Advanced sine wave modulation of continuous wave laser system for atmospheric CO$_2$ differential absorption measurements

Joel F. Campbell,* Bing Lin, and Amin R. Nehrir

NASA Langley Research Center, Hampton, Virginia 23681, USA

*Corresponding author: joel.f.campbell@nasa.gov

Received 11 September 2013; revised 16 December 2013; accepted 18 December 2013; posted 24 December 2013 (Doc. ID 197456); published 4 February 2014

In this theoretical study, modulation techniques are developed to support the Active Sensing of CO$_2$ Emissions over Nights, Days, and Seasons (ASCENDS) mission. A continuous wave (CW) lidar system using sine waves modulated by maximum length (ML) pseudo-noise (PN) codes is described for making simultaneous online/offline differential absorption measurements. Amplitude and phase-shift keying (PSK) modulated intensity modulation (IM) carriers, in addition to a hybrid-pulse technique are investigated, which exhibit optimal autocorrelation properties. A method is presented to bandwidth limit the ML sequence based on a filter implemented in terms of Jacobi theta functions, which does not significantly degrade the resolution or introduce sidelobes as a means of reducing aliasing and IM carrier bandwidth.

OCIS codes: (280.1910) DIAL, differential absorption lidar; (300.6360) Spectroscopy, laser; (170.4090) Modulation techniques.

http://dx.doi.org/10.1364/AO.53.000816

1. Introduction

NASA Langley Research Center in collaboration with ITT Exelis has been evaluating continuous wave (CW) laser absorption spectrometer (LAS) systems as a means of retrieving atmospheric CO$_2$ column amounts from space to support the Active Sensing of CO$_2$ Emissions over Nights, Days, and Seasons (ASCENDS) mission. The U.S. National Research Council’s report, entitled Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, identified the need for an active lidar space mission. This report targeted a need for measurements of column CO$_2$ to gain a better understanding of CO$_2$ sources and sinks through the active sensing of ASCENDS [1]. The primary objective of ASCENDS is to make CO$_2$ column-mixing ratio measurements during day and night over all latitudes and all seasons and in the presence of thin or scattered clouds. These measurements would be used to significantly reduce the uncertainties in global estimates of CO$_2$ sources and sinks, to provide an increased understanding of the connection between climate and atmospheric CO$_2$, to improve climate models, and to close the carbon budget for improved carbon and chemical weather forecasting and as well as implementing policy decisions [1].

The critical component in the ASCENDS mission is the LAS for CO$_2$ column measurements with weighting toward the mid- to lower troposphere using the integrated path differential absorption (IPDA) technique from the surface and cloud backscattered returns.

IPDA lidar systems measure the difference of the total gas absorption along the path length between two or more laser wavelengths in the path of the laser to a scattering target. For the Earth’s climate system, there are large changes in the atmospheric variables such as cloud and aerosol distributions, and the profiles of water vapor and other absorbing gases from regional to global scales. Also, the surface reflectance varies with many factors and variables, such as surface type, roughness, slope, vegetation,
and soil moisture. The variations in these atmospheric and surface variables can generate large changes in received lidar powers from the surface or other targets, which can introduce large biases and artifacts in determination of CO₂ column amounts when only a single laser with a single wavelength is utilized. The advantage of IPDA LAS system is that the strong effects of the variations in atmospheric states and surface reflection on lidar return powers would be minimized from the ratio of the simultaneously received lidar powers from the two closely spaced online and offline IPDA wavelengths. All the influences on the lidar return signal from these environmental variables are the same for the two or more closely spaced wavelengths. The ratio tends to minimize the environmental effects except the differential absorption optical depth at the desired LAS wavelengths. Thus, with all kinds of the considerations including experimental factors, atmospheric temperature and pressure profiles, and range estimations, high accuracy CO₂ column measurements can be achieved from the CO₂ differential optical depth.

For the purposes of this paper, we consider the two-wavelength approach, which uses one laser with the wavelength tuned to the CO₂ gas absorption line (online), and the other wavelength tuned away from the absorption feature but still within 100 pm from that same absorption line (offline). The differential absorption optical depth for the CO₂ column may be derived from the ratio of the received optical power from each wavelength. There are several different approaches to making this measurement.

CO₂ IPDA LAS measurement approaches include pulsed [2,3] and CW systems. Direct-detection pulsed systems use a technique where measurements for each of the IPDA wavelengths are sampled successively via sequential laser shots sampled through different atmospheric columns and surface spots at the surface. The CW approach described here transmits the on- and offline lidar signals to the surface simultaneously from a single laser transmitter, thereby allowing for simultaneous sampling of the atmospheric column background noise and surface reflectivity within the laser spot, thus reducing the potential errors. In addition, a major advantage in this approach is the significant reduction of low-frequency variations in background radiation and colored noise, which are filtered out in a natural way with IM carrier-based matched-filter approaches by effectively limiting the detection frequency and bandwidth. This is done through the demodulation process shown in Fig. 2, which limits the frequency range to a band of frequencies about the carrier frequency, thus filtering out unwanted low-frequency noise. Within the CW approach, IM techniques are used to separate the simultaneous return signals from the different IPDA wavelengths, providing ranging capability and discriminate surface returns from other atmospheric returns, such as aerosols and clouds. The IM technique is also useful for extracting the laser signals from solar background noise, which occurs as a by-product of the matched-filter technique used to retrieve the science signal of interest. Potential IM-CW techniques include PN code CW [4,5], linear swept sine wave CW [6–11], unswept sine wave CW [12–16], nonlinear swept sine wave CW [17], etc. The baseline technique used by the NASA Langley Research Center in partnership with ITT Exelis for the ASCENDS mission is the linear swept frequency CW [11]. The utility of this technique for both range and CO₂ column absorption measurements has been demonstrated over the past several years through ground-based and airborne demonstrations [11].

A drawback of various modulation techniques, such as the linear swept frequency, is the potential for generation of sidelobes in the autocorrelation function, which could introduce small bias errors when intermediate scatterers such as thin clouds are close to the measurement target. This is the main advantage of digital techniques such as PN codes and nonlinear frequency modulation. Previously, PN codes have been described using a pure maximum length (ML) sequence with time shifting to separate channels [4]. This technique may have the advantage of eliminating the sidelobe issue for certain hardware configurations in LAS systems, while allowing for multiple channels that share the same bandwidth. Additionally, the time-shifting approach allows for simple and straightforward data processing of multi-channel laser returns since only a single matched-filter correlation needs to be performed for all channels instead of a separate matched-filter correlation for each channel. One potential disadvantage with this technique is that not all hardware configurations have a suitable low-frequency response for transmitting and/or receiving a pure ML sequence without distortion. Since an ML sequence has a main power band at zero frequency, using such a technique in band-limited hardware (high-pass or bandpass filtered in particular) will filter out the main power band and corrupt the autocorrelation properties, which can result in nonorthogonality between channels in a time shifted multichannel configuration. This problem could be addressed by imparting a PN code amplitude modulation on a narrowband carrier at a predetermined frequency, performing amplitude demodulation, and conducting matched filter correlation with the reference PN code to obtain the range capability. A low-cost demonstration of this was implemented using Hilbert transforms for amplitude demodulation. This demonstrated much better autocorrelation properties than a linear swept frequency but still exhibited minor artifacts, and the demodulated PN code exhibited minor distortions and ringing [5]. This was a student demonstration implemented in audio and was instrumental in demonstrating remote sensing concepts using PN codes on a carrier and was also a nice demonstration outside of the classroom as well. In this paper, we eliminate those minor issues by implementing...
an improved demodulator. We also introduce improved modulation techniques beyond amplitude modulation, which improves the signal-to-noise ratio (SNR). To further simplify our swept frequency IM-CW LAS system for a future ASCENDS mission, we concurrently investigate advanced IM techniques and develop modulation/demodulation techniques that are simple and have autocorrelation properties that rival those of the pure ML sequence. This includes a quadrature correlator demodulator, PSK modulation techniques, and hybrid-pulse modulation. This results in near-perfect autocorrelation properties and an increase in SNR by up to a factor of 2.

Different modulation techniques are discussed in the following sections appropriate for the hardware shown in Fig. 2. Key results are summarized in the conclusion. Details of the actual instrument, atmospheric models, etc. have been discussed elsewhere [11,18] and will not be repeated here.

2. Measurement Technique

The concept of our IM-CW system for CO₂ IPDA measurements is illustrated in Fig. 1. For the system described in this figure, amplitude modulated on- and offline seed lasers with good spectral properties are optically combined using fiber optics and used to simultaneously seed a single erbium-doped fiber amplifier (EDFA) to increase the transmitted power. A small fraction of the transmitted beam is picked off via an optical tap inside of the EDFA and sent to a reference detector for energy normalization. The backscattered science signal from the on- and offline wavelengths from the surface as well as aerosols and clouds are simultaneously collected with a telescope, optically filtered with a narrowband optical filter, and detected by a single detector. The science and reference signals are amplified, electronically filtered, and then digitized for retrievals of column CO₂.

Postprocessing of the digitized science and reference data allow for discrimination between ground and intermediate scatterers (from clouds, etc.) using the matched-filter technique and also to obtain differential absorption power ratios for inference of CO₂ column amounts as well as range estimates to the scattering targets.

The on- and offline laser modulation signals using the time-shifted approach for this study can be expressed as

$$\Lambda_{\text{off}} = 1 + m\xi(t), \quad \Lambda_{\text{on}} = 1 + m\xi(t - \Delta t).$$  \(1\)

where \(m\) is the modulation index with a value between 0 and 1, and \(\xi(t)\) is the repeating modulation waveform where \(-1 \leq \xi(t) \leq 1\), \(\xi(0) = 0\), and \(\Delta t\) is the time shifting of online modulation waveform related to that of offline. In general, we choose \(\Delta t > 2r_{\text{max}}/c\) to avoid interference between the time-shifted channel and intermediate scatterers, where \(r_{\text{max}}\) is the maximum distance among potential scatterers, including the Earth’s surface, and is also called the unambiguous range, while \(c\) is the speed of light.

Given the modulation waveforms presented in Eq. \((1)\), and using the relation \(P_{\text{off/on}}(t) = \bar{P}_{\text{off/on}}(t)\Lambda_{\text{off/on}}(t)\), the instantaneous received lidar on- and offline received optical powers \([P_{\text{off}}(t)\) and \(P_{\text{on}}(t)\), respectively] from a single scattering target (i.e., the surface) at range \(r\) (which may also vary with time depending on the experiment) are given by

$$P_{\text{off}}(t) = \frac{K}{r^2} \bar{P}_{\text{off}} \exp \left(-2\epsilon \int_0^r \beta(r)\,dr\right) \times \exp(-2r)(1 + m\xi(t - 2r/c))$$

$$P_{\text{on}}(t) = \frac{K}{r^2} \bar{P}_{\text{on}} \exp \left(-2\epsilon \int_0^r \beta(r)\,dr\right) \times \exp(-2r)\exp(-1m\xi(t - 2r/c - \Delta t)).$$  \(2\)

where \(r\) is the total one-way column optical depth for the on- and offline measurement, which is determined by the gas absorption from molecules other than CO₂, and \(r'\) is the one-way column optical depth resulting from CO₂ absorption only, \(\beta\) is the backscatter coefficient (km⁻¹ sr⁻¹), \(\epsilon\) is the extinction to backscatter ratio (sr⁻¹), \(K\) is a constant, where \(\bar{P}_{\text{off}}\) and

---

Fig. 1. Baseline instrument block diagram with either bandpass or low-pass-filtered receiver subsystem. Current baseline uses a bandpass filter.
\( \bar{P}_{on} \) are the average transmitted on- and offline powers averaged over one period of the modulation waveform, respectively.

When multiple targets including the surface exist in the lidar path length, each target would generate similar received lidar signals like those in Eq. (2). The signals from these multiple targets are combined at the detector and converted to an electronic signal \( S(t) \). For an AC coupled receiver subsystem, we have

\[
S(t) = \sum_{k} (C_{1k} m(t-2r_k/c - \Delta t) + C_{2k} m(t-2r_k/c)),
\]

where we sum over all \( k \) scatterers and where,

\[
C_{1k} = \frac{K'}{r^2} \bar{P}_{on}^r \exp \left( -2t \int_0^{r_k} \beta(r')dr' \right) \exp(-2r_k) \exp(-2r_k),
\]

\[
C_{2k} = \frac{K'}{r^2} \bar{P}_{off}^r \exp \left( -2t \int_0^{r_k} \beta(r')dr' \right) \exp(-2r_k),
\]

where \( K' \) is a constant. These \( C_{1k} \) and \( C_{2k} \) returns can be uniquely discriminated from other returns from different scatterers using a matched filter, which involves correlating the received waveform with a reference signal. These are also proportional to the average received optical power (\( \bar{P}_{off,k} \) and \( \bar{P}_{on,k} \)) from the \( k \)th scatterer. For a ground target, solving for \( r'_g \) gives

\[
r'_g = \frac{1}{2} \ln \left( \frac{C_{2g} \bar{P}_{on}}{C_{1g} \bar{P}_{off}} \right) = \frac{1}{2} \ln \left( \frac{\bar{P}_{off,g} \bar{P}_{on}^r}{\bar{P}_{on,g} \bar{P}_{off}^r} \right),
\]

where \( C_{1g} \) and \( C_{2g} \) are the ground returns. Once \( C_{1g} \) and \( C_{2g} \) are determined, the column optical depth for CO\(_2\) can be found using Eq. (5). In general, this is done by cross correlating the reference waveform with the return signal \( S(t) \), i.e., by the matched-filter technique as mentioned before. This results in multiple peaks for the multiple scatterers including the surface. The ground return can be identified through range gating (or maximum range). By choosing the reference modulation waveforms with perfect auto-correlation properties (ML sequence, for instance), channel separation is achieved through the time-shifting approach since each channel correlates at a different apparent range without interference between channels or scatterers. That is, each scatterer results in two peaks in its matched-filter output: one for the offline returns; the other for online returns with the \( \Delta t \) shifted time. The details of this approach are discussed elsewhere [4].

3. Bandpass-Filtered Modulation Waveform Generation and Demodulation

We now present the improved modulation approach for CO\(_2\) IPDA measurements. This new method modulates an IM sine wave carrier by an ML sequence as in Eq. (6). To effectively represent each bit of ML sequence in the transmitted waveforms and allow flexibility in the unambiguous range for a particular code length and sample rate, we oversample the ML sequence by an integer, \( M \), such that each code bit of the ML sequence is represented by \( M \) points. We then use that bit to modulate a sine wave that has exactly \( P \) cycles per code bit. In order to satisfy the Nyquist rule, we must have \( M > 2P \). Let \( z_0(n) \equiv z(n) \) be the original ML sequence and \( Z_0(n) \equiv Z(n) \) be the oversampled ML sequence such that \( Z_0(n) = z(n)(n-1)/M) + 1 \), where \( \text{int}(x) \) is a function that represents the integer part of \( x \). These definitions are used in the following sections wherever PN code modulation is discussed. The following two subsections describe amplitude and PSK IM carrier modulation and demodulation techniques. Amplitude modulation is achieved by modulating the IM carrier with an ML sequence. With PSK modulation, we use the ML sequence to shift the phase of the IM carrier by 180 degrees, which is equivalent to changing the sign of the IM carrier. The PSK modulation technique makes more efficient use of available optical amplifier power and results in better SNR.

A. IM Carrier Amplitude Modulation

In the case of amplitude modulation, the modulation waveform takes on one of two forms:

\[
\xi_a(n) = Z(n) \cos(2\pi(n - 1/2)P/M),
\]

\[
\xi_b(n) = Z(n) \sin(2\pi(n - 1/2)P/M).
\]

The \( \frac{1}{2} \) factor is introduced to make the autocorrelation function symmetric.

Demodulation is accomplished as in Fig. 2 using the reference kernel of \( 2Z(n) - 1 \). This particular reference kernel is necessary to obtain a resulting autocorrelation function with off main lobe values of zero with the amplitude modulation described by Eq. (6). That is to say, with this reference kernel the autocorrelation function is zero except at the main lobe. This is a quadrature demodulator [19] with the low-pass filter replaced by a matched-filter correlation with the oversampled ML sequence. The

---

Fig. 2. Quadrature-filtered correlator used for demodulation. The approach is similar to a standard quadrature demodulator except the low-pass filter section has been replaced by a matched-filter correlation with the reference PN code kernel.
correlations can be calculated using a DFT, which are most efficiently computed with a fast Fourier transform (FFT) and take the form \[R(\text{ref}, \text{data}) = \frac{1}{N} \sum_{m=0}^{N-1} \text{ref}^*(m)\text{data}(m+n)\]

\[= \text{DFT}^{-1}(\text{DFT}^*\text{ref}\text{DFT}\text{data}),\]

where \(*\) denotes the complex conjugate, \(N\) is the length of oversampled ML sequence, \(\text{ref}\) is defined as the reference signal, and \(\text{data}\) is defined as the received data.

As an example of a modulation technique that is suitable for a bandpass-filtered LAS system presented in Fig. 1, we take the special case where \(P = 1\) (one cycle per code bit) and \(M = 8\) (8 samples per code bit). We generate an 8th order (or 255 length) ML sequence using the recurrence relation \(z_{n+8} = z_n \oplus z_{n+2} \oplus z_{n+5} \oplus z_{n+6}\) with the seed \(\{1, 0, 1, 0, 1, 1, 1, 1\}\), where \(\oplus\) is the “exclusive or.” We then oversample this by a factor of 8. Figure 3 shows a comparison between the cosine wave versus sine wave modulation as in Eq. (6). Note that each waveform in this and all subsequent plots are plotted with eight samples per code bit with linear interpolation between points.

The autocorrelation function may be computed according to the block diagram in Fig. 2. The autocorrelation function for the cosine and sine wave cases are shown in Fig. 4, using a 35 sample delay. For larger values of \(P\), the autocorrelation for the sine and cosine modulation cases approaches that of the pure PN code autocorrelation function.

It can be shown that the half-height width of the cosine and sine wave cases is identical to the pure ML-sequence case, though the sine wave case may be better suited for absorption measurements because of the more rounded (flatter) peak. The maximum value for an autocorrelation function with a sharp peak will show sensitivities to under-sampling and therefore may result in a bias. The reason for this is because the return signal will be delayed in such a way that the exact peak location of the autocorrelation function may not always fall on a sample point. As a result, there will be a range-dependent jitter in the height measurement. If the peak is sharp, the jitter will be worse than if it is relatively flat. Another reason is that typically one uses some sort of interpolation to get a better estimate of the peak height, and this is much easier to do if the peak is a continuous function. The sine wave autocorrelation function is less sharp than the pure ML sequence. In both cases, and within the accuracy of numerical calculations, the off main lobe values are essentially zero with the chosen parameters and normalization. The resolution of discriminating different targets is close to the pure ML-sequence case of \(c/2B_s\) [4,5], where \(B_s\) is the code bit rate (number of code bits per second), which is related to the effective modulation bandwidth. It should be noted that the autocorrelation function in the case of more than two cycles per code bit results in an autocorrelation function that looks very similar to the pure ML sequence case [4,5].

B. PSK IM Carrier Modulation

Our particular PSK modulation takes one of two forms:

\[\zeta_a(n) = (2Z(n) - 1)\cos(2\pi(n - 1/2)P/M),\]

\[\zeta_b(n) = (2Z(n) - 1)\sin(2\pi(n - 1/2)P/M).\]

As in the previous case, the \(1/2\) factor is introduced to make the autocorrelation function symmetric. This is very similar to the amplitude modulation case except that \(Z(n)\) has been replaced by \(2Z(n) - 1\). Figure 5 shows a comparison between the cosine wave versus sine wave modulation as in Eq. (8).

Demodulation is accomplished as in Fig. 2, which results in the best autocorrelation properties. Another PSK demodulation technique sometimes used in radar is to correlate the outgoing PSK modulated sine wave signal with the reference PSK modulated sine wave, but that results in less than ideal autocorrelation properties that contain side lobes [20]. In this case, we use the reference kernel \(Z(n)\) with aforementioned demodulator in Fig. 2.

![Figure 3](image-url)

**Fig. 3.** Comparison between cosine (a) and sine wave modulation (b) for amplitude-modulation case as in Eq. (6). Each modulation waveform produces a different autocorrelation function.
reference kernel is necessary to get a resulting autocorrelation function with off main lobe values of zero in the case of PSK. The autocorrelation function is identical to that shown in Fig. 4 in the case where \( P = 1 \) (one cycle per code bit) and \( M = 8 \) (eight samples per code bit). In spite of identical autocorrelation functions, this case has a major advantage. Since we are using \( Z(n) \) as a kernel, this method has better noise immunity. The reason is because \( Z(n) \) is 0 for half of the length, which means we correlate over half the noise compared to the amplitude modulated case. As a result, the SNR should be better by a factor of \( \sqrt{2} \) since the amplitude for each is the same. However, if we were to use \( 2Z(n) - 1 \) for the kernel in the PSK modulation case, the SNR would be even better. The noise increases by a factor of \( \sqrt{2} \), but the amplitude increases by a factor of 2. The net result is that the SNR is a factor of 2 better than the amplitude modulation case. These results are explained in detail in Appendix A. The downside is there is a slight constant bias in the background that is inversely proportional to the bit length of the code \([4]\), which makes it undesirable for the amplitude measurements necessary for \( \text{CO}_2 \) retrievals. This should present little or no problem for range measurements, however. As a result, the \( Z(n) \) kernel version used here is the likely candidate for amplitude measurements because of the superior autocorrelation properties.

4. Low-Pass Filtered Modulation Waveform Generation and Demodulation

A. Hybrid Pulse/CW Modulation

In lidar designed to detect pulses such as those with photon-counting receivers or other types of DC coupled receiver subsystems, other types of modulation are possible. For an alternative implementation, we consider the case where we have DC-coupled electronics, which, for the purpose of this paper, are achieved by using the low-pass-filtered configuration in Fig. 1, combined with an EDFA in the transmitter. An EDFA amplifier utilizes an internal pump laser with a wavelength typically at 980 nm that continually pumps the erbium atoms in the erbium-doped fiber to an excited state. An external signal laser (which in our case has a wavelength of approximately 1570 nm) is used to pass photons through the erbium-doped fiber. Excited erbium atoms then give up some of their energy at the same wavelength and phase as the signal laser, which has the effect of amplifying the laser signal. There are two types of modulation possible with an EDFA amplifier. The first is pump modulation. Modulating the intensity of the pump laser produces a change in gain of the

![Fig. 4. Comparison between autocorrelation function for cosine wave (a) and sine wave case (b) with ML-sequence modulation with one cycle per code bit and eight samples per code bit using a 35 sample delay. Half-height width is identical to the pure ML-sequence case.](image)

![Fig. 5. Comparison between cosine (a) and sine wave modulation (b) for PSK modulation case as in Eq. (8). Each modulation waveform produces a different autocorrelation function.](image)
signal, which modulates the output of the EDFA. With pump modulation there is a low-pass-filter response [21]. This would be an ideal configuration, but it requires an optical amplifier for each channel. However, with signal modulation as has been used in the past [11], which allows one to use the same amplifier for all channels, this system behaves as a lead compensator circuit [21], which is similar to a high-pass filter. This is accomplished by modulating the input laser signal directly. In our case, we modulate this signal by varying the intensity. The step response of such an amplifier configured for signal modulation, which is shown in Fig. 6, does not fully follow the input step function. This is due to an excited state population depletion, which then settles to a steady state.

Even if the transmitter and receiver electronics are DC coupled, the nature of the optical amplifier used here conspires to degrade the autocorrelation properties due to the distortion of the PN code. By modulating the PN code by an offset carrier, one can potentially avoid those sorts of distortions because the optical amplifier responds better to quick changes in signal input than it does to maintaining a constant level as is required by a pure PN code. One solution is to build this as a pulse train into the modulation signal. One could do this with a square wave if one has enough bandwidth. A particularly simple method would be to modify Eq. (6) to

\[
\xi(n) = (1 - \cos(2\pi(n - 1/2)P/M))Z(n) - 1 = 2\sin^2(\pi(n - 1/2)P/M)Z(n) - 1. \quad (9)
\]

An example of this modification is shown in Fig. 7, with the first 100 points of the modulation waveform. For this example, \(P = 1\) and \(M = 8\) are selected. This modification has one main advantage over the previous technique: for the same amplitude, it uses only half the average optical power. The reason is because in order to transmit the previously described modulation it must be offset. This means that a logical zero in that case is half the peak laser power whereas in this case a logical zero is zero power.

Demodulation of this is particularly simple. Instead of demodulating the result using the block diagram in Fig. 2, one simply has to correlate the return signal with \(2Z(n) - 1\). The resulting autocorrelation function is shown in Fig. 8.

As in the previous case, the half-height width is identical to the pure ML-sequence modulation. However, due to the more rounded peak of the autocorrelation function, the height measurement, which is desirable for accurate \(\text{CO}_2\) absorption measurements, should be less sensitive to sampling than with the sharp peak of the pure ML-sequence case or the ML-sequence modulated sine wave case in the previous section. Yet the autocorrelation properties are near perfect as those in previous cases with the chosen parameters and normalization.

5. Bandwidth-Limited Codes

In most LAS implementations it is desirable to limit the bandwidth of the PN code to reduce aliasing and thereby minimize sampling errors in the measurement of the amplitude height of the autocorrelation function. This is particularly important in a space application because there is an upper limit to how fast one can sample. For instance, as of the time of this writing, the fastest one can digitize with space qualified parts, at 16 bit resolution is 10 MHz, which essentially sets the upper limit for the spatial distance between points at about 15 m. Since we need 3 m resolution accuracy to get 1 ppm \(\text{CO}_2\) concentration accuracy, it puts severe limits on what we can do without interpolation, and interpolation doesn’t work well with an aliased signal. Since we can’t sample at any rate we want to avoid aliasing, the only option left is to somehow band limit the signal. This helps with the absorption measurements and makes the peak height less sensitive to small changes in range. The question is how to accomplish this requirement? Many filters can introduce sidelobes to the autocorrelation function, such as a square

Fig. 6. EDFA optical amplifier step response (signal modulation).

Fig. 7. Plot of hybrid sine wave pulse laser modulation signal, using modulation waveform shown in Eq. (9), assuming \(m = 1\). Note that the received signal in a DC coupled system will be proportional to the modulation signal.

Fig. 8. Autocorrelation function for hybrid pulse modulation using a 35 sample delay.
window filter or some other high-order filter. This also potentially includes a nonoptimal bandpass filter, which may already be present so filtering the PN code in the correct way before hand in combination with putting the PN code on a carrier can prevent the problem. Gaussian filters are sometimes used in digital communications to bandwidth limit digital data, minimize channel bleed-over, and carry the maximum amount of information with the least bandwidth [22]. A convenient implementation is to use a periodic version of a Gaussian filter so that we can terminate the filter at the end of a cycle. Since a discrete Fourier transform (DFT) contains the frequency content in the first half of the array and a mirror image in the second half, and the PN code itself is transmitted as a repeating waveform, a cyclic filter convolution is appropriate.

We may construct a periodic Gaussian in the following way with

\[ G(x) = \sum_{n=-\infty}^{\infty} \exp(-\frac{(x-nL)^2}{2\sigma^2}). \quad (10) \]

Here \( L \) is the period, \( \sigma \) is the standard deviation, \( x \) is an as yet undefined variable, and \( L \) is the period. This sum can be done analytically (see Appendix B) as

\[
\sum_{n=-\infty}^{\infty} \exp(-\frac{(x-nL)^2}{2\sigma^2}) \\
= \sqrt{2\pi} \left[ \frac{\sigma}{L} \right] \left[ 1 + 2 \sum_{n=0}^{\infty} q^n \cos(2\pi nx/L) \right] \\
\equiv \sqrt{2\pi} \frac{\sigma}{L} \theta_3(ax/L,q). \quad (11)
\]

where \( \theta_3 \) is the Jacobi theta function of the third kind, and \( q \) is given by

\[ q = \exp(-2\pi^2 \sigma^2/L^2). \quad (12) \]

Theta functions are widely used in the theory of elliptic functions [23] and other applications [24,25]. In order to scale this so that it is 1 at the maxima, we have

\[ W(x,L,q) = \theta_3(ax/L,q)/\theta_3(0,q). \quad (13) \]

To apply this as a filter in the frequency domain, we simply multiply the DFT by the array \( W(n-1,N,q) \) and take the inverse DFT, where \( N \) is the length of the array and \( 1 \leq n \leq N \). Here we have taken \( x = n-1 \) with period \( N \) since the analytic continuation of a DFT has period \( N \) if \( n \) is the array index.

To demonstrate the usefulness of this filter, we first show how to limit the bandwidth of an ML sequence and how the autocorrelation properties and resolution are preserved. We then use that bandwidth-limited sequence in the two modulation examples discussed in the previous sections to show

the effect of bandwidth limiting and preservation of the near-perfect autocorrelation properties for each case with only a minor degradation in resolution. Because of these characteristics, the discussed modulation codes could be easily implemented in current \( \text{CO}_2 \) LAS systems and the concept of future space instruments.

A. Bandwidth-Limited ML Sequence

We start with the consideration of the simple case of a pure ML sequence with length 255 and oversampled by a factor of 8 as in the previous section. We filter this with the theta function filter with \( q = 0.9 \) and \( N = 255 \times 8 = 2040 \). A plot of the unfiltered versus filtered PN code frequency content is shown in Fig. 9.

If the original PN code modulation were applied to a LAS system with a nonoptimal filter, aside from the spread in the autocorrelation function, the spectrum content could be distorted in a way that could introduce possible artifacts such as sidelobes. By bandwidth limiting the codes, it is possible for the modulation bandwidth to fit well within the bandwidth of the bandpass or low-pass receiver filter such that the receiver filter impact is minimal to virtually eliminate the sidelobe issue, which has been confirmed by simulations. These concepts point out practical ways to implement our PN codes to current and future \( \text{CO}_2 \) LAS instrumentation.

The waveform shapes of these unfiltered versus filtered PN codes are shown in Fig. 10. The waveform of the filtered code is obtained from the circular convolution between the cyclic theta function filter and the cyclic PN code.

The results of the correlation between \( Z(n) - 1 \) and the filtered and unfiltered PN codes are shown in Fig. 11. In spite of the strong filtering, the main lobe width of the autocorrelation function is similar, and the peak is more rounded, which is a desirable trait for accurate measurements of \( \text{CO}_2 \) column amounts. Also, if the waveform is sampled at higher than the Nyquist rate, the peak height of the

Fig. 9. Comparison of frequency content of the filtered versus unfiltered PN code. Filtered PN code bandwidth is well under the Nyquist sampling rate demonstrating aliasing is no longer an issue.
The autocorrelation function can be determined more accurately through interpolation because the continuous function on which it is based can always be reconstructed exactly. Shannon’s version of the Nyquist theorem states exactly that [26]. Of course, noise and other errors would tend to make this process less exact. But it does show that bandwidth limiting the waveform to below the Nyquist sample rate is desirable if one is interested in absorption measurements. Because of the required IM carriers for PN modulation in the instrumentation as discussed previously, we further investigate the characteristics of sine wave and hybrid modulations for the filtered PN codes to ensure the usefulness of current discussed techniques for LAS applications.

B. Amplitude Modulated IM Carrier Using Theta Function Filtered PN Code

In this case, we use the filtered PN code in the previous subsection to modulate a sine wave IM carrier as in $\tilde{z}_b$ of Eq. (6) with $P = 2$ (two cycles per code bit) and $M = 8$ (eight samples per code bit). A plot of the resulting modulation is shown in Fig. 12.

The frequency spectra of the modulated sine wave are plotted in Fig. 13. This shows that the resulting frequency content of the filtered modulation is bounded within the Nyquist range with no aliasing.

The autocorrelation functions of these filtered versus unfiltered PN modulations are illustrated in Fig. 14. This shows the modulation produces more robust matched filter output for peak determination by producing a more rounded peak, which is desirable for amplitude measurements in differential absorption estimation because it produces less sampling error. The resolution degradation is very minor.
PSK Modulated IM Carrier Using Theta Function Filtered PN Code

In this case, we use the filtered PN code described earlier to modulate a sine wave IM carrier as in $\xi_k$ of Eq. (8) with $P = 2$ (two cycles per code bit) and $M = 8$ (eight samples per code bit). A plot of the resulting modulation is shown in Fig. 15.

Figure 16 is a plot of the frequency spectra of the modulated sine wave. This shows the resulting frequency content of the filtered modulation is also bounded within the Nyquist range with no aliasing as in the amplitude modulation case.

The plot of the autocorrelation function of the filtered versus unfiltered PN PSK modulation is identical to Fig. 14 of the previous section.

Hybrid Pulse Modulation with Theta Function Filtered ML Sequence

We take the case using Eq. (9) with $P = 1$ (one cycle per code bit) and $M = 8$ (eight samples per code bit) and use the same filtered PN code as in the previous two examples with $q = 0.9$. Figure 17 shows the resulting modulation.

The resulting frequency spectra are shown in Fig. 18. Here we see that the original modulation with the chosen parameters already does a good job at limiting the high frequencies as one might deduce from Fig. 6. This is because there are no sharp transitions as there are with the pure ML sequence. The filtered version limits the spectra even more.

Figure 19 is a plot of the autocorrelation function of the filtered versus unfiltered PN code modulation. Both cases show very good autocorrelation properties.

Discussion and Summary

Table 1 shows a summary of the various modulation techniques covered in this study. We have presented...
several different PN modulation techniques that can be used with different types of hardware: the baseline system in Fig. 1 with a signal-modulated EDFA, the baseline with the low-pass filter configuration, or the baseline system with the low-pass configuration and using separate EDFA's with pump modulation. In the case of DC coupled electronics combined with a laser amplifier with pump modulation, a pure or filtered ML sequence not on an IM carrier is ideal. In the case of the baseline, although the amplitude-modulated sine wave version works well, the PSK modulation case discussed here has a superior SNR. If the electronics is DC coupled, but the laser has a transient response such as the case with a signal modulated EDFA, the hybrid pulse version is best. In addition, a new type of cyclic filter based on the Jacobi theta function was introduced, which is useful for bandwidth-limiting digital PN codes without seriously degrading the resolution and resulting in autocorrelation properties completely without side lobes or other artifacts. The main benefit is to reduce aliasing, which reduces sampling error when measuring the amplitude of the autocorrelation function. It can also minimize the impact of a nonoptimal filter by minimizing side lobes caused by that filter as has been confirmed by numerical simulations with filters used in previous experiments. By using a combination of IM carrier modulation and theta function filtering, one can design a modulation scheme that is flexible enough to work with most hardware while maintaining the integrity of the autocorrelation properties. Internal studies have shown the previous linear-swept-frequency technique, which exhibits a sidelobe level of up to 13 dB, used in our previous test range and flight campaigns, has shown a degradation in accuracy due to sidelobe interference from thin clouds. Although one could potentially develop numerical techniques to deal with this issue, it is much simpler to eliminate the problem by using a modulation and demodulation that doesn’t have sidelobes. This helps meet ASCENDS program requirements by reducing computational burden. Another benefit that has been discussed in a previous publication [4] is the use of the time-shifted approach to separate channels. That approach works with every modulation technique presented here and has a major advantage in that the demodulation process can be done with a single demodulation step for all channels. This further simplifies the computational burden when compared to the linear-swept-frequency technique, which requires a separate demodulation for each channel. Numerical simulations based on a previous model for the system used up to this point [18] show that these techniques could potentially benefit the ASCENDS program with little or no change to the existing baseline hardware and will be explored in a future experimental study.

Appendix A: SNR for AM, PSK, PN, and Hybrid Pulse Modulation

Let’s suppose we have a sine wave signal amplitude modulated by a PN code, $Z$ with amplitude $\sigma$ in the presence of white noise (for example) $\eta$ with variance $\sigma^2$ and 0 mean. For simplicity, we look at the pure electronic signal as pure signal plus white noise. Note that in reality the noise variance will also be a function of the average received optical power if we have detector shot noise, which is normally modeled as white noise. If we assume equal average power, this simple white noise with variance $\sigma^2$ assumption is valid. This is sufficient for determining whether amplitude or PSK modulation has better SNR, and what it is under those assumptions.

1. Noise through the Quadrature Correlator

A. 2Z-1 Kernel Case

We first look at the noise level at the output of the lower and upper correlators in Fig. 2. The variance of the correlation of PN code kernel $2Z-1$ with the lower sine reference signal multiplied by the noise (as per the demodulation process in the lower leg in Fig. 2) is

$$\sigma_{EL}^2 = \frac{1}{N} \langle (2Z - 1)^2 \sin^2(2\pi f t) \eta^2 \rangle$$

$$= \frac{1}{N} \langle \sin^2(2\pi f t) \eta^2 \rangle \approx \frac{1}{N} \frac{1}{2} \sigma^2,$$  \hfill (A1)

so the standard deviation is

$$\sigma_E \approx \frac{1}{\sqrt{2N}} \sigma.$$  \hfill (A2)

For the noise at the output of the upper correlator, we have

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
 & Total BW & PBCF & SNR & SB Offset \\
\hline
PN & Wide & DC & $2\sqrt{2}$ & No \\
FPN & Variable & DC & $<2\sqrt{2}$ & No \\
AM & Wide & Carrier & 1 & No \\
FAM & Variable & Carrier & $<1$ & No \\
PSK & Wide & Carrier & $\sqrt{2}$ & No \\
PSK1 & Variable & Carrier & $<\sqrt{2}$ & No \\
PSK2 & Wide & Carrier & 2 & Yes \\
PSK2 & Variable & Carrier & $<2$ & Yes \\
H pulse & Less wide & DC & $2\sqrt{2}$ & No \\
FH pulse & Variable & DC & $<2\sqrt{2}$ & No \\
\hline
\end{tabular}
\caption{Comparison of Several Different Modulations$^a$}
\footnotesize{$^a$Total bandwidth (total BW), power band center frequency (PBCF), signal-to-noise (SNR) using AM modulation as a reference, and signal background offset (SB offset). modulations include pure PN code (PN), filtered PN code (FPN), amplitude modulation (AM), PSK with Z kernel (PSK1), filtered PSK with Z kernel (PSK1), PSK with 2Z-1 kernel (PSK2), and filtered PSK with 2Z-1 kernel (PSK2).}
\end{table}
\[ \sigma_U^2 = \frac{1}{N} (2N - 1)^2 \cos^2(2\pi ft_j) \eta^2 \]
\[ = \frac{1}{N} (\cos^2(2\pi ft_j) \eta^2) \approx \frac{1}{N} \frac{1}{2} \sigma^2. \quad \text{(A3)} \]

Here the standard deviation is
\[ \sigma_U \approx \frac{1}{\sqrt{2N}} \sigma. \quad \text{(A4)} \]

B. Z Kernel Case
As before, we look at the noise level at the output of the lower and upper correlators in Fig. 2. The variance of the noise at the variance of the correlation of PN code kernel Z with the lower sine reference signal multiplied by the noise (as per the demodulation process in Fig. 2) is
\[ \sigma_L^2 = \frac{1}{N} (Z^2 \sin^2(2\pi ft_j) \eta^2) \approx \frac{1}{N} \frac{1}{4} \sigma_L^2. \quad \text{(A5)} \]

so that
\[ \sigma_L \approx \frac{1}{2\sqrt{N}} \sigma. \quad \text{(A6)} \]

For the upper leg, we have
\[ \sigma_U^2 = \frac{1}{N} (Z^2 \cos^2(2\pi ft_j) \eta^2) \approx \frac{1}{N} \frac{1}{4} \sigma_U^2. \quad \text{(A7)} \]

so that
\[ \sigma_U \approx \frac{1}{2\sqrt{N}} \sigma. \quad \text{(A8)} \]

Generally speaking, the noise at the output of the lower and upper correlators \( (\eta_L \text{ and } \eta_U) \) will have the same variance but will generally be statistically independent because they are both the result of two independent correlations.

2. SNR at the Output of the Quadrature Correlator
For each modulation, we look at the effect of signal and noise through the quadrature correlator. We specifically compare the signal and noise at the correlation peak, since this is how we measure absorption. The output of the quadrature correlator is
\[ S_{\text{tot}} = \sqrt{(S_L + \eta_L)^2 + (S_U + \eta_U)^2}. \quad \text{(A9)} \]

It’s clear that the noise level will be reduced by a significant amount depending on the length of the correlation. We assume the correlation is long enough that the correlation peak of the signal is much bigger than the noise even if we start with a weak signal. Let
\[ S = \sqrt{S_L^2 + S_U^2}. \quad \text{(A10)} \]

Then
\[ S_{\text{tot}} = \sqrt{1 + \frac{2S_L}{S^2} \eta_L^2 + \frac{2S_U}{S^2} \eta_U^2 + \left( \frac{\eta_L}{S} \right)^2 + \left( \frac{\eta_U}{S} \right)^2} \]
\[ \approx \frac{S}{\left( \frac{S_L}{S^2} \eta_L + \frac{S_U}{S^2} \eta_U \right)} \]
\[ = S + \sin(\phi) \eta_L + \cos(\phi) \eta_U. \quad \text{(A11)} \]

The output noise is then
\[ \eta_{\text{out}} \approx \sin(\phi) \eta_L + \cos(\phi) \eta_U. \quad \text{(A12)} \]

The variance of the output noise is then
\[ \sigma_{\text{out}}^2 = \left( \eta_{\text{out}}^2 \right) \approx (\sin^2(\phi) \eta_L^2 + \cos^2(\phi) \eta_U^2)^2 \]
\[ = \sin^2(\phi) \eta_L^2 + 2 \sin(\phi) \cos(\phi) \eta_L \eta_U \]
\[ + \cos^2(\phi) \eta_U^2. \quad \text{(A13)} \]

Since \( \eta_L \) and \( \eta_U \) are statistically independent, \( \langle \eta_L \eta_U \rangle \approx 0 \). From our previous results, \( \langle \eta_L^2 \rangle \equiv \langle \eta_U^2 \rangle \).

Therefore
\[ \sigma_{\text{out}}^2 = (\sin^2(\phi) + \cos^2(\phi)) \eta_L^2 = \sigma_L^2 \equiv \sigma_U^2. \quad \text{(A14)} \]

The signal to noise is
\[ \text{SNR} = \frac{S}{\sigma_{\text{out}}} = \frac{S}{\sigma_L} \equiv \frac{S}{\sigma_U}. \quad \text{(A15)} \]

A. Amplitude Modulation Case with 2Z-1 Kernel
In this case, we use 2Z-1 as the kernel in the quadrature correlator and assume a signal with amplitude \( a \). The signal peak at the lower and upper correlators will be
\[ S_L = \frac{1}{4} a \sin(\phi), \quad S_U = \frac{1}{4} a \cos(\phi), \quad \text{(A16)} \]

where \( \phi \) is the phase angle between the incoming carrier and the sine reference signal. Using the results of Eqs. (A16) and (A2), we have
\[ \text{SNR} \approx \frac{S}{\sigma_L} = \frac{\sqrt{S_L^2 + S_U^2}}{\sigma_L} = \frac{a \sqrt{N}}{2 \sqrt{2} \sigma}. \quad \text{(A17)} \]

B. PSK Modulation Case with Z Kernel
In this case, we use Z as the kernel in the quadrature correlator and assume a signal with amplitude “a.” The signal peak at the lower and upper correlators will be
where $\phi$ is the phase angle between the incoming carrier and the sine reference signal. Using the results of Eqs. (A18) and (A6), we have

$$\text{SNR} \approx \frac{S}{\sigma_L} = \frac{\sqrt{S_L^2 + S_U^2}}{\sigma_L} = \frac{a\sqrt{N}}{2\sigma}. \quad (A19)$$

This is better than the amplitude modulated case by a factor of $\sqrt{2}$.

C. PSK Modulation Case with 2Z-1 Kernel
In this case, we use $Z$ as the kernel in the quadrature correlator and assume a signal with amplitude “$a$.” The signal peak at the lower and upper correlators will be

$$S_L = \frac{1}{2} a \sin(\phi), \quad S_U = \frac{1}{2} a \cos(\phi). \quad (A20)$$

where $\phi$ is the phase angle between the incoming carrier and the sine reference signal. Using the results of Eqs. (A20) and (A2), we have

$$\text{SNR} \approx \frac{S}{\sigma_L} = \frac{\sqrt{S_L^2 + S_U^2}}{\sigma_L} = \frac{a\sqrt{N}}{\sqrt{2\sigma}}. \quad (A21)$$

This is better than the amplitude modulated case by a factor of $2$.

D. SNR for the Pure PN Code and Hybrid Pulse Case
In order to make it comparable sine wave signal case for the same average optical power, we take the output signal of $2aZ$. For the case of the hybrid pulse we use $4a\sin^2(\pi(n - 1/2)/M)Z(n)$ from the results of Eq. (9). The extra factor of $2$ comes from equalizing the average optical power for each (assuming the laser amplifier has the headroom to make this possible). Using $2Z-1$ as the kernel the correlation peak for either modulation will be

$$S = \frac{N + 1}{N} a \approx a. \quad (A22)$$

The noise in this case will not have $0$ mean because transmitting a pure PN code usually requires a DC coupled receiver or photon counting system. Because the detector will output one-sided shot noise, there will always be a bias offset in the noise. The variance of the correlated noise is

$$\sigma^2_{\text{out}} = \frac{1}{N} (2Z - 1)^2 (\eta - \bar{\eta})^2 = \frac{1}{N} (\eta - \bar{\eta})^2 = \frac{1}{N} \sigma^2. \quad (A23)$$

The signal to noise is then

$$\text{SNR} = \frac{S}{\sigma_{\text{out}}} \approx \frac{a\sqrt{N}}{\sigma}. \quad (A24)$$

This is better than the amplitude-modulation case by a factor of $2\sqrt{2}$. However, unlike the other cases there is a bias offset in the noise. The bias of the correlated noise is

$$\bar{\eta}_{\text{out}} = \langle (2Z - 1)\eta \rangle = \frac{1}{N_c} \eta, \quad (A25)$$

where $N_c$ is the bit length (uninterpolated code length) of the PN code. On the other hand, this is small compared to the standard deviation. There is also a small bias term for the other carrier-based modulation cases even if the original noise isn’t biased. The small terms we threw away in Eq. (A11) represent a small bias since those noise terms are squared. This is a consequence of the quadrature demodulation, which makes the noise plus bias always positive. One possible solution to that would be to demodulate the signal with a phase-corrected reference carrier and use only one leg of the demodulator. The downside is this may not be simple to do in flight conditions where the range and phase are constantly changing.

Appendix B: Periodic Gaussian as a Jacobi Theta Function
The sum represented by Eq. (10) may be performed analytically using the Poisson summation formula [27], which states

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{k=-\infty}^{\infty} \hat{f}(k), \quad (B1)$$

where

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(n) \exp(-2\pi i kn)dn, \quad (B2)$$

is the Fourier transform of $f$. In the case of Eq. (10), this becomes

$$\hat{f}(k) = \int_{-\infty}^{\infty} \exp((x - nL)^2/2\sigma^2) \exp(-2\pi i kn)dn$$

$$= \sqrt{2\pi} \sigma \exp(-2\pi^2 \sigma^2 k^2/L^2) \exp(-2\pi kx/L), \quad (B3)$$

so that
\[ F(x) = \sqrt{2\pi} \int_{L}^{\infty} \sum_{k=-\infty}^{\infty} \exp(-2\pi^2 \sigma^2 k^2 / L^2) \times \exp(-2\pi kx / L) \]
\[ = \sqrt{2\pi} \int_{L}^{\infty} \left[ 1 + 2 \sum_{k=1}^{\infty} q^2 \cos(2\pi k x / L) \right] \]
\[ = \sqrt{2\pi} \int_{L}^{\infty} \psi(x/L, q). \quad (B4) \]

References


