laser module

Two different laser module models have been used as it has been explained in section 0. The specifications of both lasers are:

- Laser 1. HL25/MII from Lisa Laser
    - Wavelength: 635 nm
    - Output power: 10 mW
    - Divergence: < 0.4 mrad
    - Max frequency of modulation: 100 MHz
    - Modulation levels: +5V (0), -5V (1)
    - Input modulation current: 100 mA
    - Input impedance: 50 Ohm

- Laser 2. IQH1 from Laser2000
    - Wavelength: 785 nm
    - Output power: 125 mW
    - Divergence: < 1 mrad
    - Max frequency of modulation: 100 MHz
    - Input current: 100 mA
    - Input impedance: 50 Ohm
    - Modulation levels (TTL): +5V (0), 0V (1). TTL modulation.

These laser modules are equipped from manufacturers with the built-in drivers to accept a digital TTL modulation on a 50 Ω input impedance.

The interface with the laser

The output of the DSPs cannot drive directly the laser module input. The modulating circuits in the laser require voltage levels of (+ 5V, -5V) for the red laser module or TTL levels (+5V, 0V) for the IR laser, with a 100 mA current.

While the drivers of the laser modules work with 5V logic levels, the DSP works with 3.3 V logic levels. Level matching requires an interface circuit.

The first version lidar used an interface based on an operational amplifier with a large gain-bandwidth product, the same operational as in the receiver section. The selected OpAmp is the LMH6624 [33]. Some characteristics are:

- Gain Bandwidth product (LMH6624): 1.5GHz
- Input Voltage Noise: 0.92nV/√Hz
- Input Offset Voltage (limit over temp): 700μV range.
- Slew Rate: 350V/μs inverting configurations.
- Slew Rate (AV = 10): 400V/μs
- Supply Voltage Range (dual supply): ±2.5V to ±6V

This kind of operational amplifiers are designed to obtain a great gain-bandwidth product and it is necessary to take care in their use because, unlike classical operational amplifiers, they can oscillate when the gain of the circuit is below a value fixed by design. In the case of the model used in this design, this value is 10 [43], [44]. Specifically, in the design of the interface this is an advantage, because a high gain assures a perfect saturation of the output of the amplifier that is working in on-off mode.

The second version uses a MOS driver that can achieve the same features (as an example, the series MIC4426/4427/4428, which has switching times in the order of 20-40 ns, below the bit pulse length of 400-600 ns). To achieve an electric separation of circuits and avoiding ground loops the design uses a fast optocoupler at the input.

6.3. LIDAR RECEIVER

APD and first stage of the receiver

The optical receiver is based on a Hamamatsu C5460SPL5343 APD module (see table 6.1 for its specifications) coupled directly to the telescope focuser. This module contains the APD itself, the transimpedance amplifier, and a first amplification stage. It integrates the High Voltage source for the APD, and has a single 5 V power line.

The figure 6.4 shows the optical spectral response of this photoreceiver module, with a peak at 800 nm, very near of the IR laser wavelength (785 nm).

Parameter	**Value**
Active Area	3 mm diameter → 7.07×10^{-6} m^2
Spectral response / peak wavelength	400-1000 nm / 800 nm
Photo Sensitivity	0.5 $A.W^{-1}$
Cut-off frequencies	0-1.5 MHz
Photoelectric Sensitivity (Responsivity) at 905 nm	1.02×10^{7} $V.W^{-1}$
NEP at 905 nm	0.12 $pW.Hz^{-1/2}$
Minimum detection limit	0.15 nW r.m.s.
Output impedance	50 Ohm
APD F (excess noise) factor	3.45

Table 6.1. The most relevant specifications of the Hamamatsu C5460SPL5343 module.

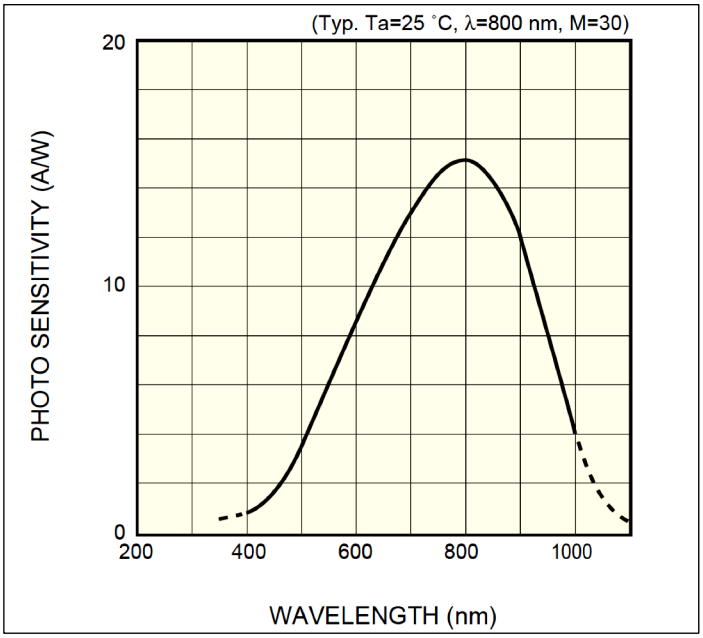


Fig. 6.4.Spectral response of the APD in the photoreceiver module Hamamatsu C5460.

The photoreceiver is mounted with an adapter for the eyepiece holder in the telescope focuser. This adapter contains a 10-nm narrowband interference filter centered at the emission wavelength of the lasers — depending on the laser used the filter is replaced —with a minimum transmittance of 50% for the infrared filter and 45% for the red filter. The red laser is used only for the first tests.

Second amplifier

The APD module output is amplified before reaching the analogue-to-digital converter.

A one-stage amplifier with adjustable gain and a controlled offset input amplifies the signal. The configuration is in the addition mode, that is, one input is the signal received from the APD, and the other input is a DC level to compensate the offset due to the sky background radiation. This DC level is generated by the DSP and an interface circuit.

The design is based on the same LMH6624 operational amplifier described in the section 0 about the laser interface circuit.

Offset compensation

In the first lidar version, the correction of the voltage offset in the amplifier circuit is done by an auxiliary circuit driven by a microcontroller. A DC level applied to the input of the amplifier corrects the voltage offset due to the background sky and other offsets. This correction is made automatically at the beginning of the integration time period. The DSP indicates to the microcontroller, based on the open source hardware platform Arduino (Atmel microcontrollers), to start with this adjustment.

As it has been explained in the section 0, the second version of the lidar does not contain a circuit for the offset adjustment.

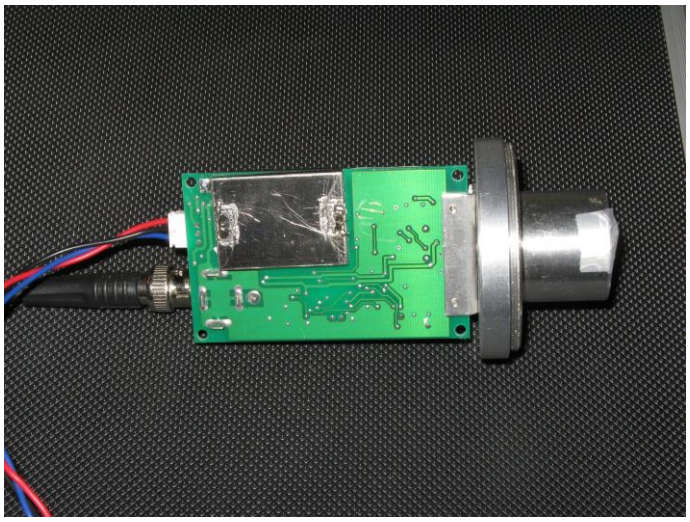


Fig. 6.5. The photoreceiver mounted with the adapter for the eyepiece holder in the telescope .The adapter holds in turn the narrowband interference filter.

Receiver amplifier

The receiver amplifier is one of the most complicated parts of the hardware design because it has to have a high bandwidth to preserve the shape, thus the energy (see section 0) of the pulses and, at the same time has to keep the added noise as low as possible. The first and most critical amplifier, the transimpedance amplifier, that converts the APD photocurrent into a voltage output, is integrated in the photoreceiver module and it has fixed characteristics. The output amplifier of the module, integrated in the same module, has 50-Ω output impedance. This module also provides the high-voltage bias to

the APD. The used photoreceiver has a bandwidth of 1.5 MHz and for this reason the second amplifier, must have at least the same bandwidth, but not much more, to avoid increasing the noise.

The two functions of this amplifier are:

- Adapting the levels at the photoreceiver output to the A/D converter input levels.
- Compensating the offset created by the sky background light at the output of photoreceiver.

The first goal is easy to achieve, but the second one is more complicated. The high-bandwidth op-amp amplifier used —LMH6624— has to be set at a gain higher than 10 because, as it has been explained, for gains below this value the amplifier can become instable. The effective measured voltage noise at the photoreceiver output is around 2-3 mV, which means that the maximum amplitude of the noise is around 6-7 times this value (6σ criteria), that is, 15-20 mV. To adapt this maximum amplitude to the 2.5 V of the A/D converter the gain should be more than 100. As the op-amp used has a gain-bandwidth product about 1.5 GHz it is not a problem to reach this gain.

A/D conversion

In this type of lidar, the emission and the reception of the signals are done at the same time, therefore the A/D conversion is done at the same rate as the emission. Every sample sent implies another sample read from the A/D converter in the period fixed. This means that the DSP has to control these two processes accurately. In the prototypes the Direct Memory Access (DMA) capabilities of the DSPs are used. When a data bit is sent to the emitter amplifier through the DSP serial interface, concurrently a 16-bit data packet is read from the A/D converter. In practice this is difficult to achieve because the two signals run at very high rates, at the limit of DSP interfaces. The software process has to write and read a 12 or 16-bit sample at rates up to 2 or 3 Mbit/s without any time gap occurring. The maximum rate at which the system can run depends on the capability of the DSP.

The first version of lidar with the C6713 DSP uses the ADS1610 A/D [45] [46] converter, with an Evaluation Module connected to the DSP module. This A/D converter can run at speeds as high as 60 Msamples/s but with an important reduction of the signal-to-noise ratio (55 dB) and offering 12 bits of the sample signal. Using this converter at 5 or 10 Msamples/s (default mode) it achieves the maximum signal-to-noise ratio of 91 and 86 dB respectively with 16 bits of resolution at the output. The sent signal runs at 1.5 Mbits/s

which means that this low sampling rate is enough to sample correctly the received signal and to achieve the maximum resolution and signal-to-noise ratio. The second version of the lidar uses the A/D converter integrated in the C2837x running about at 2 Msamples/s with 12 bits.

6.4. POWER SUPPLIES

The system power supplies do not bear special features and do not deserve a detailed description. The only very important thing to have in mind in the designing and mounting of a system like this is the effect of ground loops in the sensitive parts of the system, that is, in the amplification and acquisition section. The presence of ground loops can ruin the operation of whole system. This is a difficult issue to deal with in a complex system like this, with a considerable number of power supplies.

The experience in the design and test of the system tells the importance of keeping further systems as compact as possible. Each cable present in the system is a risk to close a ground loop. The shortest connections probably can decrease the sensitivity of the system to ground loops.

For this reason, the version 2 of the lidar uses a smaller number of power supplies. In the lidar version 1, the DSK DSP6713 uses a 3.3, 5 and 7.5 V to operate, the A/D converter needs another too and the amplifiers, the laser and the APD module need more power supplies. The external PC has its own power supply.

The version 2 reduces drastically the number of power supplies, using one for the DSP, the Raspberry and the MOS driver, another for the APD module, the laser having its own power supply. This reduction greatly reduces the risk of inadvertently leaving ground loops.

6.5. MECHANICAL DESIGN AND ALIGNMENT

Mechanical design requires precision. The support of the telescope is fixed. The mechanical fixation of the laser has to maintain laser beam centered in the field of view of the telescope and at the same time to allow the movement for fine adjustment.

As the field of view is 13' (4 mrad) or 25' (7 mrad), depending of the lidar version, the laser beam has to be positioned with precision in the order of 1'. The divergence of the laser beam is less than 1 mrad. It is possible to reach these requirements using good mechanical elements as the used in optical banks. Figure 6.6 shows the final mounted mechanical system.

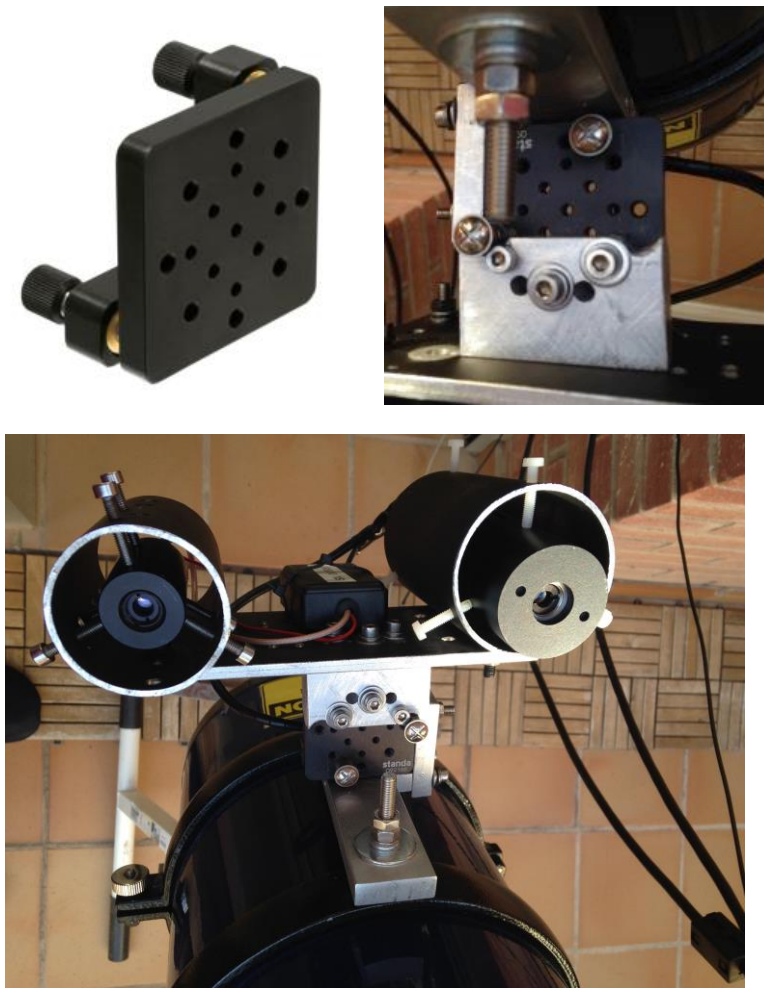


Fig. 6.6. Mechanical precision support for aligning the lasers inside the receiver field of view.

Fig. 6.7.a shows the laser beam in the field of the telescope captured by a CCD sensor in the focus of the telescope with an exposure of 8 seconds.

Fig. 6.7.b. is a superposition of the field of view of the photodetector and that of the telescope over the image of the laser beam obtained in fig. 6.7.a.

Fig. 6.7.c. shows the laser beam when clouds are present: a brighter spot is visualized at the end of the beam because of the reflected light in the clouds layer. As the spot in the cloud is wider than the diameter of the laser beam, probably it shows light backscattered from directions other than those the beam illuminates (multiple scattering).

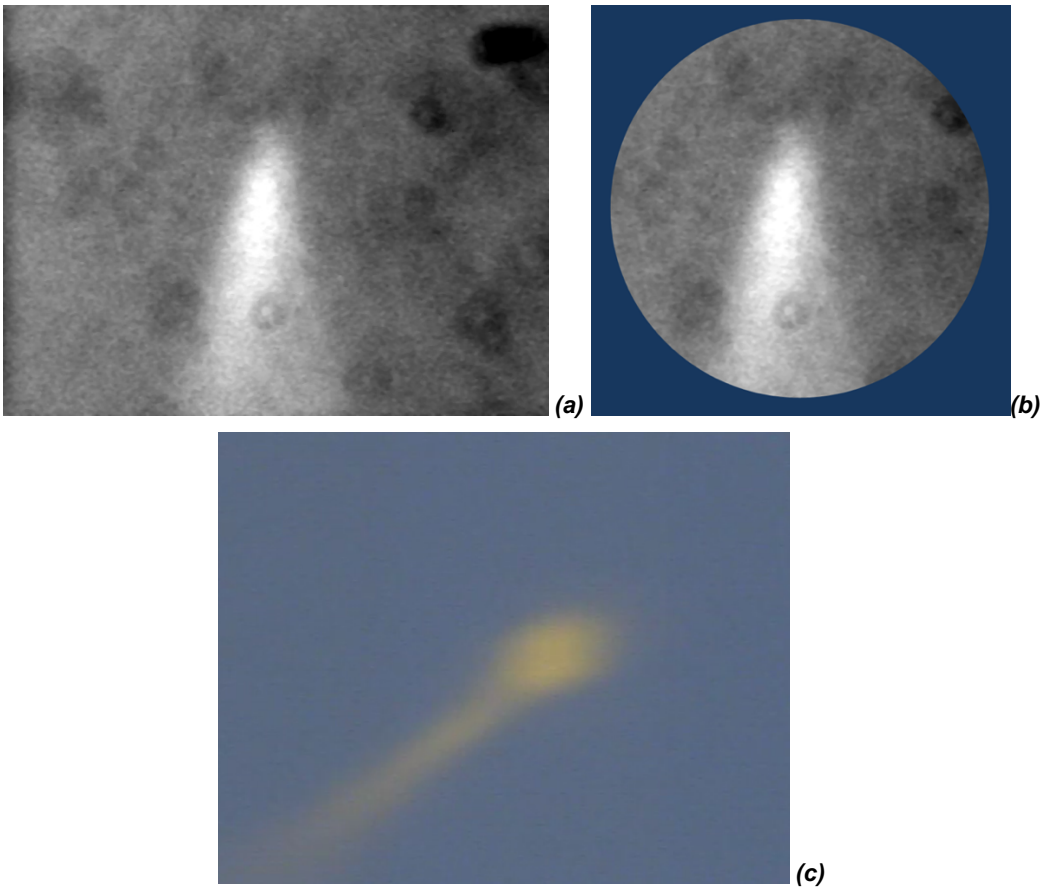


Fig. 6.7. (a) A picture of the laser beam in a clear night sky using a CCD camera. (b) The circle shows the field of view of the APD sensor at the same scale as the CCD camera used to align the laser beam (at right). (c) In a cloudy night the end of the laser beam is visible as a brilliant spot.

6.6. OPTICAL SYSTEM AND OVERLAP ESTIMATION

The experiments with this lidar have employed two telescopes. One of them is a Newton configuration with 0.15 m diameter and 0.75 m focal length. The other one is a modified Schmidt-Cassegrain, resulting in a telescope with 0.2 m diameter and 0.4 m focal length. The overlap function is shown in figure 6.8. The overlap function shows the fraction of the laser beam that is inside the field of view of the telescope as a function of the distance to it.

Assuming that the laser beam has a Gaussian profile, the fraction of power inside the field of view as a function of the range R can be calculated as [47]

$$ovf\left(R\right)=\frac{1}{\sigma^2\left(R\right)}\int_0^{r_c(R)+r_D} A_i(R,r)\exp\left(-\frac{r^2+s^2\left(R\right)}{2\sigma^2\left(R\right)}\right)I_0\left(\frac{r\,s\left(R\right)}{\sigma^2\left(R\right)}\right)r\,dr\,, \tag{6.1}$$

where

$$if\ r<\left|r_D-r_c(R)\right|\begin{cases}\alpha=\beta=\pi\\ Ai(R,r)=\dfrac{1}{\pi r_c^2(R)} \quad if\ r_c\le r_D\\ Ai(R,r)=\dfrac{1}{\pi r_D^2(R)} \quad if\ r_c>r_D\end{cases}$$

$$if\ r>\left|r_D-r_c(R)\right|\begin{cases}\alpha=\arccos\left(\dfrac{r_D^2+r^2-r_c^2(R)}{2r_Dr}\right)\\ \beta=\arccos\left(\dfrac{r_c^2(R)+r^2-r_D^2}{2r_c(R)r}\right)\\ Ai(R,r)=\dfrac{1}{2}\left[r_D^2(2\alpha(R,r)-\sin(2\alpha(R,r)))+r_c^2(R)(2\beta(R,r)-\sin(2\beta(R,r)))\right]\end{cases} \tag{6.2}$$

r_D is the radius of the field stop (which, in the optical arrangement implemented, coincides with the photodiode radius), $r_c(R)=fD/2R$ is the radius of the least confusion circle, $I_0\left(x\right)$ is the modified Bessel function of the first kind and order 0. $s\left(R\right)$ is the distance between the laser beam axis and the receiver line of sight, given by

$$s(R)=\frac{f}{R}\left[\left(R\phi_{\parallel}+d_0\right)^2+\left(R\phi_{\perp}\right)^2\right]^{1/2}, \tag{6.3}$$

with f the telescope focal length, d_0 the distance between the beam and telescope axes at $R=0$, and $\phi_{\parallel}$, $\phi_{\perp}$ the components of the misalignment angles, respectively, parallel and perpendicular to the plane defined by the telescope axis and the laser axis at $R=0$. In the case that these angles are 0 the value $s(R)=f\,d_0/R$. $\sigma(R)$ is the radius of the image on the telescope focal plane of the laser beam at $1/e^2$ of its axis intensity at range R, given by

$$\sigma(R)=\frac{f}{R}w(R). \tag{6.4}$$

$w(R)$ is the laser beam radius at range R at $1/e^2$ of its axis intensity, assuming a Gaussian beam with far-field divergence θ with its waist at $R=0$:

$$w(R)=\sqrt{\left(\frac{\lambda}{\pi\theta}\right)^2+\left(R\theta\right)^2}, \tag{6.5}$$

with λ the laser wavelength.

Figure 6.8 illustrates the parameters to calculate the overlap factor. The field of view of the telescope is the envelope of all the cones with the vertex centered in each point of the surface of the telescope.

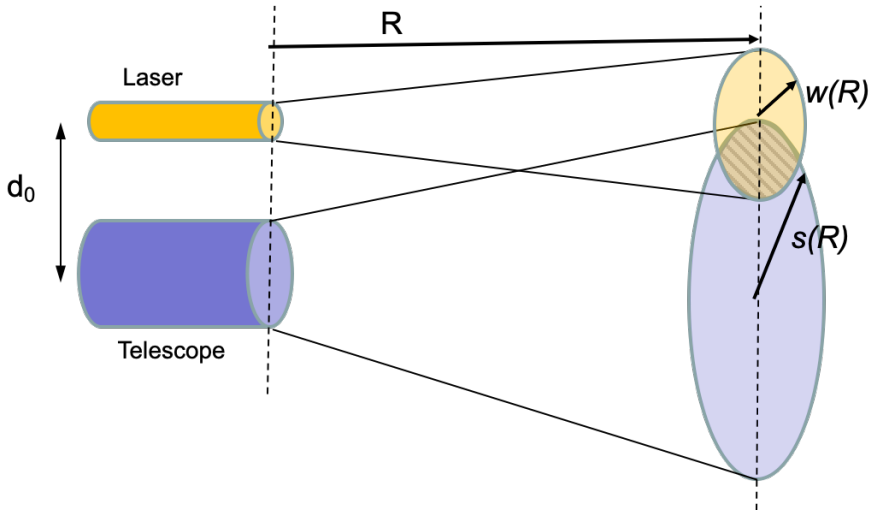
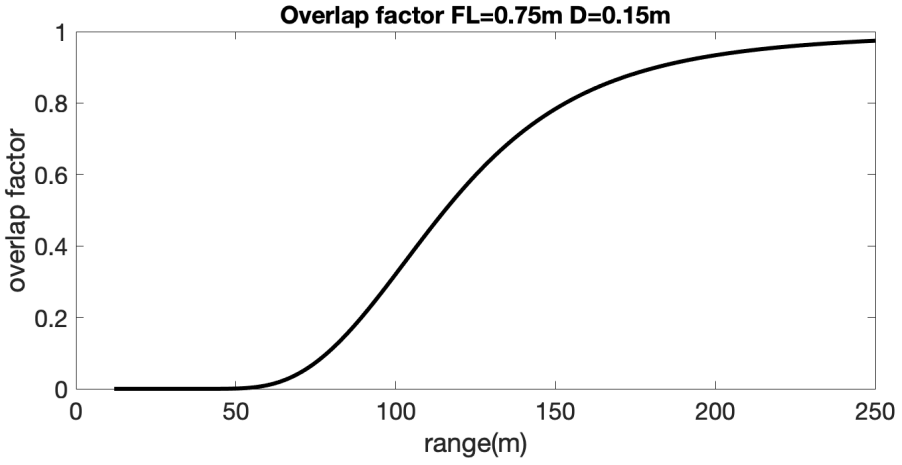
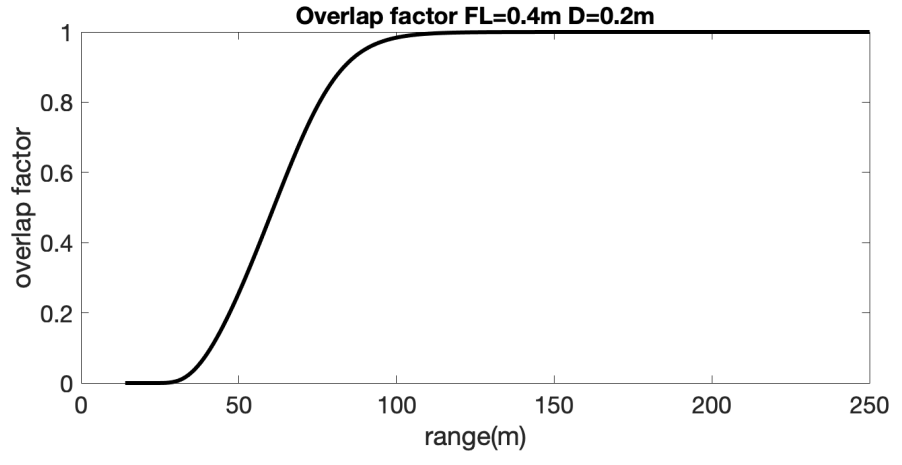


Fig. 6.8. The overlap function represents the fraction of laser beam power that is inside the telescope field of view at a certain distance. The dashed area is the overlapped surface of the two cones of the laser beam and the telescope field of view (Eq.* (6.1)*).

Figures 6.9a and 6.9b show the calculated overlap functions respectively for the 0.75-m focal length telescope and for the 0.4-m focal length. In both cases laser-beam divergence is taken as 1 mrad, the photodetector diameter as 3 mm, and the laser beam axis and the receiver line of sight are considered parallel. In the same figures it is possible to observe that the shorter is the telescope focal distance the earlier the full overlap arrives before.



(a)



(b)

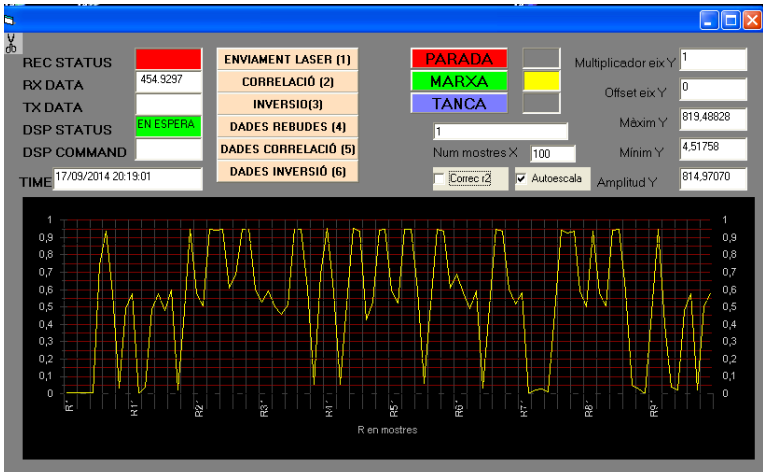
Fig. 6.9. Overlap functions of the two telescopes used. (a) for a telescope with a diameter of 15 cm and focal length of 0.75 m. (b) the plot of the overlap function for a telescope with a diameter of 20 cm and a focal length of 0.4 m. The laser beam divergence is 1 mrad.

Therefore, one important adjustment is to align the laser beam inside the telescope field of view. The appendix A.2 explains a procedure applied to perform the alignment.

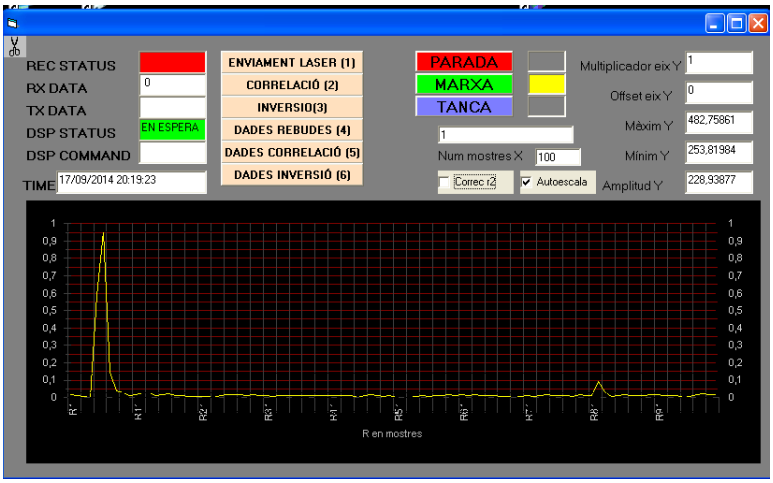
6.7. PROCESSING AND DISPLAYING THE DATA

In the demonstrated system, the display of the processed data is done using an external computer. Also, a computer is necessary to program the DSP trough the CCS interface in the Texas Instruments DSP's family that has been used in the implementation of this lidar.

Complementary software has been developed to communicate with the DSP, as well as and to sending commands and receiving measurements from the DSP, and saving it in the host computer. An example of the signal at the input of the correlator and of the correlator output is given in figure 6.10.



(a)

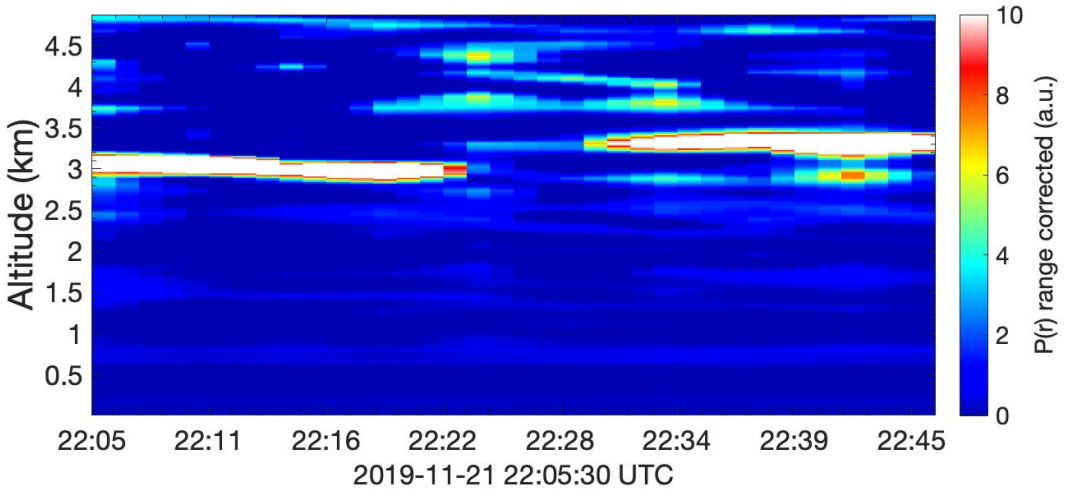


(b)

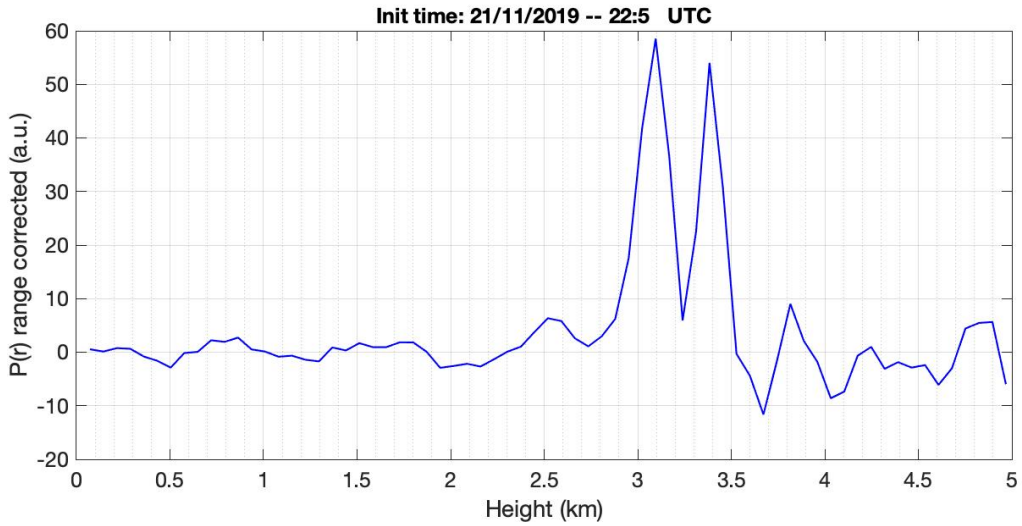
Fig. 6.10. (a) Received signal at the correlator input . (b) Correlator outputs for the input signal of pane (a) with small shift from the origin caused by a guard time at the beginning.

The test target for the results in fig. 6.10 was a hard one (a wall) located very near (some meters) so that the reflected signal was strong enough to be well above the system noise. This basic test shows that the whole system is operating according to the theory.

When the system is in operation, the results of the correlation are saved in plain text files to be processed and analyzed. In this work the saved files are analyzed with MATLAB and the results obtained are presented in 1-D and 2-D graphics. An example is given in figure 6.11 where 6.11a is a 2-D map to observe the temporal evolution of the targets (clouds in this case) in the atmosphere with a color code representing the received power (in arbitrary units). Fig. 6.11b shows the 1-D graph of the signal integrated during a period of time, in this case the complete hour of the previous 2-D graphic. In the figures the reviewed power is range corrected.



(a)



(b)

Fig. 6.11. Measurement results presented in MATLAB graphics. (a) 2-D representation of data with colors representing the range corrected received power (in arbitrary units). (b) 1-D 1-hour average of the signal in panel (a).

6.8. SUMMARY OF THE TWO LIDAR VERSIONS

In this chapter, two versions of the lidar system have been presented. A summary follows of the different characteristics of the two versions.

PROTOTYPE 1

The first prototype is made around a powerful DSP and the complementary hardware. Basically, the most prominent characteristics are:

- DSP C6713 with 1 MB RAM
- A/D converter until 60 Ms/s with at 12-bit amplitude resolution or 10 Ms/s at 16-bit amplitude resolution.
- Bit duration: 675 ns (100 m resolution)
- Hamamatsu APD-based photoreceiver module.
- Amplifier between the photoreceiver and the input of A/D converter.

- A PC connected thought USB connection to program and reading the DSP.
- IR laser of 125 mW at 785 nm.
- Telescope with 150-mm aperture and 750-mm focal length.
- Narrowband filter of 10 nm bandwidth centered on the laser wavelength.
- Hardware needs a number of different power supplies because of the A/D converter card.
- Power consumption around 50W.

PROTOTYPE 2

The second prototype uses a very cheap DSP, the Delfino C28377S with some peripherals integrated in the same chip. The most remarkable characteristics are:

- DSP C28377S with 256 kB RAM
- A/D converter integrated with 12-bit resolution at 12 Ms/s or 4 Ms/s and 16-bit resolution. This eliminates the A/D converter card and greatly simplifies the power supply design.
- Bit duration: 480 ns (72 m resolution).
- Hamamatsu APD-based photoreceiver module.
- No amplifier between the photoreceiver and the input of A/D converter.
- SBC Raspberry Pi used to communicate with the DSP, sending orders and receiving data.
- Low power consumption of the system because of low-power components (DSP and Raspberry).
- IR Laser of 125 mW at 785 nm.
- Telescope with 200-mm aperture and 400-mm focal length.
- Narrowband filter of 10 nm bandwidth centered on the laser wavelength.
- Low size and low weight of the electronics
- Power consumption below 10 W.

The first version has best features than the second, because it uses an A/D converter with more bits of resolution and the possibility of running at higher samples rates. The processor is faster and the amount of base memory available is larger. But it is more

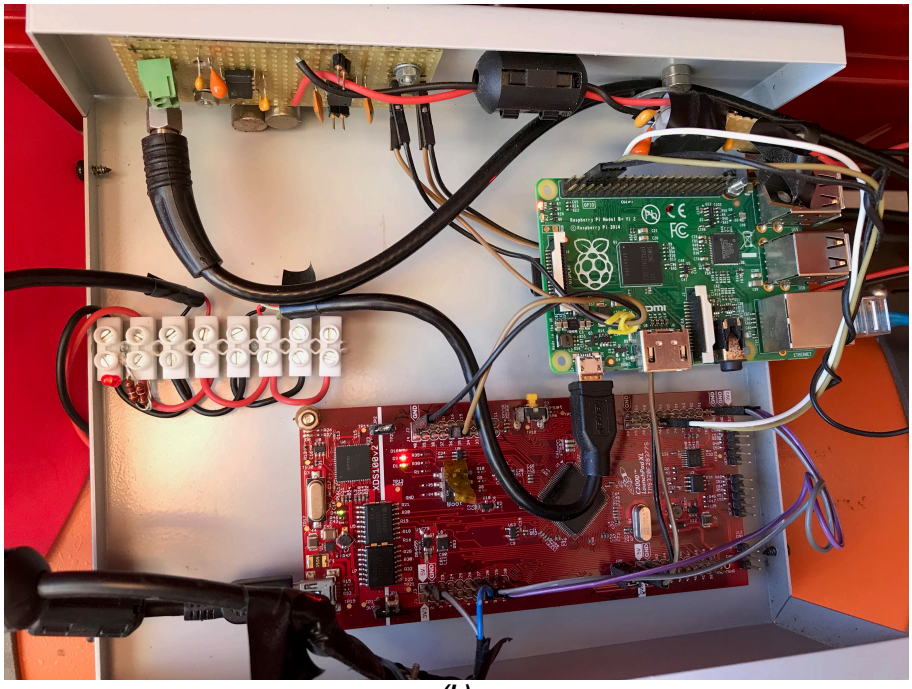
expensive (at the moment of the development in the order of 20 or 30 times) than the second because of the hardware used, it needs diverse power supplies at different voltages and demands more power than the second.

The second prototype is cheaper, lighter and much more energy-efficient than the first one, because of the DSP selected. Figure 6.12 shows a quick comparison of the size of the prototypes. The reduction in size and complexity in the second prototype is evident. The figures do not include the laser module, the APD photoreceiver and the telescope, which are the same in both prototypes. This new families of DSPs are been developed after the first prototype was made and the reduction of some features can be assumed to test the development of the first prototype in a second more compact and efficient than the one first, despite the increasing difficulty of trying put the code in less memory space and at less processor speed.

Figure 6.13 shows the complete prototype of lidar version 2 with the different parts and the system size. The lidar is completely enclosed in a box with a transparent window, allowing all-weather operation. The size could be further optimized, but this is not the goal of this research work.

In the next chapter 7 of results, it is shown results obtained with booth prototypes, evaluating the behavior of both and observing how the second version continues showing good results despite the reduction of some features.

(a)



(b)

Fig. 6.12. The hardware around the DSPs of the two prototypes at approximately the same scale. (a) First prototype showing the A/D converter over the DSP card, the input and output amplifiers, part of the power supply and the control computer. (b) Second prototype showing the three main processing and control elements: the DSP card, the MOS driver for the laser module and the SBC Raspberry Pi to control the DSP.

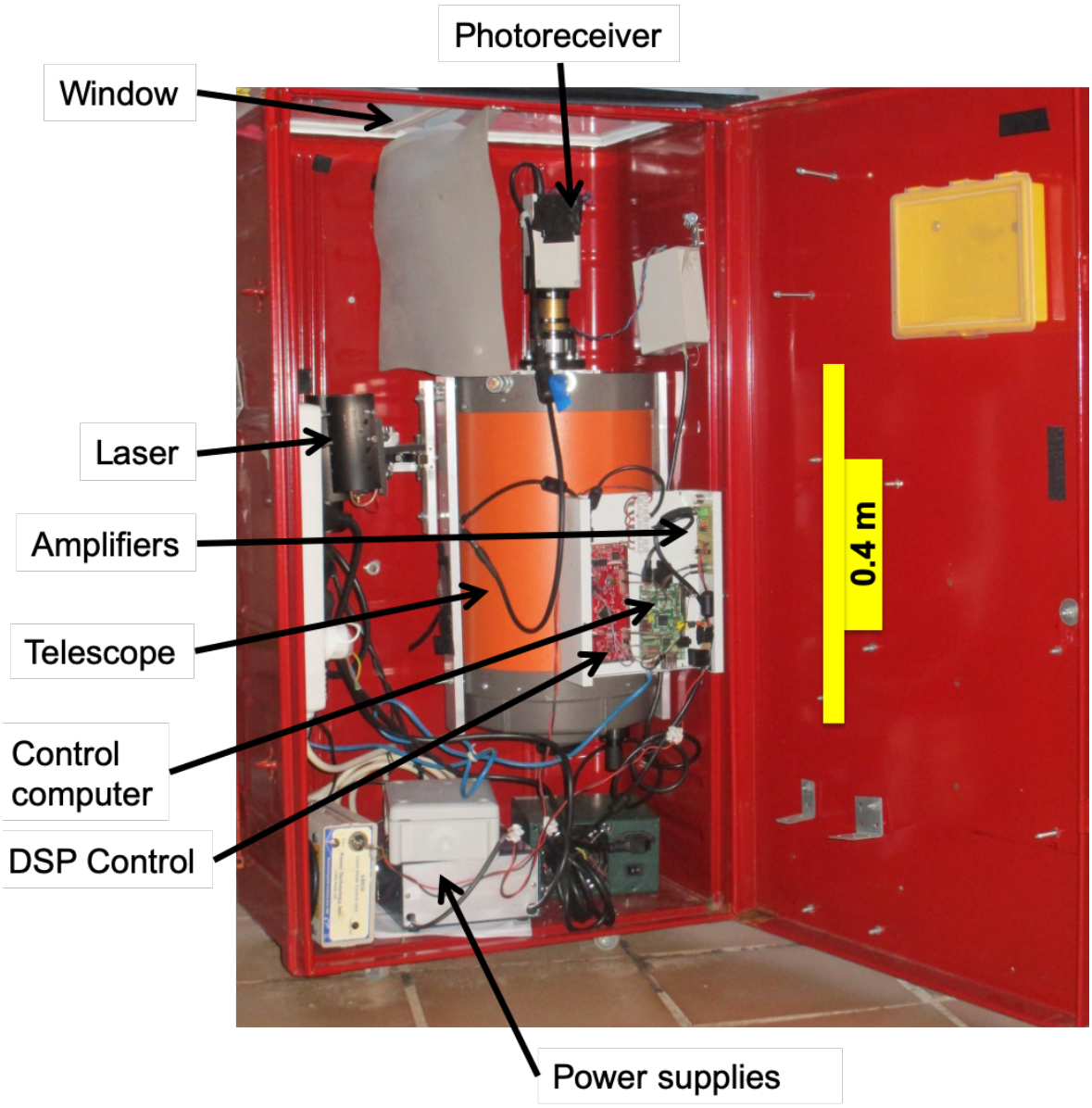


Fig. 6.13. The version 2 of the lidar prototype inside a robust box to remain outdoor.

7. TESTS AND RESULTS

The lidar designed and built according to the principles discussed in the previous chapters has been tested in two different environments: in the laboratory, where some variables are under control, and in the real open field, with continuously changing conditions. A few selected results, considered as more significant, are shown and discussed.

7.1. LABORATORY TESTS

Laboratory test are carried out to verify the agreement between the operation of the system and the theoretical principles. The test of the complete system in the laboratory is done with the 10 mW red laser. The power is enough to perform different tests and the light is visible, which makes the test process easier.

Test 1: send and receive sequences

This test is done to check the software operation and the correct synchronization of the A/D converter and the CPU.

A special sequence is designed to validate the process of sending, receiving and adding sequences. With pseudo-random sequences as those discussed in Chapter 0, it is very difficult to perform a fast test of alignment of the sequences that are sent and received, because they have a too complicated structure to inspect easily. Instead, the sequence to carry out the test is created as

0000000001000000001000000011000000011100000001111…. ,

that is, an increasing number of 1's filled with 0's to have sets of 10 bits. This number is not important in itself and the critical issue is to control the misalignment. With a sequence like this, or another one, the system is tested by sampling the sequence several times and it is easy to check the correct alignment by representing the sequence or calculating the difference between the sampled sequences.

Other tests were done on the electronic circuits, like the amplifiers, to check their correct operation. The section of the receiver amplifier dedicated to correct the sky background

(offset) in the receiver is tested too in combination with the software in the DSP that has to carry out this control. Figure 7.1a shows the adjust ramp for the offset control. Figure 7.1b compares the emitted sequence and the received sequence. The signal received reflected at different distances is shown from fig. 7.1c to fig. 7.1e with a clear wall as a solid target at different distances (2, 5, 10m approx.).

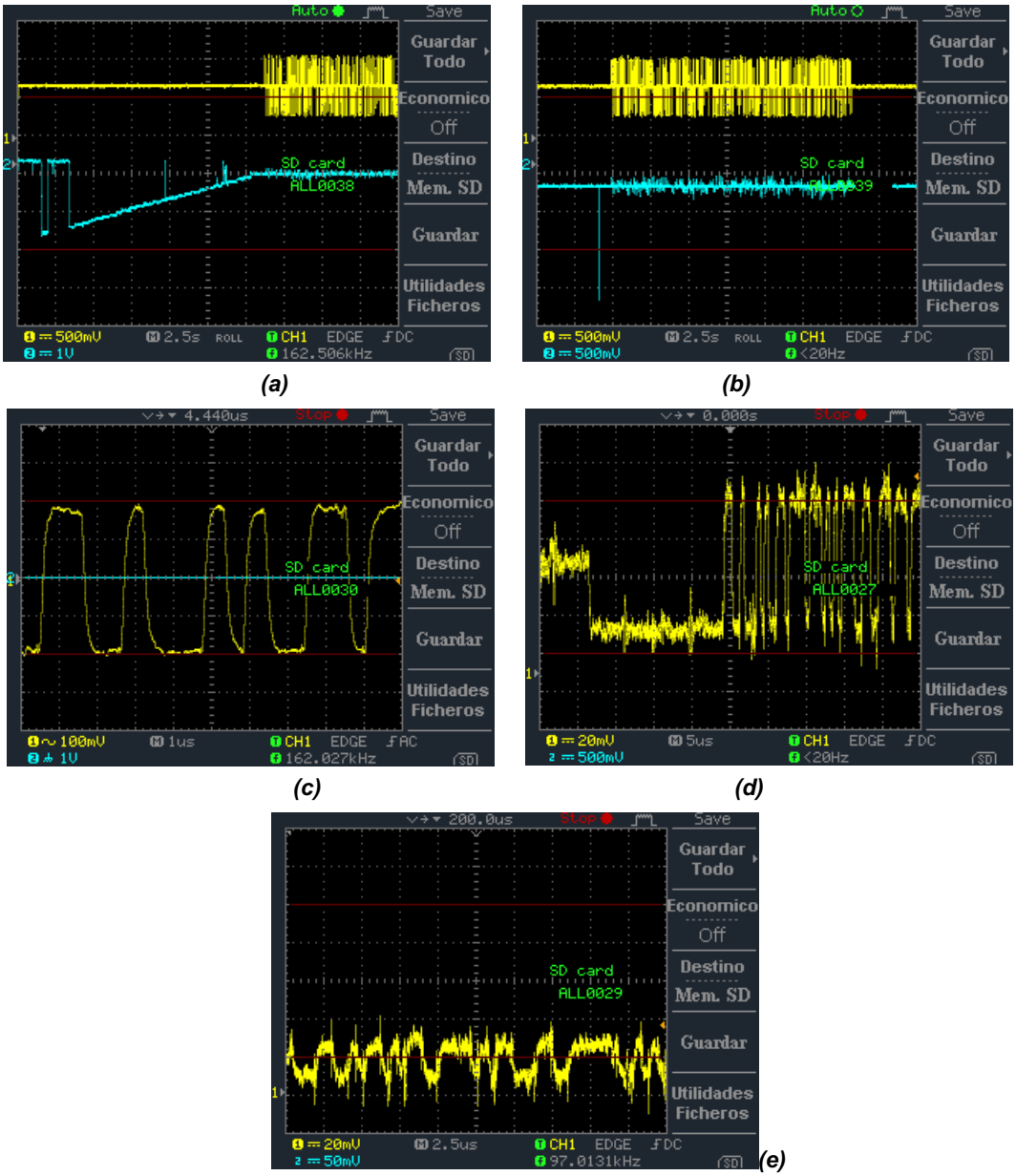


Fig. 7.1. Oscilloscope screenshots of the system operation. (a) Ramp to adjust the receiver offset immediately before the operation starts. Channel 1 (yellow): received signal; channel 2: voltage ramp for adjusting the offset. (b) A complete sequence transmitted (channel 1) and received (channel 2). (c) Received signal reflected on a very near and reflective target (clear wall at around 2m). (d) The same as (c) on a farer target (5 m approx.). When the transmission starts, a guard time is applied by means of 32 bits fixed to 0. (e) Received signal from a farther target (10 m approx.).

Test 2: Verification of the correlation process

This test is aimed to check that the result of correlation with the pseudo-random sequence is correct. For that, a target at a distance of 0.4 m was used. The graph on fig. 7.2. shows the result of correlating the received signal with the transmitted sequence. The peak at the center of the correlation function, which corresponds to a 0 samples delay, is perfectly clear. Only two hundred of samples are calculated — delay of -100 to 100 samples— with a simple correlation (no circular correlation); this is the reason for the small ripple at the ends of the correlation function.

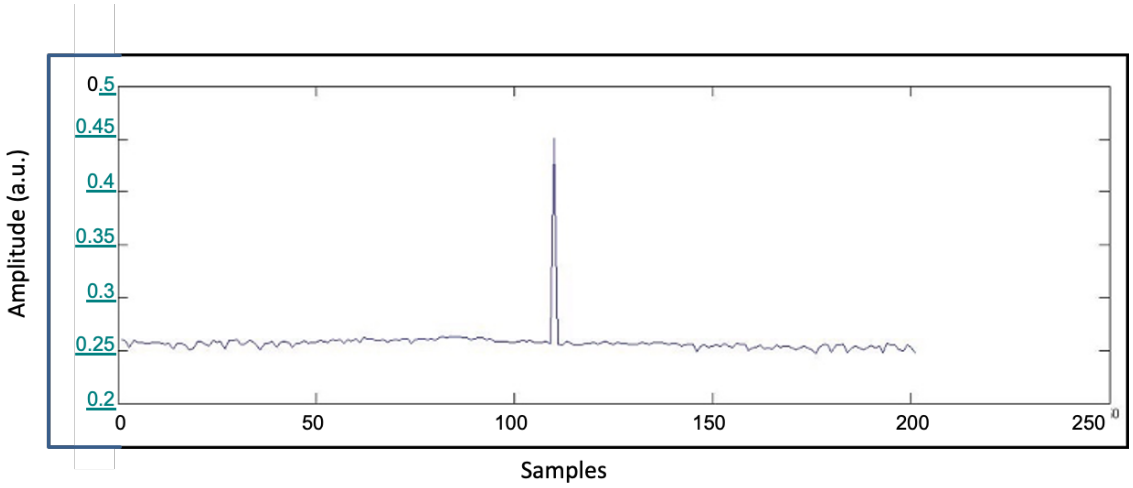


Fig. 7.2. Correlation of the received sequence reflected in a very near target (r=0.4m).

Test 3: Delay measurement

A test under controlled conditions was done to check the system ability to measure signal delays. To implement this test the laser output was coupled to a 100-m optical fiber used to create a delay of 500 ns. The applied pseudo-random sequence had 1023 bits and the parameters of the software in the DSP were adjusted to have a time resolution below the 500 ns delay. The experiment consisted in dividing the signal and sending it through two paths. One of them was the direct path, using a very near (0.4m) target. The other was the path through the optical fiber (100m). The outputs of both paths are combined at the input of the photodetector (fig. 7.3).

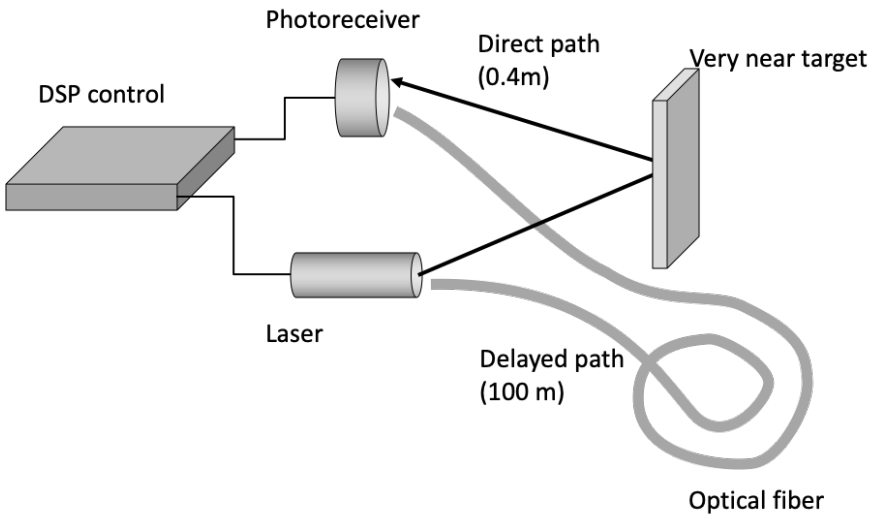


Fig. 7.3. Experimental setup in the laboratory.

The signal is acquired by the DSP and processed (correlated with the transmitted sequence) using the process designed for the lidar. The result is shown in fig. 7.4, where a temporal offset of 3 samples is visible until the first echo, the signal with delay=0 (target at 0.4 m) that appears at sample number 4. The delayed signal through the optical fiber appears at sample number 6. The bit time is adjusted to 250 ns, which taking into account the speed of light in the optical fiber (around 200.000 km/s), corresponds to a 50-m resolution. Therefore, the expected delay is two samples, as the figure shows.

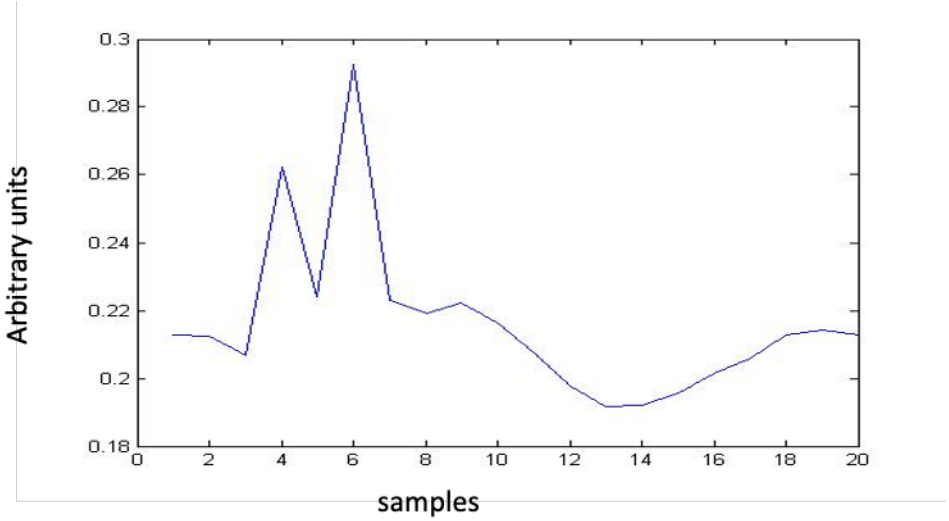
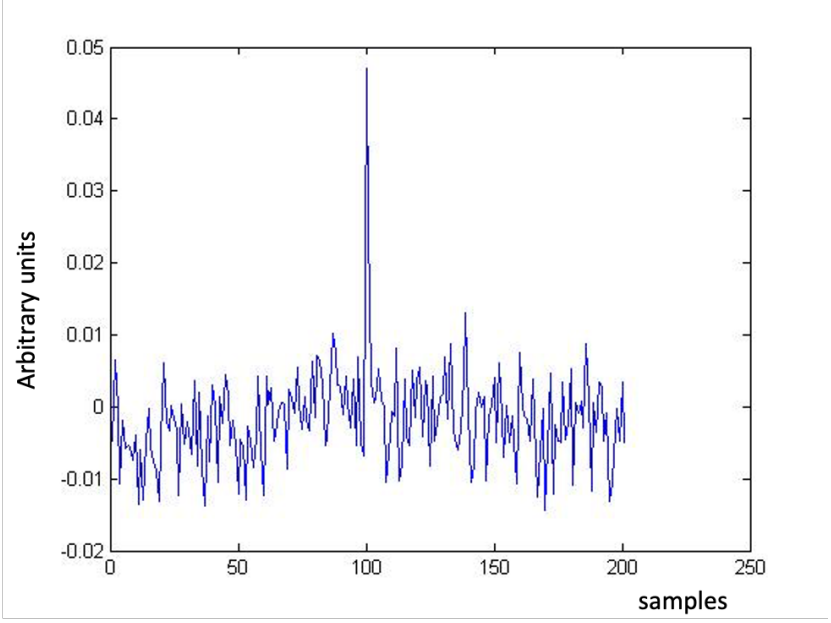


Fig. 7.4. Correlation of the received sequence composed of a direct signal and a signal delayed by a long optical fiber (100 m) to create a measurable delay. The bit time was adjusted around 250 ns (50m of resolution). After the samples showing the delay it appears the damped oscillation of the electronics impulse response.

Test 4: Complete operation

Finally, a test of the complete operation of the system, with the total system (DSP, laser emitter, photoreceiver) running is presented. Figure 7.5 shows the correlation function of the received and integrated signal —50 sequences— with the sent modified M sequence as it has been explained in chapter 3, with a wall as solid target at 6m from the lidar to obtain an intense signal return, as these experiments are performed with no optical system in the receiver. Each sequence has a duration of 545 μs (1023 bits length, 533 ns bit time), and the total integration time is 27 ms.

The system runs correctly and the peak of correlation appears at sample number 101, because the delay time in a distance of 6 m is below the range resolution (80m). Two hundred samples have been calculated, from a delay of -100 to a delay of +100 samples.



(a)

Fig. 7.5. Test of the whole system. Processing of the received signal. 50 sequences are transmitted, received and integrated, using a wall as a target at distances of r=6m. Each sequence has a duration of 545 μs(1023 bits length, 533 ns bit time) and a total 27 ms of integration time.

7.2. FIELD TEST WITH SOLID TARGET DETECTION USING THE 10 mW RED LASER AND LIDAR VERSION 1

A first field experiment to verify the operation of the system in the range of distances of some hundreds of meters to 2-3 km, consisting in the detection of solid targets, was done employing the 10 mW red laser. The measured distances in this test are given in Table 7.1 and the graphical results of the detected echoes are in 7.7 . The lidar parameters were as follows:

- Laser CW 10 mW, 635 nm
- Telescope 20-cm aperture, 2-m focal length
- APD photoreceiver module (Hamamatsu C5460SPL5343) with a 3-mm diameter photodetector
- Temporal resolution= 533 ns
- Spatial resolution = 80 m
- M-sequence duration= 1023 bits or 545 μs
- Integration: 1000 sequences= 0.545 s
- Integration 10000 sequences =5.45 s
- Lidar installed in the UPC Campus Nord (see fig. 7.6)

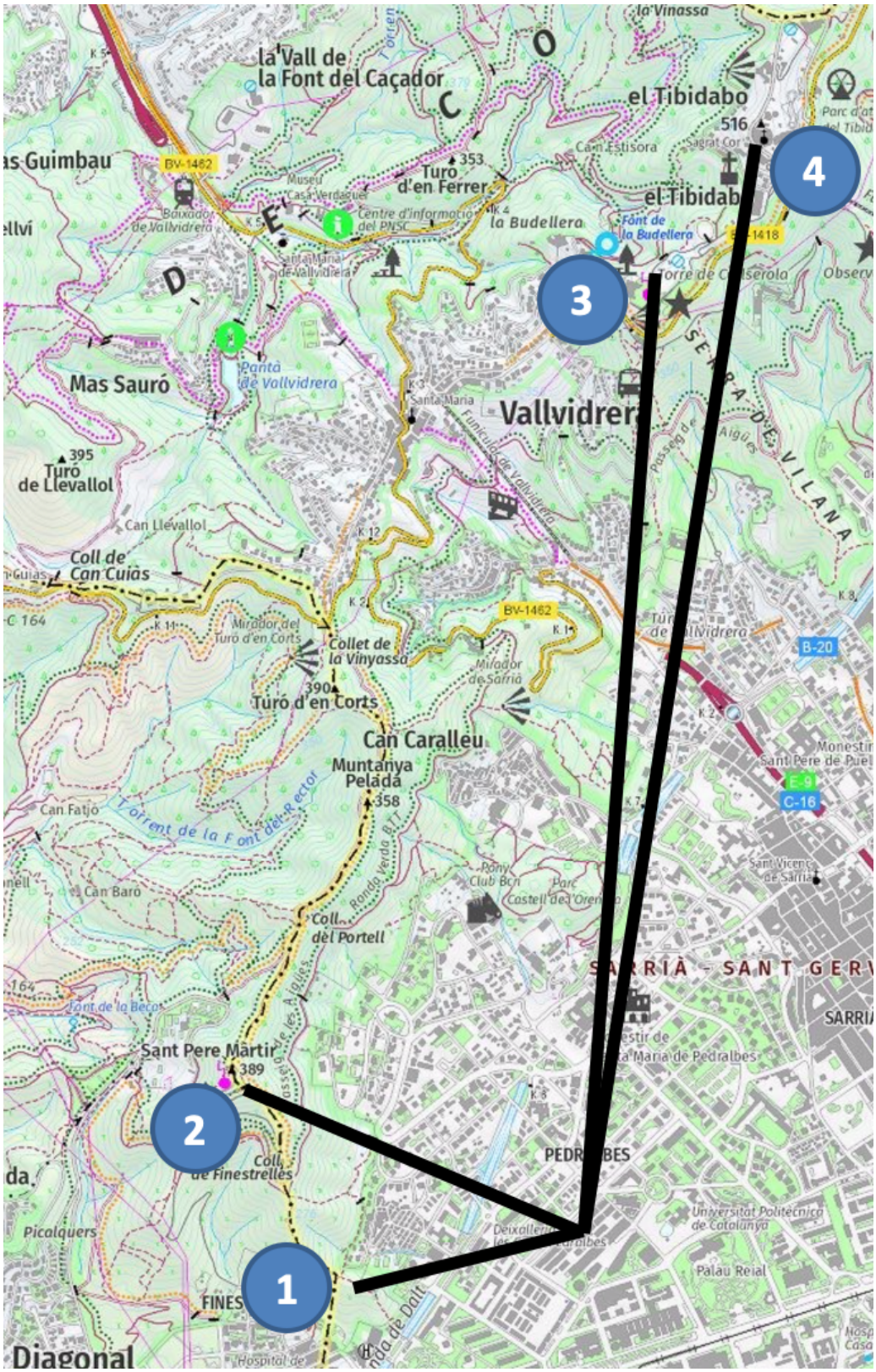
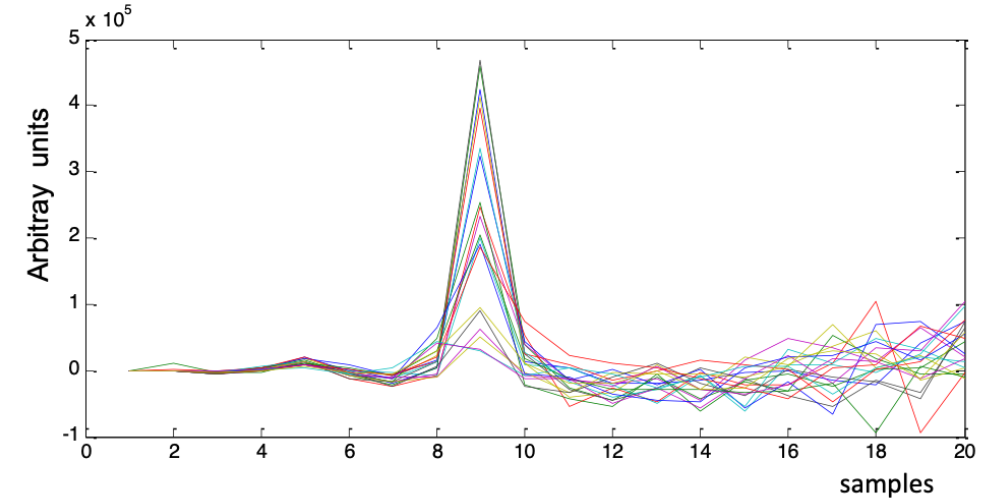


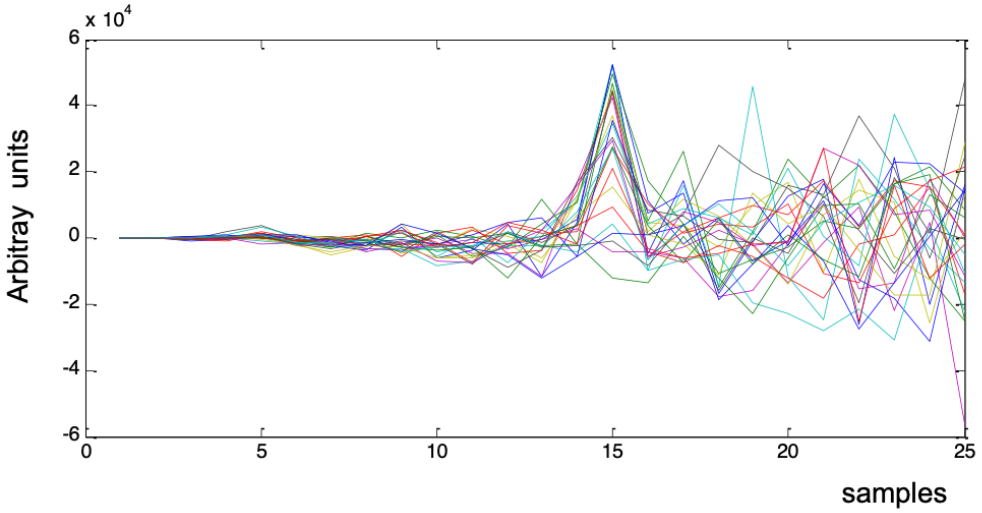
Fig. 7.6. Map of the target points.

Point on the map	Target	Distance measured on a map	Distance measured (lidar)	Figure showing the result
1	*Water tank on the hill*	0.725 km	0.72 km	7.7a
2	*Antenna tower on top of Sant Pere Màrtir hill*	1.2 km	1.2 km	7.7b
3	*Collserola Telecommunications tower*	3.1 km	3.24 km	7.7c
4	*Top of Collserola hill (Tibidabo)*	3.7 km	3.84 km	7.7d

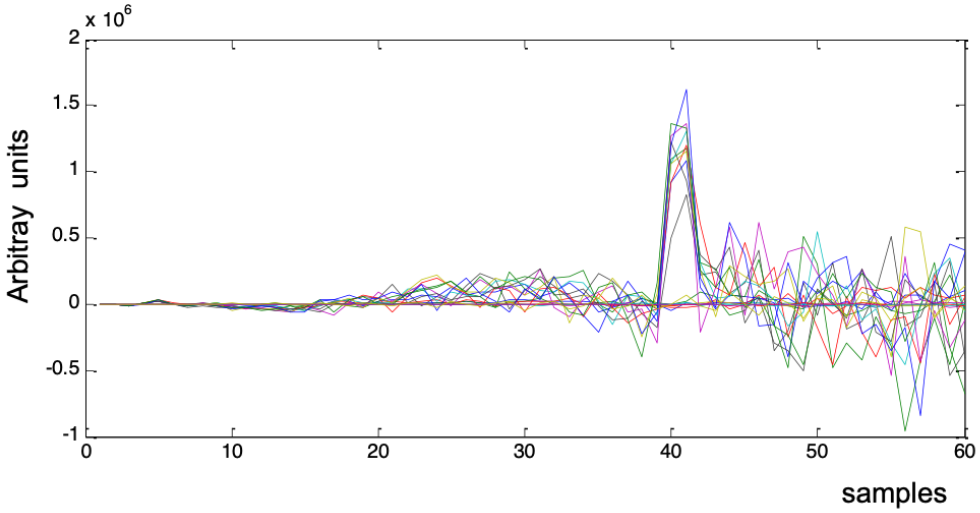
Table 7.1. Summary of ranging results obtained in the real field test.



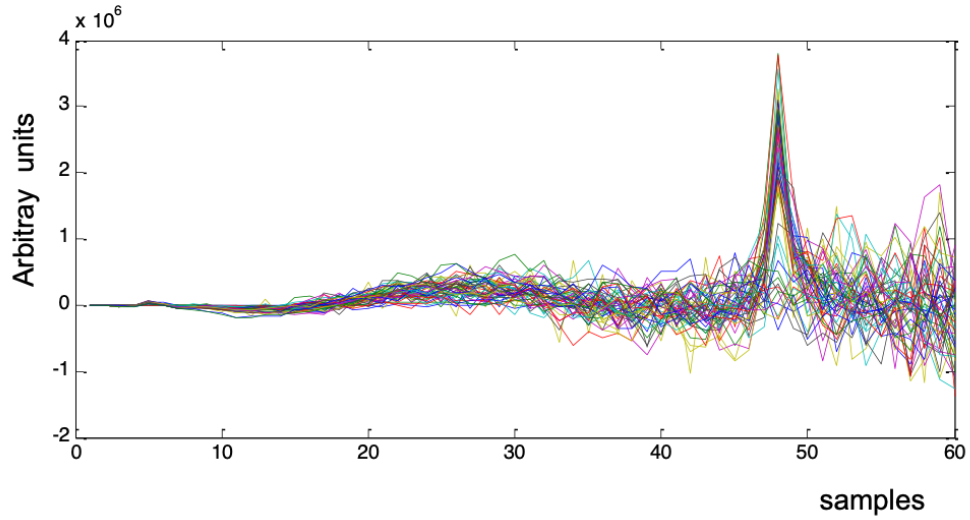
(a)



(b)



(c)



(d)

Fig. 7.7. Detection of different targets in the experiment summarized on table 1. The resolution of samples is 80m and the X-axis is in number of samples of the measured delay. Y-axis is measured in arbitrary units range corrected. (a) Water tank, (b) Antenna tower on Sant Pere Màrtir hill. (c) Collserola Tower. (d) Tibidabo hill.

7.3. FIELD TESTS WITH CLOUD AND RAIN DETECTION USING THE 10 mW RED LASER AND LIDAR VERSION 1

After the success of field tests detecting solid targets, a more complicated experiment was done seeking the detection of clouds. The experiments were performed accumulating 10^5 sequences, which is equivalent to 55s integration time with the specification of 545 μs per sequence. Low clouds, at 400-600 m altitude were detected in the sky during a part of the night as shown in fig. 7.8. Another interesting experiment was the detection of rain, in fig. 7.9. It is interesting to compare the lidar data with the registered rain detected by a near weather station. The coincidence of precipitation (fig. 7.9 bottom panel) with the detection of important echoes in the lidar (fig. 7.9 top panel) is significant.

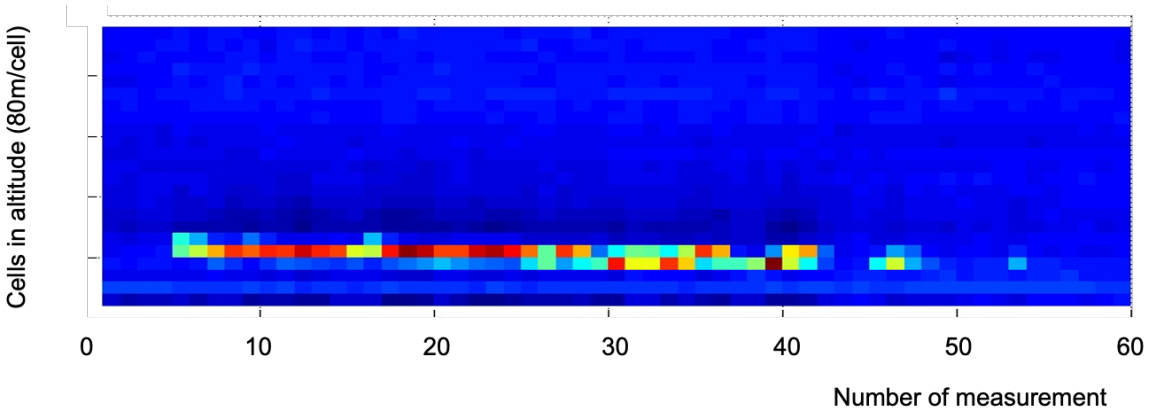


Fig. 7.8. First detection of clouds using the resolution of 80 m. According to this measurement, clouds were at 400 m height approximately. Y-axis: altitude in number of samples, X-axis time in number of measurement (55 s per measurement). Color scale: range-corrected power in arbitrary units.

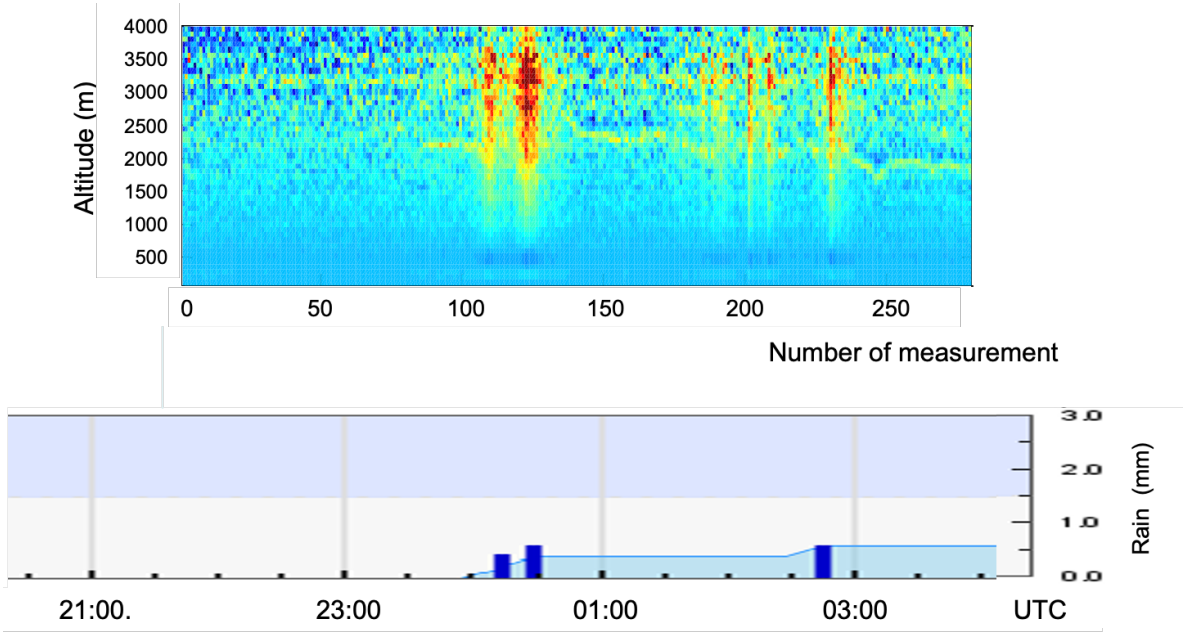


Fig. 7.9. Rain detection with the lidar. The intense echoes coincide with the precipitation registered by a near weather station. Top: Y-axis: altitude in meters, X-axis: time in the number of measurement (55 s of integration time but the interval of measurement is 1m 20s approx.), color scale: range corrected power in arbitrary units. Bottom: registered rain along the time. The two graphics are temporally aligned. M-sequence of 1023 bits used.

7.4. FIELD TESTS WITH CLOUD DETECTION USING THE 125 mW IR LASER AND LIDAR VERSION 1

The 125-mW IR laser improves the system performance as it is shown in the following results of different tests. With this laser, clouds are easier to observe. Not taking in account other changes due to the wavelength variation — figure 6.4 shows that the APD is more sensitive at 785 nm than at 635 nm and this contributes to increase the range at IR wavelengths — an increase of 3.5 times in range is expected. The lidar parameters are as follows:

- Laser CW IR of 125 mW, 785 nm
- Telescope: 15-cm diameter, 0.75-m focal length
- APD photoreceiver module, Hamamatsu C5460SPL5343, 3-mm diameter APD

- Temporal resolution= 655 ns
- Spatial resolution = 100 m
- Sequence duration= 1023 bits or 670 μs
- Integration: 10000 sequences= 6.70 s
- Integration of 60 s = 90000 sequences approximately.
- Lidar installed in Terrassa

More details about the specifications of laser and APD module are in the sections 6.4.3 and 6.5.1.

Cloud and aerosols detection

Clouds are detected in different situations as shown in the next figures. In the figure 7.10 the lidar shows clouds at different altitudes until 4 km clearly. The full overlap of the laser with the field of view of the receiver can be noticed in figure 7.11, around 300 m, below which the signal is very faint. The same figure shows a low aerosol layer or an optically thin low cloud (very light fog perhaps) causing the bright signal echo at 200-300 m height.

Clouds have been detected until altitudes of 8-10 km, as shown in figure 7.12. At higher altitude the detection is uncertain.

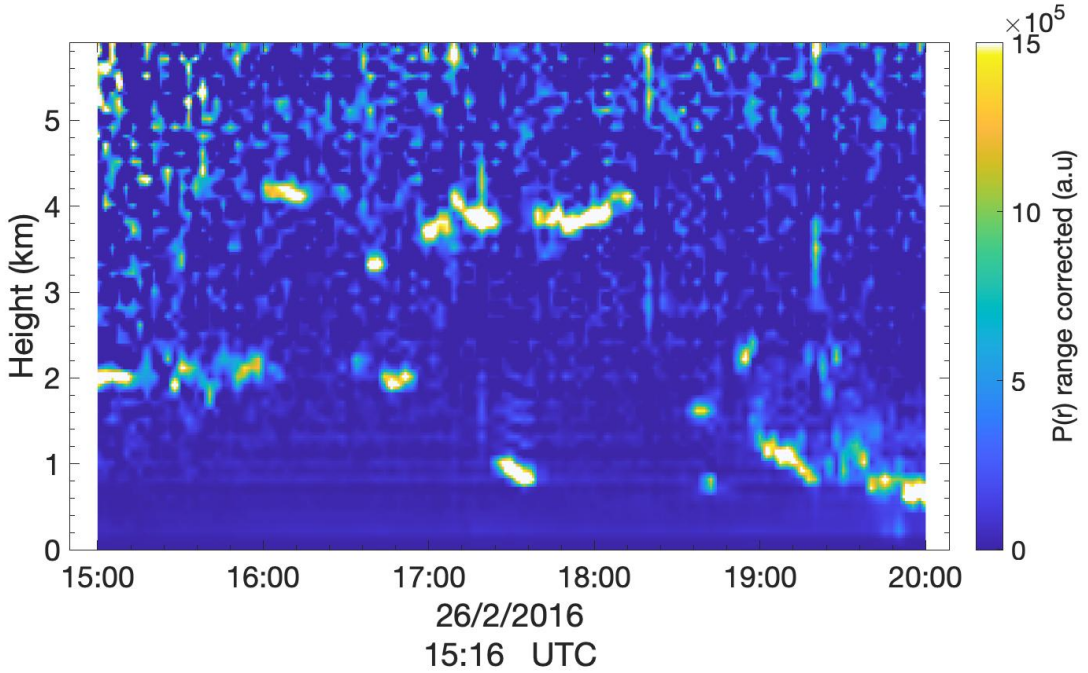


Fig. 7.10. 2-D graphic showing detected clouds at different altitudes. Layers of clouds, detected from 500 m to 4000 m. Height resolution of 100 m, temporal resolution of 4 min.

There is a ripple appearing behind a bright target producing an important echo, as figure 7.11 shows. This effect must be further investigated, but is probably due to the impulse response of some part of the electronics, like amplifiers (likely the amplifiers in the photoreceiver module). It appears only after near and bright targets and it has a pattern easily recognizable as a damped oscillation. The same figure shows a very clear layer of clouds at 5 km and there is not visible the ripple after it.

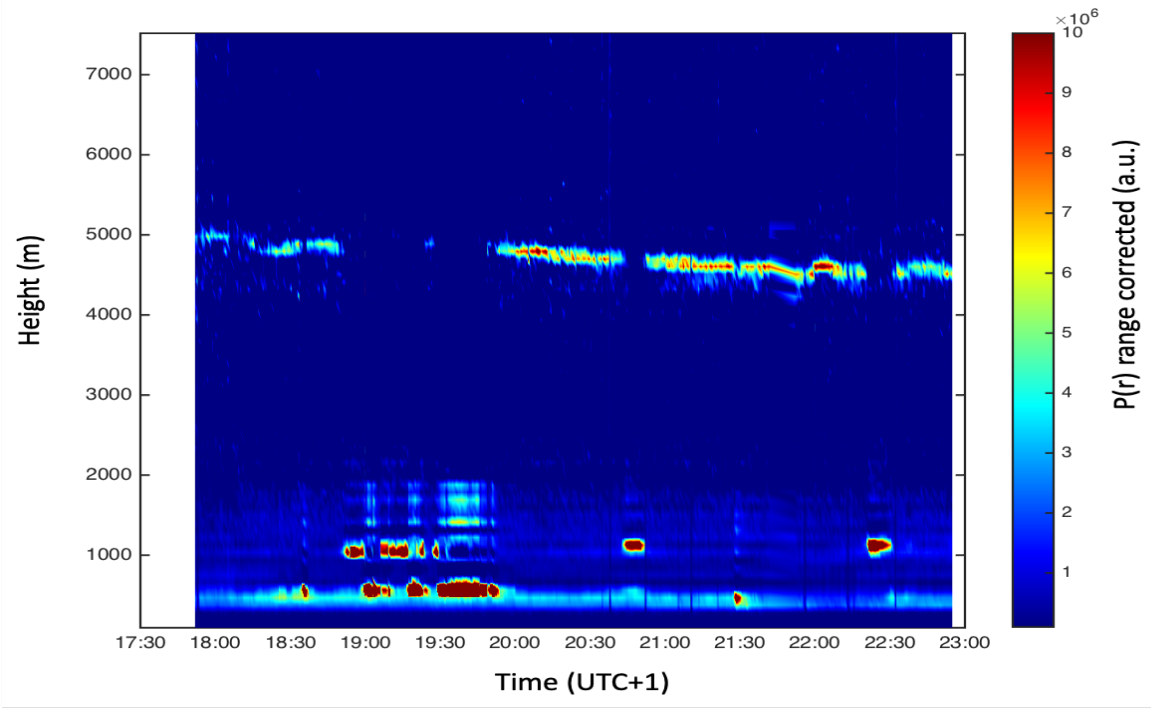
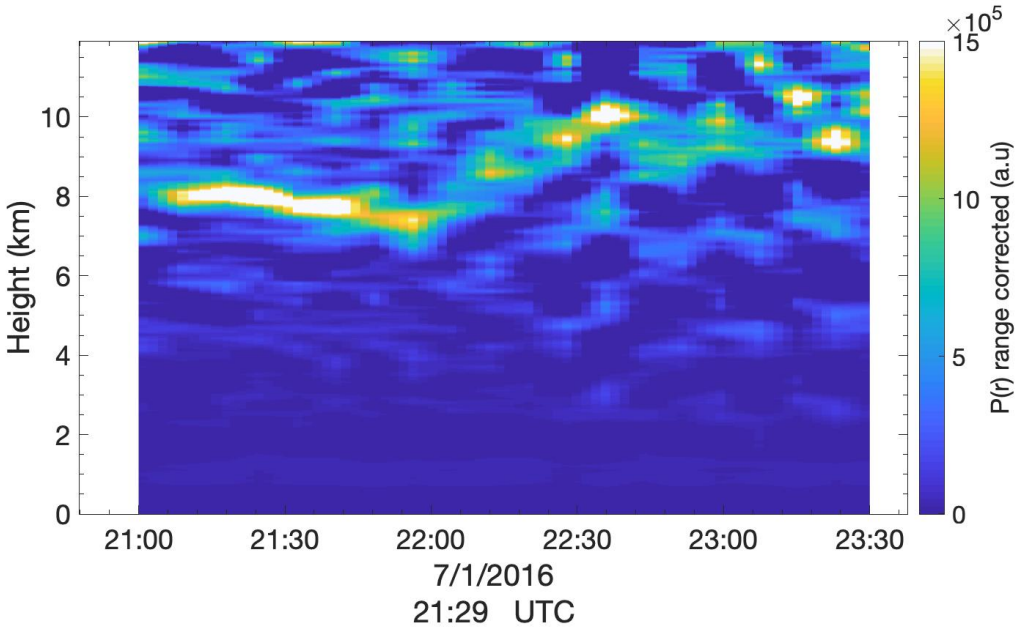
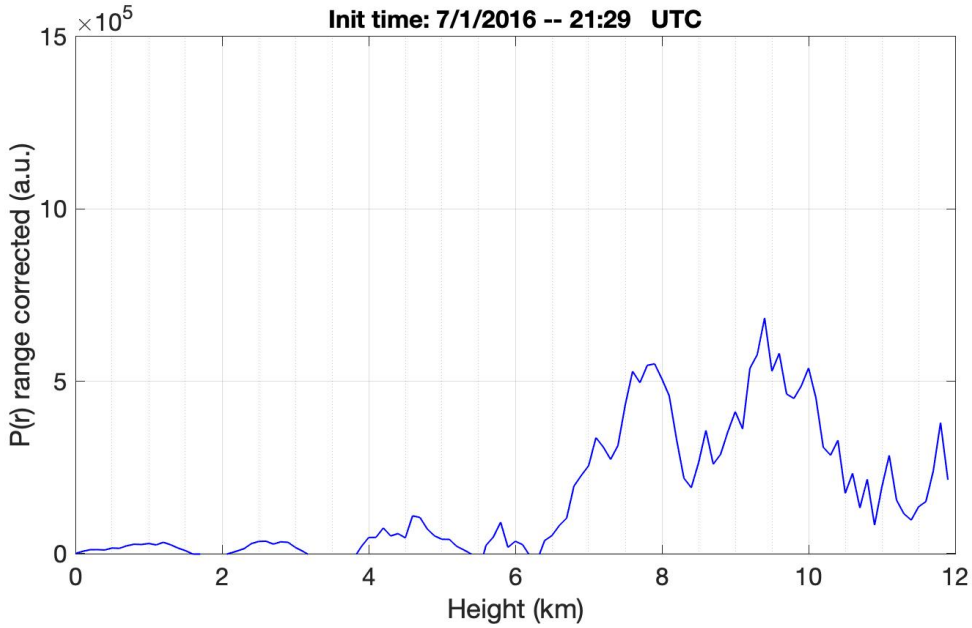


Fig. 7.11. 2-D graphic of clouds detected at 5000m mixed with some low clouds. The peak below 500 m is the full overlap attainment. Above these low clouds the damped-oscillation effect appears and the cloud layer at 4.5-5 km is invisible, masked by the low clouds.



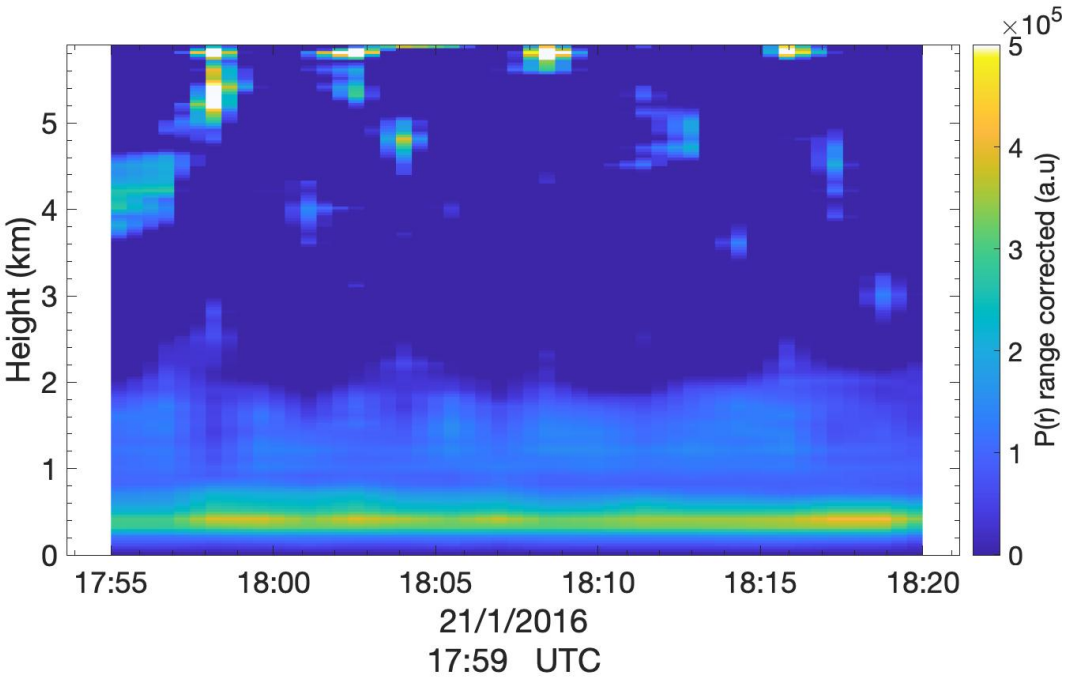
(a)



(b)

Fig. 7.12. 1-D graphic of a clear detection of clouds detected at high altitude (8-10 km). Integration time of 4 minutes per measurement. Sequence A1 of 2048 bits. (a) Evolution of clouds. (b)Total integrated measurements (2h 30m).

Light fog or aerosol layer can be detected as figure 7.13 shows (sequence A1, 2048 bits length, with 2 minutes of integration time). In this case they are clearly visible at the end of daytime and beginning nighttime. During daytime the signal-to-noise ratio decreases because of the noise induced by the sky background radiation.



(a)

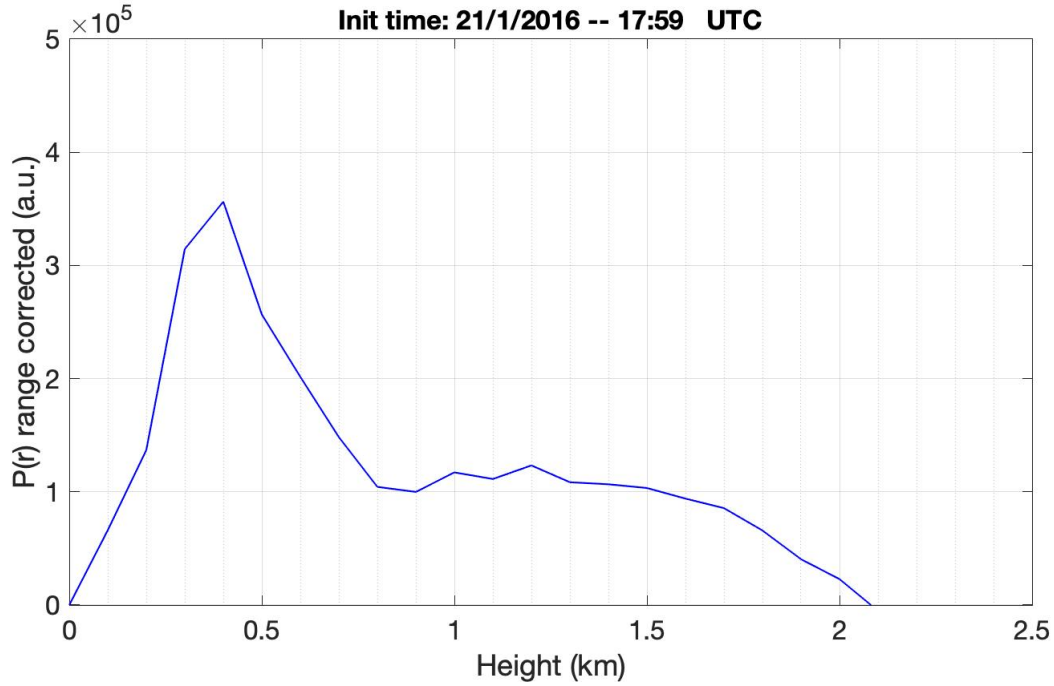


Fig. 7.13. Aerosol with A1 sequence of 2048 bits. Range resolution of 100 m, 2 minutes of integration per measure. (a) 2-D graphic. (b) Integrated data (25 minutes).

Rain detection

In some cases, rain has been detected as shown in the example of figure 7.14. At the beginning of the night there was a cloud layer around 2.5 km whose altitude decrease with time and eventually precipitation starts When the rain starts the clouds are masked by the rain. The data were obtained with an integration time of 15 s and a resolution of 100m (sequence A1, 2048 bits length).

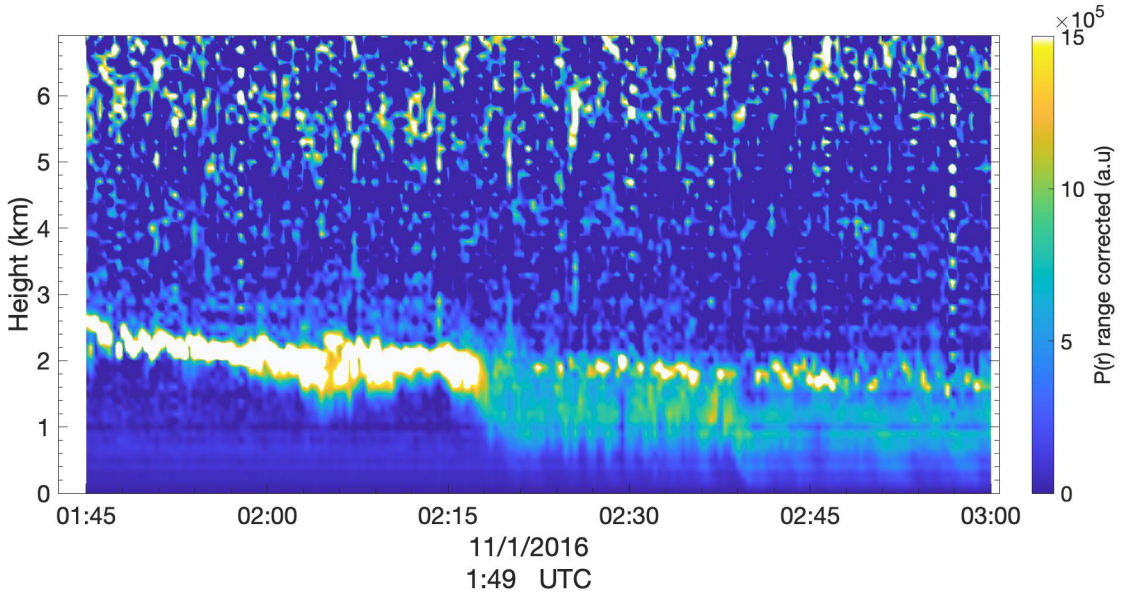


Fig. 7.14. Rain detected after the presence of a low cloud layer. Sequence A1, 2048 bits length, range resolution of 100 m, integration time of 15 s for each measure.

Wildfire smoke detection

During the period of system tests in the 2015 summer a wildfire broke out some tens of km away from the lidar observation point. Figure 7.15 shows the detection of a smoke layer. The bright spot is a sporadic cloud at low altitude at 18h 15m approx. The smoke came from a fire in Ódena, 30 km away from the lidar observation (Terrassa) on 26/07/2015. A layer between 1200-1600m of altitude shows the smoke at 1.5 km approx. The brightest point is a small cloud at low altitude intercepting the line of sight for some minutes. Range resolution of 100m. M sequence of 2047 bits. Integration time of 2 minutes per measurement.

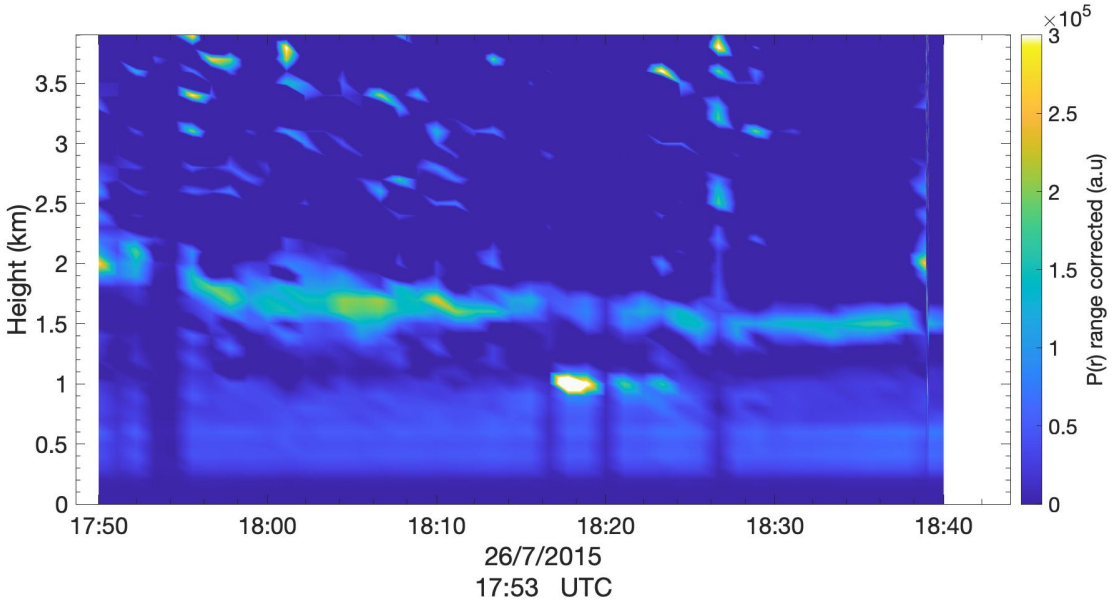


Fig. 7.15. Smoke layer (at 1.5 km approx.) coming from a fire in Ódena, 30 km from the lidar observation (Terrassa) on 26/07/2015. Range resolution of 100m. M sequence of 2047 bits. Integration time of 2 minutes per measurement.

7.5. FIELD TESTS WITH CLOUD DETECTION USING THE 125 mW IR LASER AND LIDAR VERSION 2

The new lidar version is tested with similar measurements as the precedent version. In this version some elements have been eliminated, as the amplifier between the photoreceiver module and the DSP, and the A/D converter is integrated in the DSP itself. The A/D converter is used at 12 bits resolution (the DSP specifies that it can run with 16 bits in a more complex programming). The principal goal of this evaluation with this version 2 is to test a simpler lidar system. A brief summary of the results obtained is presented to avoid the repetition of the all precedent results.

The operation of lidars during daytime shows two problems that can be observed in two simple measurements (fig. 7. and fig. 7.16). The first is the increase of noise due to the sky background radiation and the second the saturation of the photodetector due to the same background, visible in fig. 7. at sunrise. Figure 7.16 shows the loss of signal when the photoreceiver outputs become saturated by sky background radiation around noon.

The filter before the photodetector has a bandwidth of 10 nm. A narrower bandwidth might be employed, depending on the laser line stability ,

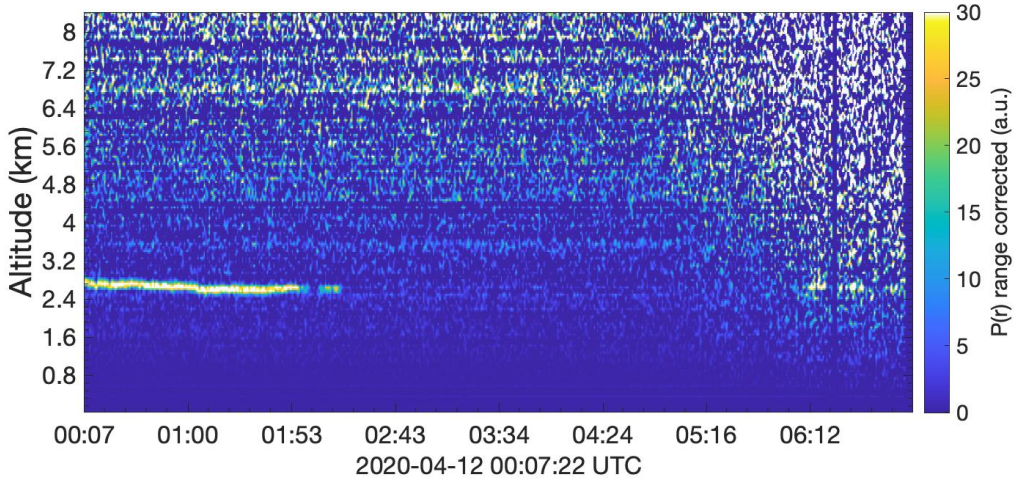


Fig. 7. Increase of the noise level with the sky background radiation at sunrise. M-sequence 2047 bits with 15 seconds of integration time. Clouds are visible during the first part of the night at 3km.

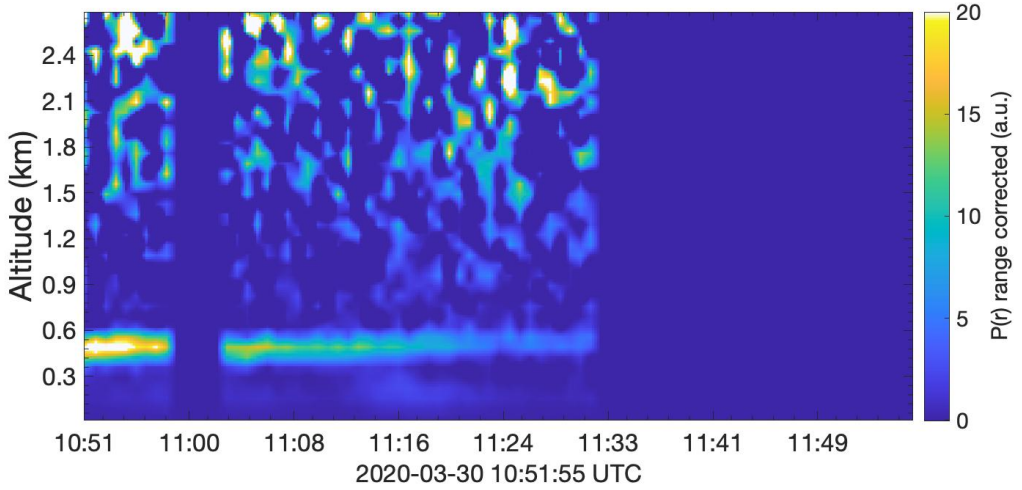


Fig. 7.16. Saturation of the electronics by sky background radiation around noon. M-sequence of 2047 bits with 15 seconds of integration time.

This simplified version 2 of the lidar can also detect clouds as it shows the figures 7. and 7.17. In figure 7.17 rain not reaching ground (virga) is also visible below the cloud

between 21:00 and 21:30. In these figures the integration time is of 15 seconds only, showing clearly the cloud level.

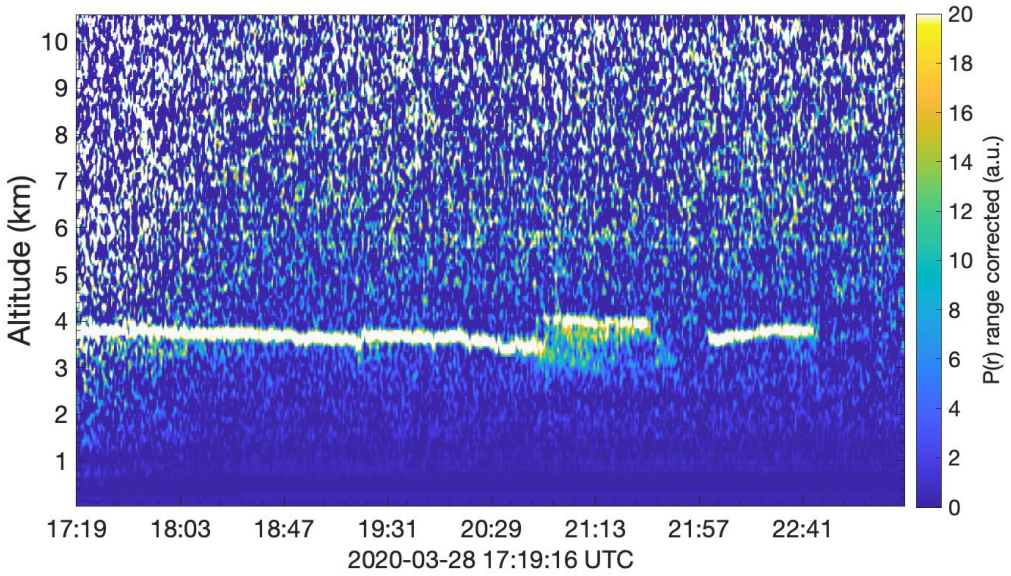
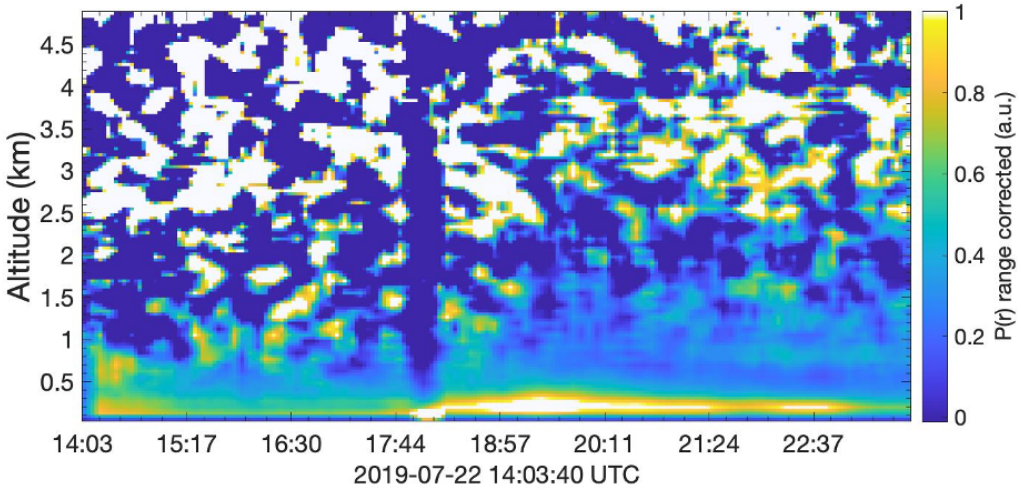


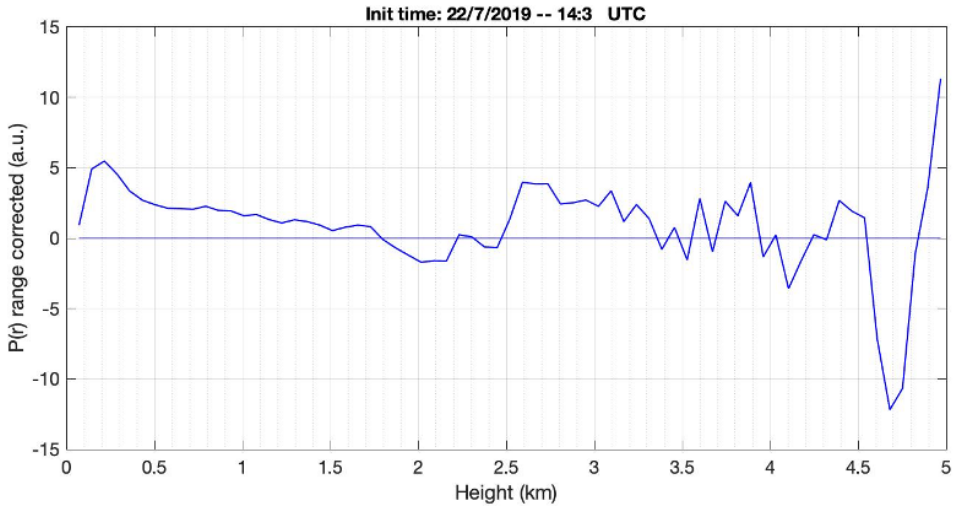
Fig. 7.17. Clouds detected at 3.5-4 km. 15 s of integration time with M-sequence of 2047 bits. Since 21:00 to 21:30 the system observes rain falling from the cloud at high altitude.

The possibility of detecting aerosols with this lidar version is evaluated too. As it has been explained at the beginning of this section, the presence of sky background radiation during daytime can hinder the observation of aerosols or dust. These elements have lesser backscatter coefficients than a cloud and they can be observed during the day basically at low altitudes, except when a strong dust outbreak occurs.

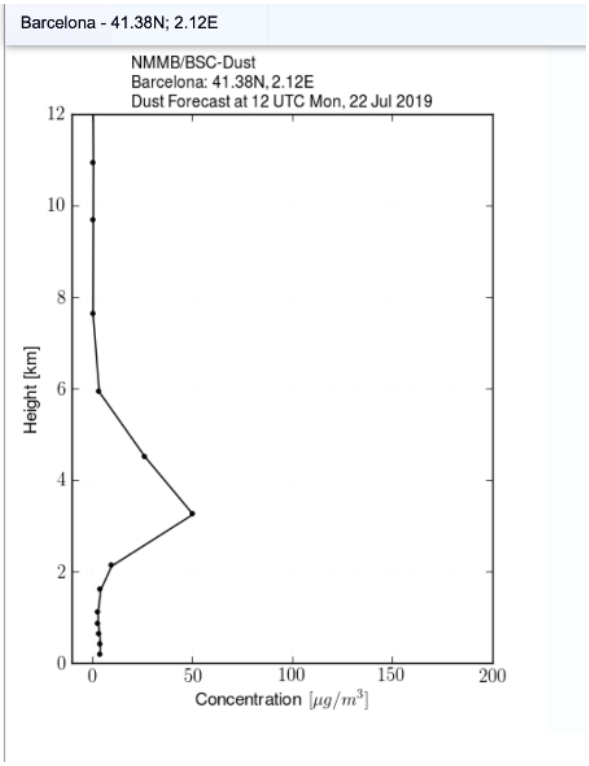
Figure 7.18a shows the presence of aerosols, more dense at low altitude that seems to extend until 1-1.5 km. This layer is better observed in the figure 7.18b, which presents the integration of all the measurements in the time span of fig 7.18a. In the same figure another layer around 3-km altitude can be noticed. Comparing this observation with the dust profile forecast of the Barcelona Supercomputing Center for that day at 12 h UTC (fig. 7.18c), the two graphs seem coherent.


Altitude (km)
4.5
4
3.5
3
2.5
2
1.5
1
0.5
14:03
15:17
16:30
17:44
18:57
20:11
21:24
22:37
2019-07-22 14:03:40 UTC
1
0.8
0.6
0.4
0.2
0
P(r) range corrected (a.u.)

(a)


Init time: 22/7/2019 -- 14:3 UTC
P(r) range corrected (a.u.)
15
10
5
0
-5
-10
-15
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Height (km)

(b)



(c)

Fig. 7.18. Detection of aerosol layer. (a) 2D data showing clearly the low layer of aerosols (integration time of 2 minutes with an M-sequence of 2047 bits). (b) 1-D data integrated over the measurement time of fig (a), revealing a second aerosol layer at around 3km. (c) Dust forecast of the Barcelona Supercomputing Center (BSC) for the same day predicting a layer peaking at around 3 km.

7.6. EVALUATION OF THE RESULTS

The precedent sections show some of the tests done with the equipment working in field conditions. These and other experiments show the capability of this lidar to detect clouds in a near to medium-far range. The maximum altitude at which a specific cloud can be detected depends of course on the reflectivity of the cloud.

At low altitudes the lidar can also detect weaker backscattering layers, as the examples of the fire smoke (fig. 7.15) or aerosols (fig. 7.18) demonstrate. In other test the lidar reveals the presence of light fog or the increasing humidity, probably by the presence of hygroscopic aerosols. The sensitivity of the instrument is limited during the day due to the sunlit sky background radiation, which could be reduced employing a narrower optical filter and/or reducing the receiver field of view. The latter measure would however increase the full overlap range.

COMPARING THE TWO LIDAR VERSIONS

The two versions use the same principle but different hardware implementations. Version 1 has a more complicated and expensive hardware, while the second version tries to simplify the system.

After the test of the two systems it seems that the results are similar, even though the comparison is difficult in field conditions because of the impossibility of having exactly the same atmospheric conditions (the two versions were implemented one after the other, reusing parts, so they couldn't be made to operate simultaneously looking at the same atmosphere).

We summarize the main differences:

DSP. The DSP has no apparent influence in the system performance. As the two processors have a high computing capability, this does not seem an important difference. The most stringent difference is the amount of chip-integrated memory, which can force the design of the code in a particular way, specifically to save and do calculations with the sequences in the memory. This could be avoided designing the circuit to add external memory.

A/D converter. There are two parameters related to this component, probably one of the most critical ones. One of them is the sampling speed, the other the resolution in bits. With the photoreceiver used in the test, with a 1.5 MHz bandwidth, this does not seem to be a restrictive parameter, except for the amount of received energy per pulse (see analysis of trade-off in sections 0 and 4.4), but this situation affects the two versions in the same way. The first version uses an A/D converter external to the DSP (with 16 bits resolution) and the second one uses the A/D converter integrated in the DSP with the option of 12 bits resolution (16 bits is possible to use too). This could make a difference, as a priori a 16-bit resolution seems to be able to detect lower signals, but in the experiments this potential advantage it is not very clear because the noise dominates over the signal. More experiments would be necessary to check if the number of bits in the A/D converter are actually affecting the performance.

Receiver Amplifier. The two versions use a different configuration of the reception amplification chain. Version 1 uses an amplifier with automatic offset correction between the photoreceiver and the A/D converter, while the second version does not (the photoreceiver output is connected directly to the A/D converter input). The use of an additional amplifier does not seem to increase the quality of the system. The function of

compensating the offset signal is in principle useful, but when the background radiation is intense the first element saturated is the built-in amplifier in the photoreceiver module, which makes impossible an offset compensation. The only solution would be to build a complete photoreceiver with offset compensation implemented in the first amplification stages.

Considering all the differences between the two versions, in a general sense no big differences show up from the results. These show that the two system can detect clouds at different altitudes. Aerosols are visible in the two systems too. Perhaps the use of the 16-bit A/D converter shows a marginal performance improvement in SNR. The second version is a cheaper and simpler solution maintaining essentially, except for the previous remark, the system performance.

8. CONCLUSIONS AND FUTURE WORK

The work developed has led to a real operative CW lidar. Two versions of the lidar with significant hardware differences have been evaluated, showing nearly equivalent performances.

This last chapter provides the conclusions and the possible future research work in this type of lidars.

8.1. CONCLUSIONS

The goal of this thesis was the theoretical study, implementation and evaluation of a complete lidar system based on a DSP, a CW laser diode and a solid-state photodetector. No optimization of the complete system is considered during this work. Under these considerations the relevant conclusions are summarized in what follows.

Theoretical evaluation of the performance of the complete lidar system. The evaluation uses the theoretical formulation and the development of a simulation tool in Matlab to include the possible real range of parameter values: noise, different bit resolution of the A/D converter, bit losses in the sampling, saturation and non-linearities of the receiver chain.

This evaluation shows that clouds are clearly possible to detect, the detection of aerosols, of course depending on their backscattering coefficient and the distance to lidar, being more difficult with the limitations of laser peak power set beforehand by the condition of CW laser-diode operation. However, it shows that it is possible to construct a lidar made all in solid state semiconductor components, working in CW mode with pseudorandom sequences as signal modulating the laser power. The low-peak power laser results in a compact and light instrument, with a low amount of energy to operate. This point is detailed in the evaluation of the lidar versions in the previous chapter (0).

The signal-to-noise ratio dependence on measurement time, range-resolution and laser power has been evaluated too. As discussed in section 4.4 for a given measurement time and a given signal-to-noise ratio, an improvement of the range resolution (smaller

ΔR) requires a corresponding increase in laser power (we can talk of peak or mean power because in this lidar the mean power is the half of peak power). The minimum ΔR is set by the receiver electrical bandwidth (Eq. (4.26)). If the bandwidth of the receivers is kept fixed, a decrease of ΔR in a given factor will require an increase of the laser power in the square root of that factor to keep the same signal-to-noise ratio. If the receiver bandwidth is matched to the sequence bit duration, then a decrease in a factor of ΔR demands for an increase in the same factor of ΔR. However, the latter case of "matched receiver" provides always a better signal-to-noise ratio than the fixed-bandwidth receiver for a given power and range resolution. The signal-to-noise ratio is proportional to the square root of the total measurement time. It should be kept in mind however that, from the point of view of the non-ambiguous distance, as well as in the consideration of immunity against non-compensated offsets, to increase the measurement time is better to increase the sequence length than to accumulate the processor outputs from several sequences.

Evaluation of the behavior of the pseudorandom sequences. Several analyses have been done about the behavior of the sequence under different conditions, from which an important conclusions result (detailed study in chapter 3) is obtained: *Non-linearity of the system creates the equivalent of intermodulation products in analogue signals, which results in ghost peaks appearing after the signal processing (correlation). A particular case is the saturation of the signal in any part of the reception system produces the same effect of giving rise to ghost peaks in the detected (after processing) signal*.

From the practical point of view, the following points have been demonstrated in a working system:

Uses of a (relatively) low power laser transmitter. The peak power emitted by the laser is lower than that of usual solid-state lidar systems. In our system, this peak power is as low as 125 mW. The use of a semiconductor (diode) laser increases the efficiency of the emitter.

Low power consumption. Related to the previous point, the use of a semiconductor laser increasing the efficiency of the system contributes to reduce the power needs. A Digital Signal Processor as a core of the system can also reduce the energy consumed in the system dramatically; therefore, it is an open door to a portable system powered by batteries for example. The amount of necessary power is only some watts. The reason to test a second version was to prove that version 2 can decrease the hardware power demand (around only 10 W for the whole lidar) while preserving essentially the system performance. At the same time the necessary hardware is reduced.

Flexibility and adaptability. Emission and reception control are made by software, running on a Digital Signal Processor. No hardware changes are necessary to change the behavior of the lidar. It is possible to change the length of sequences, time duration

of pulse (bit), different coding systems or new data retrieval algorithms. This concept could be named ***Software-Defined Lidar***. In fact, the same instrument could be designed to change continuously the operation procedure as the user or application demands. It can also run as a pulsed lidar, changing the sequence (sending only a pulse) and software used. To use this configuration, it could be necessary to use a more powerful laser, according the trade-offs analysis done in section 0 and 4.4.

Real time operation. The use of a DSP makes it possible to design a CW lidar with pseudorandom sequence working in real time —as other lidar types— because the correlation time of the signal is much smaller than the data capture time. The time used to correlate the received signal is below one second or less. In a well-optimized software for the used processor C6713, the number of mathematical operations is proportional to the product of sequence length by the number of points calculated in the correlation function. The maximum benchmark of the C6713 processor is 1.35 Gflops because mathematical operations are calculated by hardware. Other processing operations can be done in the same DSP: range correction, backscattering profile recovering by means of an inversion algorithm, etc.

In the second version with the C28379 processor (slower than the first one), for example, the speed of the processor is 400 MIPS. To calculate a correlation of 300 points with a length sequence of 2047 bits represents 614000 multiplications approximately. The goal in this work was to demonstrate the operation, but a well-optimized code for the DSP using a fine tuning at low level operations could perform the correlation in the order of milliseconds (614000/400 MIPS =1.5 ms + other operations needed). The use of the standard compiler of the software environment does not guarantees the maximum code efficiency and the use of specific optimized routines can double the time efficiency. The implantation of other strategies as double buffers to receive the data (for example the know ping-pong buffer) can achieve more time efficiency.

Robustness. The lidar is made completely in solid-state semiconductor components, therefore it is adaptable to use it in adverse environments, as well as in airborne or space-borne vehicles. The design can be compacted as the technology allows. The most voluminous part, independently of optics elements, are the power supplies. Using the state of the art of commutated power supplies they can be minimized eliminating the heavy transformers.

The conducted tests show clearly that it is possible to use the system to detect solid targets, clouds at low-medium range and low-altitude aerosols. Clouds have been detected with reasonable integration times and near aerosols too. Clouds as far as 5-8 km are visible with integration times in the order of only 15-30 seconds only. Detection of clouds as far as 10-12 km seems also to be possible, although further tests are necessary to check how reliable the detection of such remote clouds is. In the

experiments, high clouds at 8-10 km altitude have been observed. Solid targets can be detected in a very short time because of their high reflectivity.

For a full comparison with existing systems, the system developed in this thesis has been tested with telescopes of comparable diameters in the small range (diameters of 15 and 20 cm)- Eventually the parameters of this kind of lidar would be determined by the specific applications aimed at.

8.2. FUTURE WORK

This system is open to further research to improve its performance and to turn it into a real operative instrument.

Possible work around this lidar is detailed next.

Research and tests to increase the detection sensitivity, perhaps designing and ad-hoc photodetector module and introducing some corrections in the first stage preamplifier, or cooling it. The effect of cooling on the APD can also be considered. A possibility to increase the sensitivity of the lidar would be reducing the NEP photoreceiver by cooling it, as it is done in other photoreceivers as CCD cameras. Considering that the dark current, responsible of a part of the noise, is exponentially decreasing with the temperature, the effect of cooling could be tested using a Peltier element in the APD detector. The new families of SiPM solid-state photodetectors can be another way of research.

Other coding schemes to obtain more reliability or to solve some of the problems identified. In particular, the effects noticed when the system undergoes non-linear operation of the receiving amplifiers might be solved with another kind of coding, that is, that the non-linear addition of a sequence and its displaced sequence do not generate another sequence. This requires a mathematical exploration out of this work. It is important to maintain the idea that binary codes are robust related to other coding schemes because they allow working with non-linear stages in the emitter. The A2-sequences explained in chapter 3 are a kind of ternary coding and it has not been tested but it would be a possible test with the redesign of the emitter electronics to be linear.

The abundant literature about random sequences opens a way of exploration, as the possibility of new mathematical developments. In fact, the application to lidar requires

very simple conditions, that is, that the sequence autocorrelation has to be as close as possible to a delta-like function and that the number of 1's bits in the sequence has to be high enough to maintain the energetic efficiency.

Exploration of alternative ways for decoding algorithms to detect the signal and retrieve the response of the atmospheric channel. Statistics of order great than 2 (correlation) can be analyzed and tested to detect the signals and to assess their robustness in the presence of noise, as it is done in other areas of signal processing. This line of work would take into account that processors have hugely increased their computing power and that this increasing trend is expected to continue in the coming years. The possibility of using high order statistics [14][15] is pointed out here, but no personal research has been done on it. To obtain the cumulants a higher amount of calculations is necessary. An additional good feature is that, when the bispectrum is obtained from the cumulants of order 3, the symmetric noise distributions, like the Gaussian distribution, are filtered due to the characteristic of the cumulants The operation of the lidar under control of a digital signal processor can be of advantage to implement more sophisticated data processing tools and to add, for example, digital filters to suppress interferences, spurious signals, noise (for example matching the bandwidth to the sequence bit duration as commented before), etc., simplifying the hardware complexity.

CW DIAL systems. Orthogonal sequences in the laser modulation would allow working with more than one laser at different wavelengths while keeping a single photoreceiver front-end, as some authors have proposed [12][48], for the range-resolved detection of gases in open path. In this operation mode two details would require attention: the use of very stable and spectrally pure lasers at the desired wavelength and the selection of adequate sequences to have the lowest cross-correlation possible, ideally 0. Each laser would be modulated with a different pseudorandom sequence. No solution appears to exists according some limits (Welch bounds) [50][51] but a solution to study is the use of sequences that have zero cross correlation in a zone of the complete function [51][52].

After the experience of designing, building and testing this system, a lesson learned is the need of implementing very compact electronics, to avoid long cables interconnecting different parts and circuits. A compact system contributes to eliminate ground loops and other electromagnetic coupling and interference problems.

9. APPENDIX

A.1 PRODUCT OF TWO DELAYED M-SEQUENCES

Considerer two M-sequences as logic sequences A, B and developing the XOR operation

$$A \oplus B = \left(A + B\right)\left(\overline{A} + \overline{B}\right) \tag{9.1}$$

Using one of DeMorgan's theorems this expression is

$$A \oplus B = \left(A + B\right)\overline{AB} \tag{9.2}$$

On the other hand, one of the properties of the M-sequences —see the property (d) at the beginning of chapter 3— is that the XOR of two delayed M-sequences is another M-sequence, which we call C, with a different delay.

$$A \oplus B = C \tag{9.3}$$

Equation (9.2) now is

$$\left(A + B\right)\overline{AB} = C \tag{9.4}$$

Inverting the terms of equation (9.4) and using another DeMorgan's theorems the result is

$$\overline{\left(A + B\right)} + AB = \overline{C} \tag{9.5}$$

And multiplying the two terms of the equation by $A + B$

$$AB\left(A + B\right) = \overline{C}\left(A + B\right) \tag{9.6}$$

But

$$AB\left(A + B\right) = AB \tag{9.7}$$

And equation (9.6) becomes

$$AB = \overline{C}\left(A + B\right) \tag{9.8}$$

Inverting the two sides of the equation the result is

$$\overline{AB} = C + \overline{A}\overline{B} \tag{9.9}$$

A.2 ALIGNMENT PROCEDURE

To have a good alignment of the laser beam in the field of view of the receiver the alignment procedure is done in nighttime. A daytime procedure has to be based on a geometric calculation as described in the part daytime procedure of this section, except with the use of a very narrowband optical filter before the photodetector.

Nighttime procedure

The steps are:

1. Place a CCD camera (only the CCD sensor body, without optical systems) in the focuser (eyepiece holder) of the telescope and set the exposure to some fraction of second.
2. Set the laser to maximum emitting power.
3. Start in a short distance of the telescope with a solid target.
4. Centre the laser beam in the field of the CCD camera.
5. Point the system (telescope + laser + camera) to a more distant point.
6. Align the laser beam to the center of field of view of the camera again.
7. Point the system to the sky. Increase the exposure time of the camera to some seconds, depending on the camera sensitivity. With the camera used in the experiments, an exposure time from 2 to 15 seconds is enough, depending on the presence of particles of dust, droplets, etc., in the atmosphere. Steer the laser

spot to the center of field and correct the focusing. Repeat this operation until the image of figure 6.7(a) is obtained.

8. If the sky is cloudless it usually is possible to see a triangle of light with the vertex somewhere in the field of —fig. 6.7(a). Centre this vertex in the field of the camera steering softly. This vertex is the image of the farthest sections of the laser beam cone of light projected in the atmosphere and illuminating the aerosols. During a clear night this alignment is more difficult to do that in a “dusty night”.
9. If there are clouds, and depending on the camera field of view and the altitude of clouds, a bright spot is seen on the field of view —fig. 6.7 (b)— making easier the alignment task. If the clouds are illuminated by city lights a red or infrared bandpass filter had to be located in front of the CCD camera to increase the contrast of the beam with respect to the background (depending on the laser power).
10. Replace the CCD camera with the APD module and the system is aligned. The lidar must show at least the backscatter from the overlapping area of beam in the field of view of the telescope at near ranges. A fine adjustment of the telescope focus can be done with the APD in the focus. Having clouds in the sky is a good situation for refining the alignment with the APD in the focus.

Daytime procedure

This alignment procedure done at night is usually fast, except if the laser beam is very misaligned, far from the field of view of the camera. In this case a previous procedure is used (fig. A.1). This rough alignment can detect the initial misalignment of the laser respect to the axis of the telescope and it is a first approximation to the correct alignment that have to be made with a distant surface to be precise.

The steps are:

1. Place a laser pointer in the focuser of the telescope —one of the pointers used commonly to align the optical parts of a reflector telescope.
2. Measure the distance between the center of the telescope tube and the center of emitter laser.
3. Point the system to a near solid target like a wall. The two lasers will be visible (the emitter and the laser of inside the telescope). At night or in the darkness the

785 nm laser is visible, if not the alignment with the IR laser is very difficult. Then it is necessary to use an IR camera (a CCD camera usually is sensitive to near IR, and equipped with a visible light cut-off filter or, better, with a narrow band filter will do),

4. Steer the emitter laser to place its footprint on the solid target at the same distance from the footprint of the laser pointer as the distance measured in step 2.
5. At this point the emitter laser is roughly parallel to the axis of the telescope.

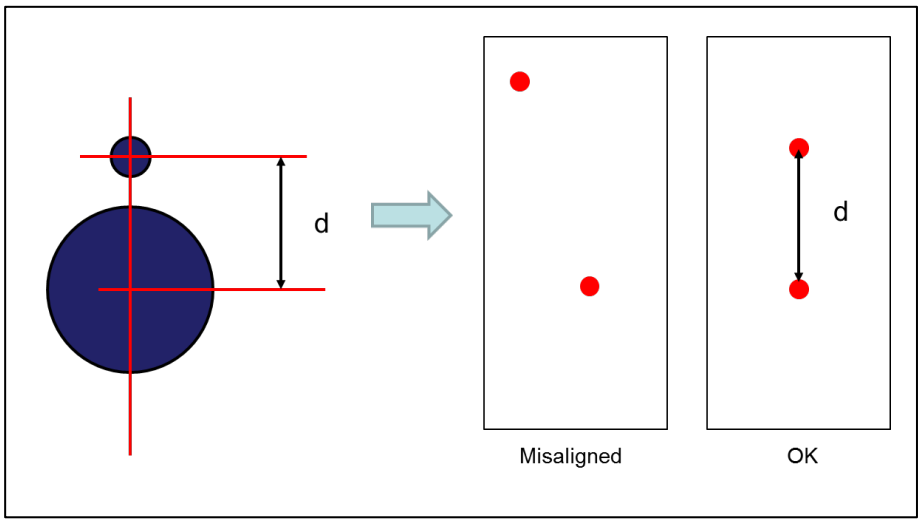
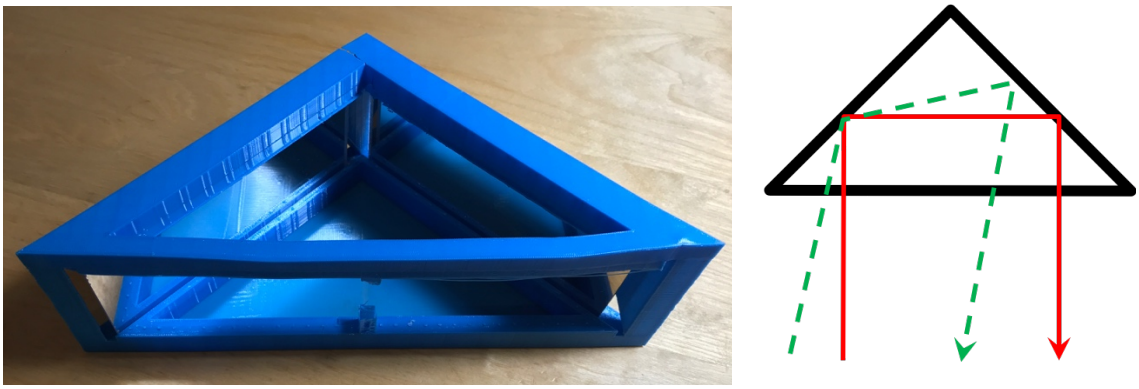


Fig. A.1. Rough alignment of the laser and the receiver axis. When a laser pointer is located in the place of the photodetector in the telescope and the beams are projected on a near flat surface, the distance between the spots has to be the same that between the axis of the telescope and the laser mounted on the telescope.

Alternatively, sometimes a cubic reflector —fig. A.2 (a)— has been used to obtain a rough alignment of the laser and the telescope field of view. The interesting property of this geometrical disposition is that the incident and reflected light rays are parallel independently of the input angle —fig. A.2 (b).



(a) *(b)*

Fig. A.2. (a) Cubic reflector for rough alignment of the laser and the receiver axis. (b) The interesting property of this geometrical disposition is that the incident and reflected light rays are parallel independently of the input angle.

10. REFERENCES

[1] D. Winker, C. Hostetler, and W. Hunt, “CALIOP: The CALIPSO lidar,” *Eur. Sp. Agency, (Special Publ. ESA SP*, vol. 2, no. 561, pp. 941–944, 2004.

[2] M. J. McGill, J. E. Yorks, V. S. Scott, A. W. Kupchock, and P. A. Selmer, “The Cloud-Aerosol Transport System (CATS): a technology demonstration on the International Space Station,” *Lidar Remote Sens. Environ. Monit. XV*, vol. 9612, p. 96120A, 2015.

[3] A. J. Illingworth *et al.*, “The earthcare satellite : The next step forward in global measurements of clouds, aerosols, precipitation, and radiation,” *Bull. Am. Meteorol. Soc.*, vol. 96, no. 8, pp. 1311–1332, 2015.

[4] X. Sun *et al.*, “ICESAT/GLAS Altimetry Measurements: Received Signal Dynamic Range and Saturation Correction,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5440–5454, 2017.

[5] L. Mona, A. Amodeo, and G. D’Amico, “Potentialities and Limits of ICESAT-2 Observation for Atmospheric Aerosol Investigation,” *EPJ Web Conf.*, vol. 119, pp. 2–5, 2016.

[6] A. G. Straume *et al.*, “ESA’s spaceborne lidar mission ADM-Aeolus; Project status and preparations for launch,” *EPJ Web Conf.*, vol. 176, pp. 1–4, 2018.

[7] R. T. H. Collis, R. D. Hake, P. B. Russell, and S. A. Bowhill, “Lidar in Space,” *Opt. Eng.*, vol. 17, no. 1, pp. 23–29, 1978.

[8] A. Arbor, “Lidar from Lidar from orbit + X i,” 2019.

[9] M. O’Brien, “Laser radar gets rugged,” *SPIE Newsroom*, vol. 3, no. 4, p. 21, 2003.

[10] N. Takeuchi, N. Sugimoto, H. Baba, and K. Sakurai, “Random modulation cw lidar,” *Appl. Opt.*, vol. 22, no. 9, p. 1382, 1983.

[11] N. Takeuchi, H. Baba, K. Sakurai, and T. Ueno, “Diode-laser random-modulation cw lidar,” *Appl. Opt.*, vol. 25, no. 1, p. 63, 1986.

[12] C. Nagasawa, M. Abo, H. Yamamoto, and O. Uchino, “Random modulation cw lidar using new random sequence,” *Appl. Opt.*, vol. 29, no. 10, p. 1466, 1990.

[13] B. Sklar, “Spread-Spectrum Techniques,” in *Digital Communications*, Prentice-Hall, 1988, pp. 536–594.

[14] J. Vidai and J. A. Rodriguez Fonollosa, “Impulse response recovery of linear systems through weighted cumulant slices,” *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2626–2631, 1996.

[15] A. Yamani, M. Bettayeb, and L. Ghouti, “High-order spectra-based deconvolution of ultrasonic NDT signals for defect identification,” *Ultrasonics*, vol. 35, no. 7, pp. 525–531, 1997.

[16] J. Hee, “Impulse response measurements using by Change log,” *Program*, no. August, 2003.

[17] G. B. Stan, J. J. Embrechts, and D. Archambeau, “Comparison of different impulse response measurement techniques,” *AES J. Audio Eng. Soc.*, vol. 50, no. 4, pp. 249–262, 2002.

[18] A. R. Møller, “Systems identification using pseudorandom noise applied to a

sensorineural system," *Comput. Math. with Appl.*, vol. 12, no. 6 PART A, pp. 803–814, 1986.

[19] L. Rufer, S. Mir, E. Simeu, and C. Domingues, "On-chip testing of MEMS using pseudo random test sequences," 2004.

[20] W. Charlton, "Linear System Identification using Pseudo Random Binary Signals," 1968.

[21] A. Farina and F. Righini, "Software implementation of an MLS analyzer , with tools for convolution , auralization and inverse filtering," *Signals*, pp. 1–24, 1997.

[22] F. Corsi, C. Marzocca, G. Matarrese, A. Baschirotto, and S. D'Amico, "Pseudo-random sequence based tuning system for continuous-time filters," *Proc. - Des. Autom. Test Eur. Conf. Exhib.*, vol. 1, no. c, pp. 94–99, 2004.

[23] A. Comerón and R. Agishev, "Dimensionless parameters for lidar performance characterization," *Remote Sens. Clouds Atmos. XIX; Opt. Atmos. Propag. Adapt. Syst. XVII*, vol. 9242, p. 92420U, 2014.

[24] R. R. Agishev *et al.*, "Application of the method of decomposition of lidar signal-to-noise ratio to the assessment of laser instruments for gaseous pollution detection," *Appl. Phys. B Lasers Opt.*, vol. 79, no. 2, pp. 255–264, 2004.

[25] R. Agishev and A. Comerón, "Assessment of capabilities of lidar systems in day- and night-time under different atmospheric and internal-noise conditions," *EPJ Web Conf.*, vol. 176, pp. 0–3, 2018.

[26] M. Goresky and A. Klapper, "Pseudonoise sequences based on algebraic feedback shift registers," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1649–1662, 2006.

[27] Z. Tirkel, "Cross-correlation of M-sequences - some unusual coincidences," *IEEE Int. Symp. Spread Spectr. Tech. Appl.*, vol. 3, pp. 969–973, 1996.

[28] D. V. Sarwate and M. B. Pursley, "Crosscorrelation Properties of Pseudorandom and Related Sequences," *Proc. IEEE*, vol. 68, no. 5, pp. 593–619, 1980.

[29] F. J. MacWilliams and N. J. A. Sloane, "Pseudo-Random Sequences and Arrays," *Proc. IEEE*, vol. 64, no. 12, 1976.

[30] J. H. Lindholm, "An Analysis of the Pseudo-Randomness Properties of Subsequences of Long M-Sequences," *IEEE Trans. Inf. Theory*, vol. 14, no. 4, pp. 569–576, 1968.

[31] J. G. Proakis, "Digital Communications," McGraw-Hill, 1989, pp. 396–399.

[32] F. R. K. Chung, J. A. Salehi, and V. K. Wei, "Optical orthogonal codes: design, analysis and applications," *IEEE Trans. Inf. Theory*, vol. 35, no. 3, pp. 595–604, May 1989.

[33] A. Rybaltowski and A. Taflove, "Superior signal-to-noise ratio of a new AA1 sequence for random-modulation continuous-wave lidar," *Opt. Lett.*, vol. 29, no. 15, p. 1709, 2004.

[34] A. Rybaltowski and A. Taflove, "New modulation sequence for random-modulation continuous-wave lidar," *Lidar Remote Sens. Ind. Environ. Monit. II*, vol. 4484, no. January 2002, p. 216, 2002.

[35] A. Rybaltowski and A. Taflove, "Signal-to-noise ratio in direct-detection mid-infrared Random-Modulation Continuous-Wave lidar in the presence of colored additive noise," *Opt. Express*, vol. 9, no. 8, p. 389, 2001.

[36] J. Vanderkooy, "Aspects of MLS Measuring Systems," *J. Audio Eng. Soc.*, vol.

42, no. 4, pp. 219–231, 1994.

[37] M. Wright, “Comments on `{Aspects} of {MLS} Measuring Systems’,” *Journal of Audio Engineering Society*, vol. 43, no. 1/2. pp. 48–50, 1995.

[38] C. Dunn and M. O. Hawksford, “Distortion immunity of MLS-derived impulse response measurements,” *AES J. Audio Eng. Soc.*, vol. 41, no. 5, pp. 314–335, 1993.

[39] Texas Instruments, “Floating point Digital Signal Processor TMS320C6713B - SPRS294B,” no. December 2007, 2006.

[40] Spectrum Digital, “TMS320C6713 DSK Technical Reference,” 2003.

[41] Texas Instruments, “TMS320F2837xS Microcontrollers Technical Reference Manual,” p. 2740, 2019.

[42] Texas Instruments (Texas Instruments), “LAUNCHXL-F28377S Overview User’s Guide,” no. January, pp. 1–25, 2015.

[43] Texas Instruments, “AN-1604 Decompensated operational amplifiers,” no. May, pp. 1–15, 2013.

[44] Texas Instruments, “LMH6624/LMH6626 Single / Dual Ultra Low Noise Wideband Operational Amplifier,” no. November 2002, 2013.

[45] Texas Instruments, “ADS1610 ANALOG-TO-DIGITAL CONVERTER,” no. August 2005, 2006.

[46] Texas Instruments, “ADS1610EVM,” no. May, 2006.

[47] T. Halldórsson and J. Langerholc, “Geometrical form factors for the lidar function,” *Appl. Opt.*, vol. 17, no. 2, p. 240, 1978.

[48] J. M. G. Tijero *et al.*, “Random-modulation CW lidar system for space-borne carbon dioxide remote sensing based on a high-brightness semiconductor laser,” vol. 10563, no. October, p. 103, 2017.

[49] L. Welch, “Lower bounds on the maximum cross correlation of signals (Corresp.),” *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 397–399, May 1974.

[50] J. L. Massey, “On welch’s bound for the correlation of a sequence set,” *IEEE International Symposium on Information Theory - Proceedings*. p. 385, 1991.

[51] P. Z. Fan, N. Suehiro, N. Kuroyanagi, and X. M. Deng, “Class of binary sequences with zero correlation zone,” *Electron. Lett.*, vol. 35, no. 10, pp. 777–779, 1999.

[52] B. Fassi, A. Taleb-ahmed, A. Djebbari, and L. Dayoub, “A New Class of Binary Zero Correlation Zone Sequence Sets,” *IOSR J. Electron. Commun. Eng.*, vol. 5, no. 3, pp. 15–19, 2013.