

# Pseudorandom noise code-based technique for cloud and aerosol discrimination applications

Joel Campbell, Narasimha S. Prasad, Michael Flood, Wallace Harrison

NASA Langley Research Center, 5 N. Dryden St., Hampton, VA 23681

## Abstract

NASA Langley Research Center is working on a continuous wave (CW) laser based remote sensing scheme for the detection of CO<sub>2</sub> and O<sub>2</sub> from space based platforms suitable for ACTIVE SENSING OF CO<sub>2</sub> EMISSIONS OVER NIGHTS, DAYS, AND SEASONS (ASCENDS) mission. ASCENDS is a future space-based mission to determine the global distribution of sources and sinks of atmospheric carbon dioxide (CO<sub>2</sub>). A unique, multi-frequency, intensity modulated CW (IMCW) laser absorption spectrometer (LAS) operating at 1.57 micron for CO<sub>2</sub> sensing has been developed. Effective aerosol and cloud discrimination techniques are being investigated in order to determine concentration values with accuracies less than 0.3%. In this paper, we discuss the demonstration of a PN code based technique for cloud and aerosol discrimination applications. The possibility of using maximum length (ML)-sequences for range and absorption measurements is investigated. A simple model for accomplishing this objective is formulated, Proof-of-concept experiments carried out using SONAR based LIDAR simulator that was built using simple audio hardware provided promising results for extension into optical wavelengths.

**Keywords:** ASCENDS, CO<sub>2</sub> sensing, O<sub>2</sub> sensing, PN codes, CW lidar

## 1. INTRODUCTION

The National Research Council's (NRC) Decadal Survey (DS) of Earth Science and Applications from Space has identified the Active Sensing of CO<sub>2</sub> Emissions over Nights, Days, and Seasons (ASCENDS) as an important Tier II space-based atmospheric science mission. The CO<sub>2</sub> mixing ratio needs to be measured to a precision of 0.5 percent of background or better (slightly less than 2 ppm) at 100-km horizontal resolution overland and 200-km resolution over oceans. To meet this goal, the ASCENDS mission requires simultaneous laser remote sensing of CO<sub>2</sub> and O<sub>2</sub> in order to convert CO<sub>2</sub> column number densities to average column CO<sub>2</sub> mixing ratios (XCO<sub>2</sub>). As such, the CO<sub>2</sub> column number density and the O<sub>2</sub> column number density will be utilized to derive the average XCO<sub>2</sub> column. The anticipated benefits of ASCENDS as discussed in the decadal survey include (a) quantification of global spatial distribution of atmospheric CO<sub>2</sub> on scales of weather models as well as terrestrial and oceanic sources and sinks of CO<sub>2</sub> during day/night over all seasons, and (b) establish a scientific basis for future projections of CO<sub>2</sub> sources and sinks through data-driven enhancements of Earth-system process modeling.

Accordingly, NASA Langley Research Center (LaRC), with its partners, is working on a CW laser absorption spectrometer based remote sensing scheme operating in the 1.57 micron spectral band for the detection of CO<sub>2</sub> and 1.26 micron spectral band for the detection of O<sub>2</sub>. For concentration determination, differential absorption lidar (DIAL) scheme at a selected transition for each gas is adopted. Hence, two are more wavelength at a known transition are utilized. The 1.5 micron spectral band lying in the telecom region was chosen for CO<sub>2</sub> detection due to spectroscopic properties combined with architectural advantages. The selection of 1.26 micron band for O<sub>2</sub> detection that is close to 1.57 micron spectral band instead of 0.76 micron spectral band is anticipated to provide processing advantages. Within these spectral bands, an optimal transition that is least sensitive to environmental parameters for each of these trace gases is selected. With limitations on CW laser powers, the remote sensing systems may have to operate in the photon starved regions. However, an Intensity modulated continuous wave (IMCW) technique is adopted to achieve highest possible

SNR for processing return signals to minimize the influence of various noise sources within the filter bandwidths for each channel.

Unlike pulsed lasers based DIAL systems, CW laser based remote sensing systems are affected by cloud and aerosol interferences. Hence, fast and effective discrimination techniques have to be adopted during data retrieval process. Accordingly, various realizable aerosol and cloud discrimination techniques are being investigated for determining accuracies less than 0.5%. In this paper, the adaptation of a Pseudorandom Noise (PN) code based technique to discriminate returns from clouds, aerosols and ground is discussed. This is followed by proof-of-concept experiments carried out using acoustic frequencies.

### 1.1 Resolution and Range

The minimum resolution is given by

$$r = \frac{c}{2 * B_R} \tag{1}$$

where c is the speed of light and B<sub>R</sub> is the bit rate of the PN code. The maximum range is given by

$$R = (N - 1) * r \tag{2}$$

where N is the code length. For example, for a PN code of length 255 bits and a bit rate of 50 kHz, the maximum range would be 762 km, which is the range we would need for a satellite. However, there are advantages to making the code longer as discussed in Section below.

## 2. PN CODES

PN codes are well known in the communications devices for their use in encoding communication channels such as in spread spectrum communications. One of the simplest PN codes is the m-sequence. Furthermore, m-sequences are basis of many other PN codes. We also refer them as mL-maximum length sequence. Besides radar and sonar, the m-sequences have been used in Lidar applications for measuring range<sup>1,2</sup>. Pulse lidars inherently provide range information. However, in the case of CW laser based remote sensing systems, range information on clouds and hard targets can be obtained by transmitting amplitude modulated PN code. Range profiles can be extracted by cross correlation between the reflected signal and the transmitted code. The correlation peaks of range profiles will indicate the location of scattering centers such as clouds and ground.

New codes that are orthogonal to each other over a limited operational range can be generated. This gives us the ability to design multichannel DIAL systems. As such, it allows one to measure relative absorptions of two or more laser wavelengths at a particular range to detect specific atmospheric species. Besides m-sequences, other codes are also useful for ranging<sup>3</sup>. However, use of PN codes at multiple laser wavelengths for measuring absorption at specific ranges in order to discriminate the return of a cloud from the ground is supposed to be practically effective and simple from implementation standpoint. For our application, the m-sequences are especially advantageous because they have very good autocorrelation properties.

The mL sequences can be represented in a number of different ways. One popular method is by using linear feedback registers with modulo 2 additions<sup>4</sup> and it is shown in Figure 1.



Note that some people define this differently by dividing by the length. This works but it requires something of the order of  $N^2$  steps. A better and faster way is to use Fourier transforms. In this case,

$$R[n] = F^{-1} \left[ F[A] F[B]^* \right], \quad (7)$$

where the \* indicates the complex conjugate and F is the FFT. This is superior because it only takes something of the order of  $N \log N$  steps. As an example we take the case in the previous section and compute the cross correlation between it and a code shifted 128 places to the right and the result is shown in Figure 2.

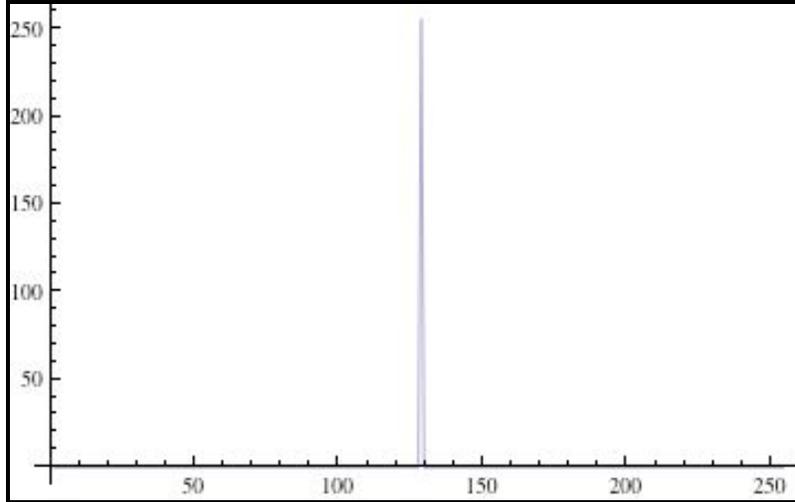


Figure 2. Cross correlation of code and shifted code computed with FFT.

## 2.2 Signal-to-noise ratio estimates

We first consider white noise. It is possible to estimate the signal to noise ratio improvement before and after correlation. In this situation, we have a DAQ sampling the signal at

$$M = \frac{S}{B_R} \quad (8)$$

Here, M is the samples per bit,  $S_R$  is the DAQ sample rate and  $B_R$  is the PN code bit rate. This increases the apparent code length by M, which must be taken into account when applying the correlation calculation and the new code length becomes MN samples long.

Let suppose we have an AC coupled signal which we represent by

$$S = \eta + \alpha pn^\Delta \quad (9)$$

where  $\eta$  is the noise,  $\alpha$  is the attenuation factor and

$$pn_i^\Delta = 2PN_{i+\Delta} - 1. \quad (10)$$

In Eq. 10, the term PN is the PN code represented by 1's and 0's and  $\Delta$  represents the phase shift where as the term pn contains 1's and -1's.. The standard deviation of the noise we represent by

$$\sigma = \sqrt{\frac{\sum_{i=1}^{MN} (\eta_i - \eta_{ave})^2}{MN}}, \quad (11)$$

where we have summed over the total length of the code. After performing the correlation calculation, we see that the signal strength increases by a factor of MN where they are correlated. We now ask the question what correlation does to the standard deviation of the noise. The correlation function of the total signal is,

$$R(j) = \sum_{i=1}^{MN} (\eta_i + \alpha pn_i^\Delta) pn_{i+j}^0 = \sum_{i=1}^{MN} \eta_i pn_{i+j}^0 + \alpha \sum_{i=1}^{MN} pn_i^\Delta pn_{i+j}^0 = \eta_j' + S_j' \quad (12)$$

The question now is to find the standard deviation of the transformed noise  $\eta'$  which we assume to have 0 mean (AC coupled). We first find the variance by

$$(\sigma')^2 = \langle (\eta')^2 \rangle = \frac{1}{MN} \sum_{j=1}^{MN} (\eta_j')^2 \quad (13)$$

To find the sum, we write

$$(\eta_j')^2 = \sum_{m=1}^{MN} \eta_m pn_{m+j}^0 \sum_{n=1}^{MN} \eta_n pn_{n+j}^0 = \sum_{n,m} \eta_m \eta_n pn_{m+j}^0 pn_{n+j}^0 \quad (14)$$

so that

$$(\sigma')^2 = \frac{1}{MN} \sum_{j=1}^{MN} \sum_{n,m} \eta_m \eta_n pn_{m+j}^0 pn_{n+j}^0 \quad (15)$$

If the noise has 0 mean, the only part of the sum that will contribute is when n=m so

$$(\sigma')^2 = \frac{1}{MN} \sum_{j=1}^{MN} \sum_m \eta_m^2 (pn_{m+j}^0)^2 = \frac{1}{MN} \sum_{j=1}^{MN} \sum_m \eta_m^2 = \frac{1}{MN} \sum_{j=1}^{MN} MN \sigma^2 = MN \sigma^2 \quad (16)$$

In the above equation, since pn is either a 1 or -1, its square is always 1. The result is that

$$\sigma' = \sqrt{MN} \sigma \quad (17)$$

The new correlated snr' in terms of the original snr is

$$\frac{snr'}{\sqrt{MN}\sigma} = \frac{NM\alpha}{\sqrt{MN}\sigma} = \sqrt{MN} snr . \quad (18)$$

Expressed differently in terms of previously defined variables, we find

$$\frac{snr'}{snr} = \sqrt{MN} = \sqrt{\frac{samplerate}{codebitrate}} N . \quad (19)$$

We may use this in combination range and resolution equations in order to optimize the system.

In the event the signal processed with P codes back to back it is easy to verify

$$\frac{snr'}{snr} = \sqrt{MNP} = \sqrt{\frac{samplerate}{codebitrate}} NP \quad (20)$$

The Eq. 20 is useful because it indicates we may process an entire block of data at once to get an averaging effect and we may arbitrarily increase the signal to noise by simply taking large amount of data. For noise with a 1/f power spectrum we have been able to verify that

$$\frac{snr'}{snr} \geq \sqrt{MNP} = \sqrt{\frac{samplerate}{codebitrate}} NP \quad (21)$$

### 3. NUMERICAL RESULTS

The numerical computations were carried out using Mathematica 7. As a comparison, we compute white noise as a series of pseudorandom numbers ranging from -1 to 1. To generate pink noise we may take white noise and filter it with an FFT filter using  $1/\text{Sqrt}[f]$  as a filter kernel, thereby giving it a power spectrum of  $1/f$ .

As a simple simulation we may represent the return signal from a hard target as a combination of the delayed PN code in combination with noise. If M is the number of samples per bit of the DAQ and N is the length of the code, the strength of the signal after cross correlation will be  $N*M$  when the return signal and PN code are in sync, and  $-M$  when they are out of sync in the absence of noise.

As an example, we choose 1 as the amplitude of the return signal in combination with random noise 10 time that level. We choose a code that is 511 bits long, a bit rate of 50000 bits/sec, and a DAQ sample rate of 500000 samples/sec. This gives a resolution of 3 km and a range of 1533 km. We use a code with the generating function  $G = x^9 + x^7 + x^5 + x^2 + 1$  and a seed of (1,0,1,0,1,1,1,1). We also choose a delay corresponding to 500 km. Figure 3 shows the return signal in absence of noise. Figure 4 shows the same signal buried in noise that is 10 times the amplitude of the signal. After correlation the signal can be easily extracted.

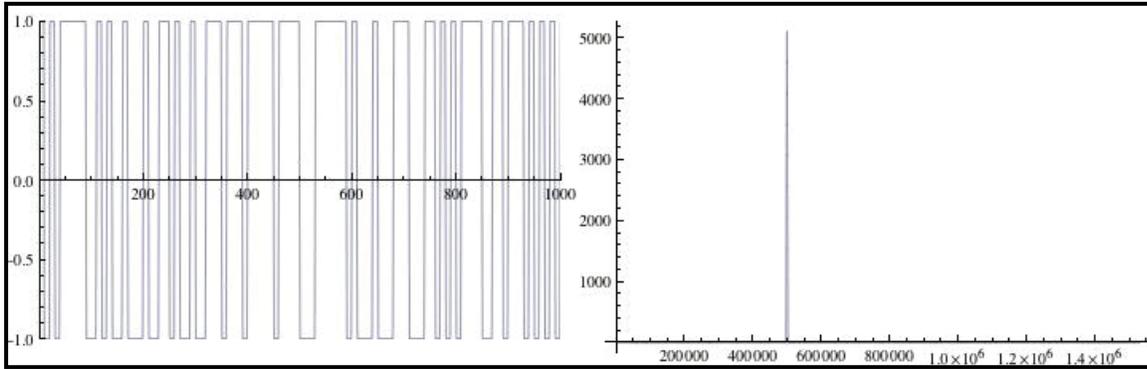


Figure 3. Return signal less noise before and after correlation.

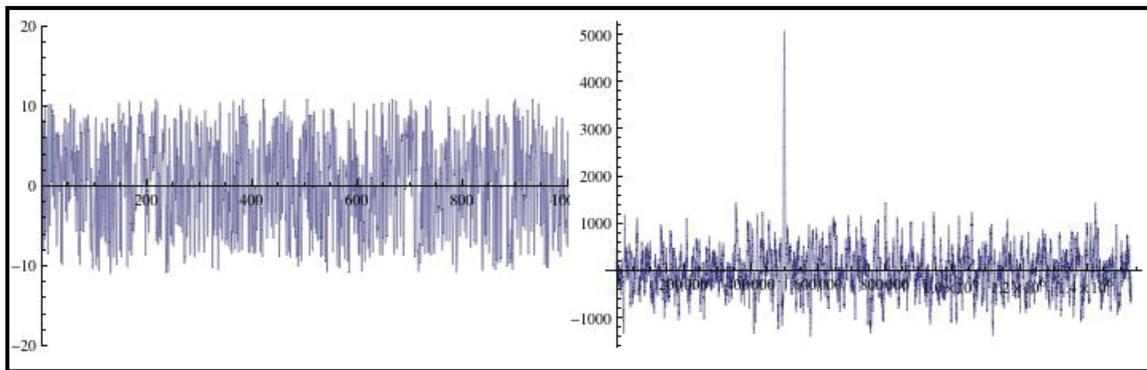


Figure 4. Return signal with 10x white noise before and after correlation.

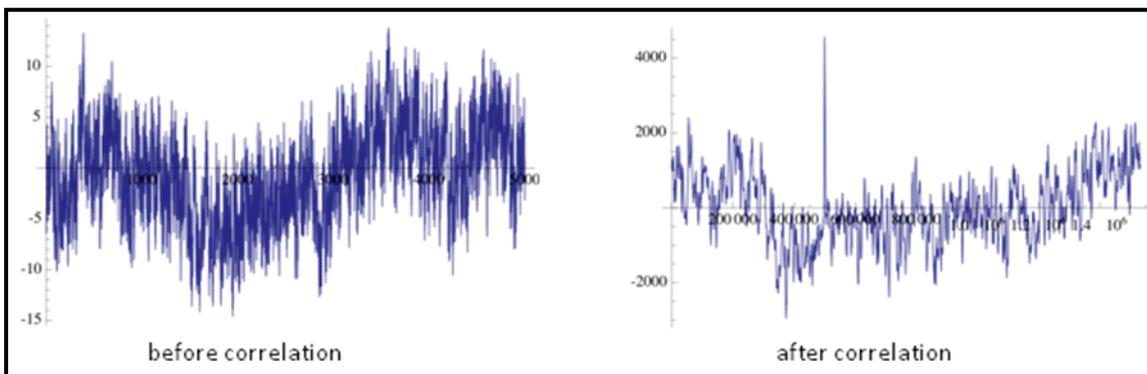


Figure 5. Return signal with 10x pink noise before and after correlation.

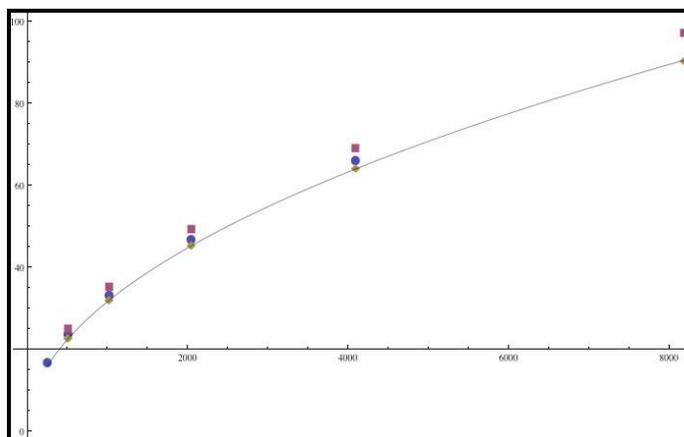


Figure 6. Plot of SNRgain as a function of MNP for different cases. The solid line is for ideal white noise and the other points are for different cases of pink noise.

### Multiple channels

In order for this to be useful for detection of atmospheric species, it is advantageous to be able to decode multiple channels. This is still a work in progress but we believe that there is a solution that works. We may take advantage of two facts to use m-sequences (for instance) to construct a system with two or more channels. Since m-sequences have very good autocorrelation properties shifted versions of them will cross correlate well except for a small DC component where they are uncorrelated and the big spike where they correlate. The second aspect we have with this application is when the LIDAR is looking down from space there will be a final hard target (the ground) below which there will be no further returns possible. That opens up some interesting possibilities. If we know there will be no further returns past a certain distance we can size a system with twice the range we need and use that further range to add a second shifted code that is shifted by half the doubled range. To get twice the range we simply use a PN code twice as long while maintaining the same code bit rate. As long as there are no returns beyond that half range there will be no interference between the channels. The first half of the autocorrelation gives the range information for the first un-shifted channel and the second half of the autocorrelation gives the range information for the shifted second channel and we can do that in a single step. By doubling the length of the code we also increase the SNR as we see from the previous section. We can make the code as long as we need to add more channels.

In the sample below we use a code length of 511 with a sample rate of 500000 samples/sec and a code bit rate of 50000 bits/sec. This gives us a total range of  $511 \cdot 3 \cdot 10^5 / (2 \cdot 50000) = 1530$  km. To generate the PN code we use a code with the generating function  $G = x^9 + x^7 + x^5 + x^2 + 1$ , a seed of (1,0,1,0,1,1,1,1,1) and a shifted version of that code shifted 256 places to the right. Figure 5 illustrates return signal with 10x pink noise before and after correlation.

These results illustrate that the signal to noise ratio improves significantly with the length of the code, number of code blocks processed simultaneously, and the amount of oversampling.

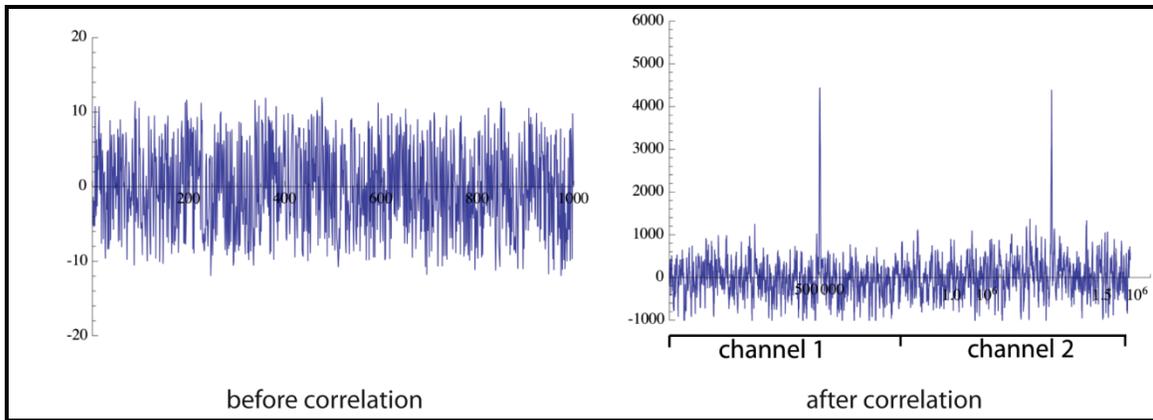


Figure 7. Example of using two codes in a double ranged system with two channels.

#### 4. EXPERIMENTAL RESULTS AND DISCUSSION

To test these concepts we implement the system shown in Figure 8 in audio. Because an audio speaker is not capable of reproducing the very low frequencies required by the PN code, one must first put the PN code on a carrier, then modulate that carrier with the PN code. Here the resolution and range are represented by Eqs. 2 and 3 except that  $c$  becomes the speed of sound in this case.

In the initial development, this system was first implemented in Mathematica<sup>5</sup> then in Labview. In order to accomplish this, special LabVIEW audio drivers written by Christian Zeitnitz<sup>6</sup> were utilized. This increased both the performance and functionality of the Labview audio interface. A Samson Servo 201a audio amp was used for the audio amplifier. This had a bandwidth of 65 kHz and is designed for accurate sound reproduction. The audio card was an M-Audio 1814 with a 1394a interface. This card is capable of sample rates of up to 192 kHz.

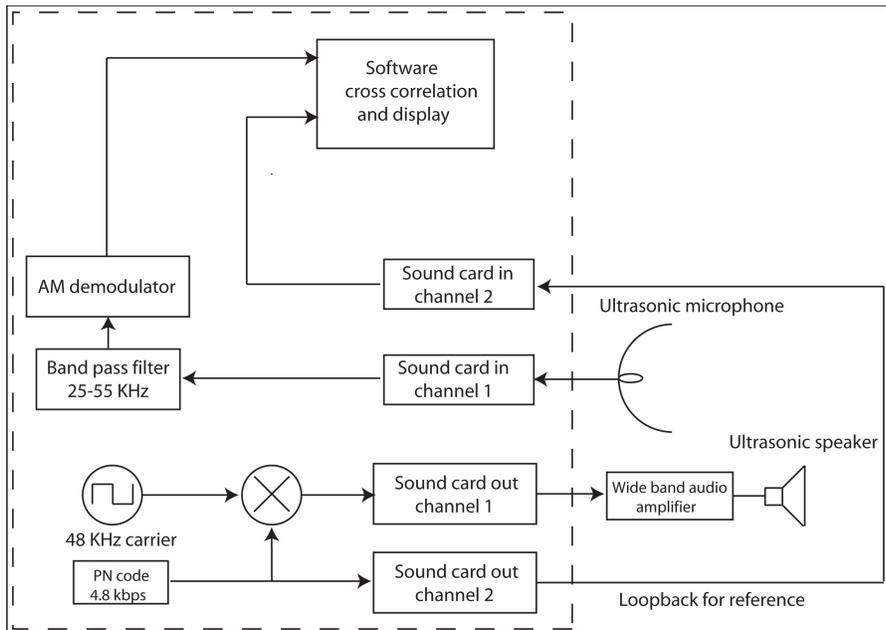


Figure 8. Audio implementation for single channel audio ranging system. All modulation/demodulation, filtering, and cross correlation functions were implemented in Labview software.

A dual channel implementation is also possible which could potentially be used for differential absorption measurements. This is analogous to the proposed LIDAR implementation and can potentially be used for differential absorption measurements.

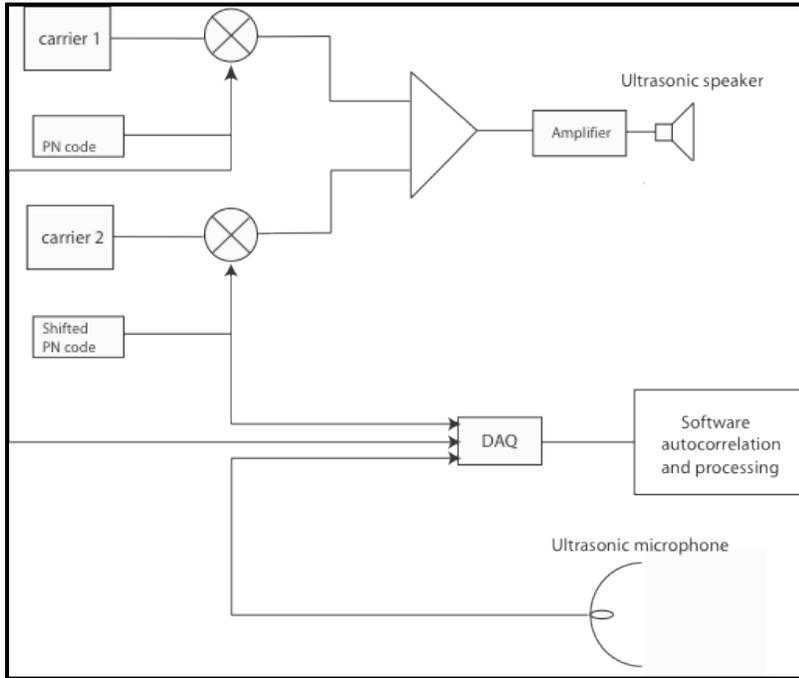


Figure 9. Dual channel implementation of audio ranging system can potentially be used for differential absorption measurements. This is analogous to how it could potentially be implemented in LIDAR.

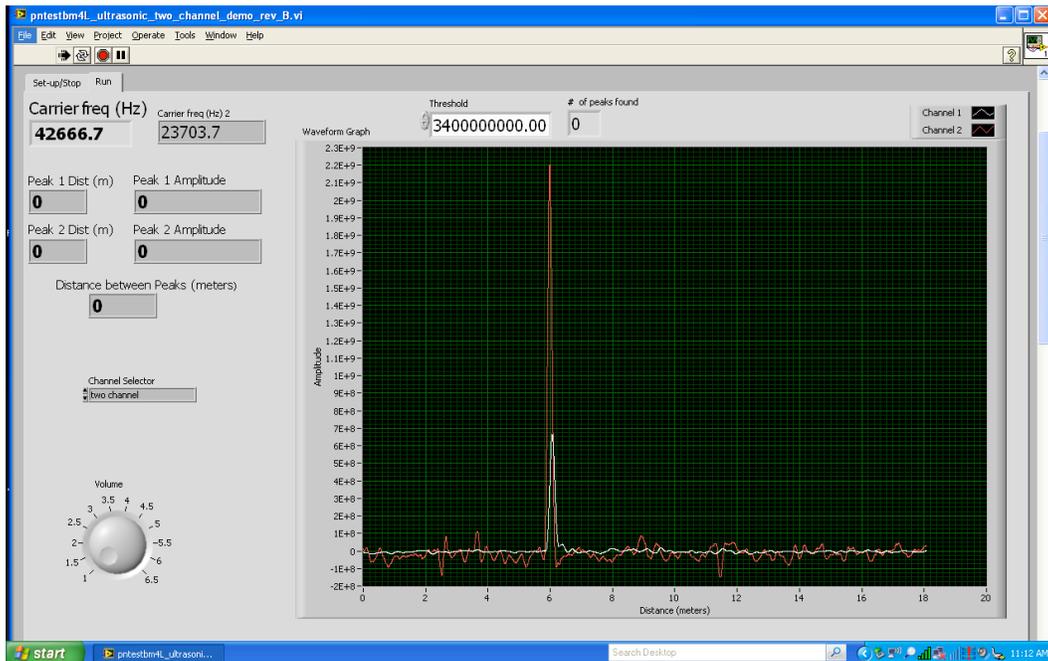


Figure 10. Labview interface and display for the dual channel implementation.

The proposition is to use two or more PN codes where each is shifted in time in a LIDAR implementation. The shifting of the codes is done such that there is equal spacing. For instance with two codes of length 511, the second is shifted to the right 256 places with respect to the first code. The shifting is done in such a way that as the end of the code is shifted past the boundary it wraps to the beginning. The maximum distance to the farthest target shall be 255 in that case. With

four codes of length 511, each is shifted 128 places to the right of the preceding code. The maximum distance to the target in that case shall be 127. In this way one may have multiple channels for a given code length. The only restriction is that the LIDAR/RADAR/SONAR must be pointed at targets with a hard surface (such as the surface of the earth) no farther than the distance to the first code. Without that one will have range-wrapping issues. In addition to this one should either subtract the average from the code or from the data prior to processing. This helps eliminate noise/bias issues.

## 5. CONCLUSIONS

In this paper, a novel method based on PN codes to discriminate clouds and aerosols from ground returns suitable for CW laser based remote sensing systems for space based platforms is discussed. The intended use of this technique is in the ongoing ACTIVE SENSING OF CO<sub>2</sub> EMISSIONS OVER NIGHTS, DAYS, AND SEASONS (ASCENDS) program effort at NASA Langley Research Center. It is shown by numerical methods that time-shifted PN codes would be valuable for obtaining range information of interfering clouds and aerosols so that returns from ground or topographic background can be discriminated during data retrieval processes. Proof-of-concept prototype operating at audio frequencies has been successfully demonstrated. Plans are underway for demonstration at optical frequencies.

## REFERENCES

- [1] N. Takeuchi, N. Sugimoto, H. Baba, and K. Sakurai, "Random modulation CW lidar," *Applied Optics*, Vol. 22, No. 9, May. 1, 1983.
- [2] Nobuo Takeuchi, Hiroshi Baba, Katsumi Sakurai, and Toshiyuki Ueno, "Diode-laser random-modulation cw lidar," *Applied Optics*, Vol. 25, No. 1, Jan. 1, 1986.
- [3] Adam Rybaltowski and Allen Taflove, "Signal-to-noise ratio in direct-detection mid-infrared, Random-Modulation Continuous-Wave lidar in the presence of colored additive noise," *Optics Express*, Vol. 9, No. 8, Oct. 8, 2001.
- [4] Linear feedback shift registers, New Wave Instruments  
[http://www.newwaveinstruments.com/resources/articles/m\\_sequence\\_linear\\_feedback\\_shift\\_register\\_lfsr.htm](http://www.newwaveinstruments.com/resources/articles/m_sequence_linear_feedback_shift_register_lfsr.htm).
- [5] Joel F. Campbell, N. Prasad, Michael Flood, Fenton Harrison, "Experiments in Mathematica with spread spectrum SONAR," Wolfram Technology Conference, October 14, 2010.
- [6] Christian Zeitnitz, WaveIO: A soundcard interface for Labview: <http://www.zeitnitz.de/Christian/waveio>.