Design and Analysis of a First-Generation Optical Pulse-Position Modulation Receiver

V. Vilnrotter,[1] A. Biswas,[1] W. Farr,[1] D. Fort,[2] and E. Sigman[2]

In this article, the requirements and design of a prototype pulse-position modulation (PPM) receiver are considered. This article focuses on a prototype optical receiver concept that will be used to demonstrate and validate optical reception under conditions representative of deep-space communications, where Earth and spacecraft dynamics must be taken into account. The system will rely on PPM, which is a viable high-data-rate format that enables the use of high-peak-power lasers for interplanetary communications. The design incorporates an amplifier front end that provides the required dynamic range, along with a capability to evaluate various detectors, including photo-multiplier tubes and avalanche photodiode detectors. The digital section is designed to achieve and maintain PPM slot and frame synchronization, detect PPM symbols, estimate range, and eventually interface with decoding equipment. These receiver functions will be implemented using field-programmable gate arrays. This prototype receiver will provide a test bed for validating optical communications concepts in the laboratory, over mountain-to-mountain terrestrial links, and eventually with retroreflected signals from Earth-orbiting satellites.

I. Introduction

Ground reception technology for deep-space optical communications is being developed at JPL in order to service the expanding set of future NASA missions. The increased data rates with a reduction in payload mass and volume are drivers for an optical approach to retrieving information from interplanetary spacecraft. In order to receive measurable laser signals from distant spacecraft, high peak powers must be relied upon [1]. High peak powers are achieved by Q-switched lasers [2] that can be operated at 10- to 100-kHz repetition rates. Further increases in laser repetition frequencies (data rates) without compromising peak power can be pursued by resorting to cavity dumping [3] and other novel wave-guiding schemes. However, in the near term (5 to 7 years), Q-switched lasers are the most likely candidates for the first generation of flight lasers. Given the limited repetition rates in order to achieve required peak powers, pulse-position modulation (PPM) is the technique currently being considered for transmitting data. In an operational scenario, PPM laser pulses transmitted from deep space will be incident upon large-area (∼10-m-diameter) collecting apertures and then brought to focus upon large-area (∼2- to 3-mm-diameter) optical detectors. These detectors, preceded by appropriate narrow (0.1- to 0.2-nm) optical

[1] Communications Systems and Research Section.

[2] Tracking Systems and Applications Section.

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

bandpass filters and followed by low-noise amplifiers, will serve as the PPM receiver front end. For large collecting apertures, the irradiance fluctuations (scintillation) induced by atmospheric turbulence will tend to be averaged out [4,5]. However, the large collecting apertures are not currently available and, in the intervening technology development phase, PPM receivers will be tested using 0.6- to 1-m-diameter telescopes with laser beams propagating near horizontal atmospheric paths. Experience [6] has shown that the aperture averaged signal through these telescopes can exhibit 8- to 10-dB fades from the mean irradiance levels; therefore, adequate dynamic range must be designed into the front end of the PPM receivers. The capability of performing ranging on the received laser pulses is also an important secondary goal.

In this article, the requirements and design of a PPM receiver planned for development over the next year will be elaborated. The development is driven by the need to demonstrate and validate PPM reception using a number of different optical detectors that display a range of sensitivities.

The planned tests using the PPM receiver include

(1) End-to-end laboratory testing so that the photons per bit versus bit-error rate (BER) can be measured under varying simulated background light levels. The dynamic range over which the receiver can operate will be characterized. Testing with variable laser repetition rates in order to simulate the pulse time of arrivals while tracking a satellite will be performed in order to validate synchronization. The overall performance must be understood in terms of losses contributed from different functional units of the receiver.

(2) Initial field demonstrations relying on static links where laser beams will be transmitted across a near horizontal path and received by a 0.6- to 1-m-diameter telescope. Alternatively, the laser pulses could be transmitted through the receiving telescope and retroreflected back prior to incidence on the receiver.

(3) Retroreflected links to satellites where short duration (∼15-ms) bursts of pulses are transmitted and retroreflected in order to test receiver functions while tracking satellites.

By first assembling the prototype PPM receiver and then performing some of the tests described above, risks involved in implementing optical communications as a means of retrieving data from spacecraft at planetary distances will be minimized. This will also provide a theoretically sound and experimentally validated basis for more advanced designs of future receivers required for providing a robust high-data-rate service for NASA's future missions. In the following sections, we will describe the external effects that impact receiver design, provide an overview of the optical PPM format, characterize the optical detection process, describe the functions to be performed by the optical PPM receiver, and conclude with hardware considerations.

II. Link Effects due to Spacecraft Dynamics

A. The Effects of Spacecraft Dynamics on Transmitted Laser Pulses

When attempting to track a fast-moving object such as a spacecraft, particularly one with rapidly changing velocity, the impact of spacecraft dynamics on the PPM pulse train must be taken into account. Consider the communications link depicted in Fig. 1, showing a spacecraft in relative motion with respect to the receiver. The velocity vector of the spacecraft can be decomposed into radial and tangential components with respect to the receiver, which is assumed to be stationary on the ground. The tangential component gives rise to angular motion that must be tracked by the receiving telescope, while the radial component gives rise to temporal dynamics. We assume here that the telescope tracks the tangential component, and concentrate only on temporal effects generated by the radial component.

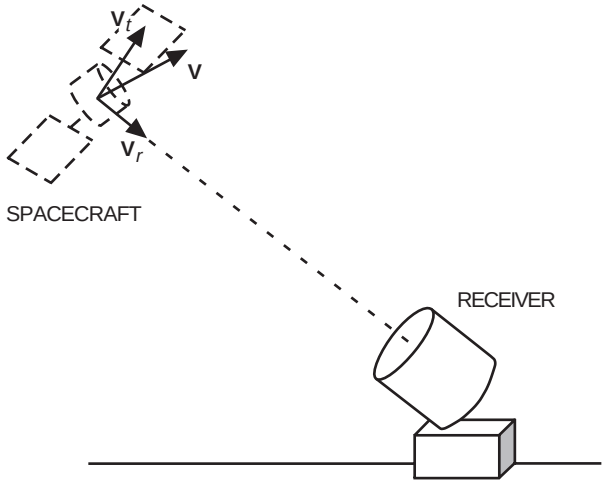

Fig. 1. Optical space communications link geometry.

Suppose that the radial velocity $\mathbf{v}_r$ is towards the receiver, as in Fig. 1, with magnitude v_r (much less than the velocity of light, c, so that relativistic effects can be ignored) and that the transmitter onboard the spacecraft is sending a train of pulses towards the receiver. The PPM signal format is described in greater detail in Subsection III.A. For this analysis, it is easiest to assume that the pulse is always in the first of $M + N$ slots, which includes M information and N "dead-time" slots (see Fig. 2), where each slot is τ seconds in duration. The clocks at the ground receiver and onboard the spacecraft are assumed to be synchronized. Although we are ignoring relativistic effects, we can state this condition precisely by assuming that the two clocks were synchronized at some earlier time when the relative radial velocity component was zero and that the clocks are stable enough to keep any relative drift between them negligibly small. There are three cases to consider: zero radial velocity, positive radial velocity (taken to be towards the receiver), and negative radial velocity (away from the receiver).

1. Zero Radial Velocity. Let's set the time to zero at the transmitter, $t = 0$, when the first transmitted pulse leaves the spacecraft. If the distance to the receiver is R meters, the first pulse arrives at the receiver at a time $\Delta = R/c$ seconds later. If the time of arrival was known at the receiver (through the use of accurate predicts, for example), then we can assume that this delay was taken into account and the receiver's clock adjusted so that the first pulse arrives exactly at the center of the first slot of a frame, which we designate as the first frame. Since there is no relative radial velocity between the receiver and transmitter, all subsequent pulses arrive exactly in the middle of the first slot of each subsequent frame, merely delayed by a constant propagation time.

2. Positive Radial Velocity. Next consider the case where the spacecraft travels towards the receiver with constant radial velocity v_r. After traversing a distance R, the first pulse arrives at the spacecraft $\Delta = R/c$ seconds later. The second pulse is launched a time T seconds after the first, during which time the spacecraft has moved $v_r T$ meters closer to the receiver. The second pulse covers the lesser distance $R - v_r T$ to get to the receiver, and therefore arrives $T + (R - v_r T)/c$ seconds later. The third pulse is sent $2T$ seconds after the first and arrives $2T + (R - 2v_r T)/c$ seconds after the first, and so on. Ignoring the common delay R/c, the arrival times of the pulses at the receiver are seen to be related as $0, T(1 - [v_r/c]), 2T(1 - [v_r/c]), \cdots, (N-1)T(1 - [v_r/c])$. We can see, therefore, that the effect of positive relative radial velocity is to reduce the frame duration, effectively multiplying the frame, slot, and subslot durations, by the compression factor $(1 - [v_r/c])$. Therefore, to stay synchronized with the incoming pulses, the receiver has to reduce the frame duration or, equivalently, increase the frame repetition rate by the inverse of this factor. Since the frames are built up from subslots defined by the clock frequency, this can be accomplished most directly by increasing the receiver clock frequency from f_c to $f_c/(1 - [v_r/c]) \simeq f_c(1 + [v_r/c])$ in order to compress the timing by the required amount.

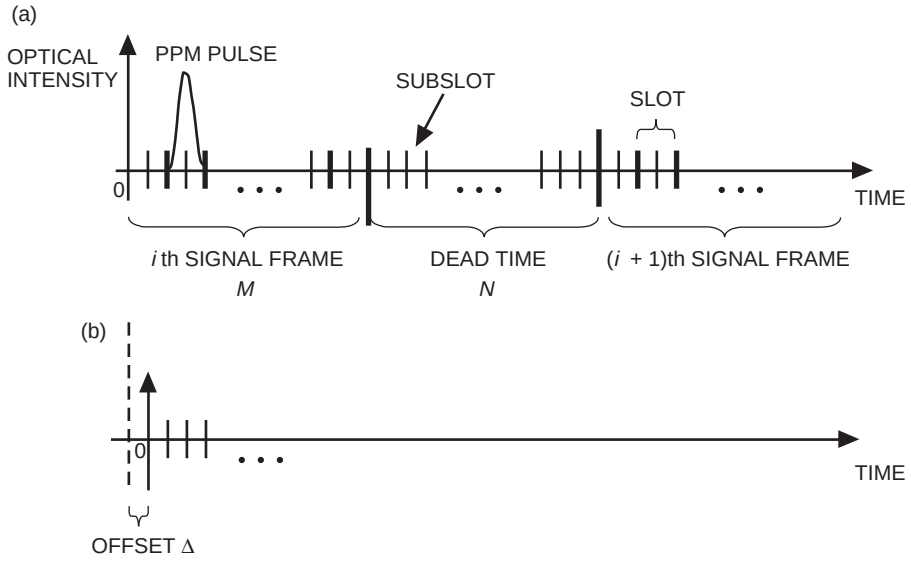

Fig. 2. PPM signal format: (a) the received PPM signal format and (b) its relation to the receiver clock before slot synchronization.

3. Negative Radial Velocity. The analysis for negative radial velocity is analogous to the positive radial velocity case, except that subslot, slot, and frame durations now have to be expanded by a factor of $(1+[v_r/c])$ to account for the increasing relative distance as the spacecraft moves away from the receiver. The frequency of the receiver clock is reduced in this case by applying the factor $1/(1+[v_r/c]) \cong 1-(v_r/c)$ to the nominal operating frequency.

B. The Effects of Spacecraft Dynamics on Retroreflected Pulses

Another mode of operation for experimentally demonstrating an optical PPM link is by means of retroreflectors on orbiting satellites. In this mode, the transmitter is located near the receiver and launches PPM pulses towards the spacecraft, using a somewhat broadened beam to mask atmospheric random steering effects and taking into account the "point ahead" required to hit the spacecraft. This scenario is illustrated in Fig. 3.

The point-ahead angle is a function of the tangential velocity component only and is defined relative to the direction of the received (retroreflected) light in the plane of the trajectory. The magnitude of the point-ahead angle is approximately $\theta_{pa} \cong 2v_t/c$, which is on the order of 20 μrad for tangential velocities of 3000 m/s (11,000 km/h). Assuming that the point-ahead angle has been taken into account by the telescope tracking system, the intensity of the reflected component at the receiver and the Doppler-induced delay on the reflected pulses remain to be determined.

Consider the link geometry of Fig. 4. The ground station transmits an optical beam towards the spacecraft, taking into account point ahead. If the uplink is a Gaussian beam, the magnitude of the electric field component at the points (x, y) in a plane perpendicular to the beam axis, a distance z from the transmitting laser as depicted in Fig. 2, is given by an expression of the form [7]

$$E(x,y,z) = E_0 \frac{W_0}{W(z)} \exp\left[-\frac{x^2+y^2}{W^2(z)}\right] \exp\left[-jkz - jk\frac{x^2+y^2}{2R(z)} + j\varsigma(z)\right] \tag{1}$$

where $E_0 = E(0,0,0)$, λ is the wavelength, $k = 2\pi/\lambda$, W_0 is the effective radius of the beam waist, $z_0 = \pi W_0^2/\lambda$, $W(z) = W_0\sqrt{1+(z/z_0)^2}$, $R(z) = z\left[1+(z_0/z)^2\right]$, and $\varsigma(z) = \tan^{-1}(z/z_0)$. The diver-

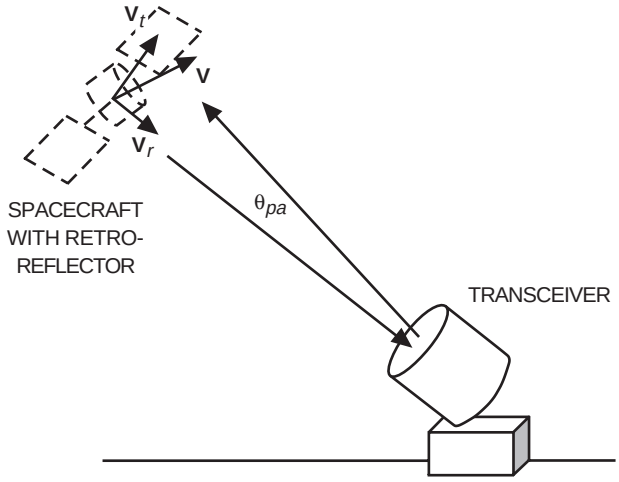

Fig. 3. Optical communications experiment using a retroreflector.

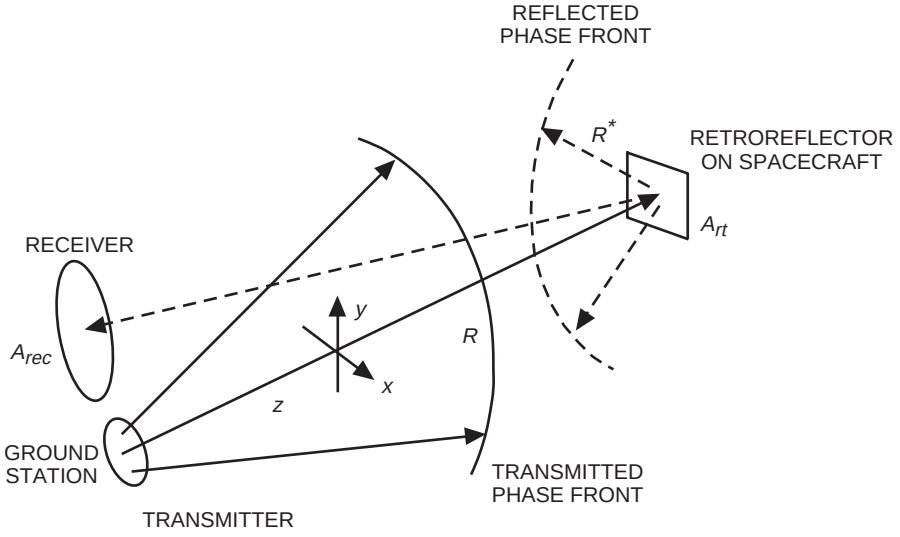

Fig. 4. Transmitter, receiver, and retroreflector geometry for determining the retroreflected signal power and Doppler effects induced on the pulses.

gence of the Gaussian beam (half angle) can be expressed in terms of the beam waist diameter as $\theta_0 = (2/\pi)(\lambda/2W_0)$. The optical intensity anywhere in the beam can be written as

$$I(x, y, z) = I_0 \left[\frac{W_0}{W(z)}\right]^2 \exp\left[-\frac{2(x^2 + y^2)}{W^2(z)}\right] \tag{2}$$

where $I_0 = |E_0^2|$. It can be shown that the total power in the Gaussian beam is $P = (1/2)I_0(\pi W_0^2)$; hence, Eq. (2) can be rewritten as

$$I(x, y, z) = P\frac{2}{\pi W^2(z)} \exp\left[-\frac{2(x^2 + y^2)}{W^2(z)}\right] \tag{3}$$

Therefore, if the transmitter generates a Gaussian beam with total power P, the total power collected by an aperture of A_{rt} square meters on axis at a distance R is $I(0,0,R)A_{rt} = 2PA_{rt}/\pi W^2(R)$. Using the definition of Rayleigh distance, defined as $z_0 = \pi W_0^2/\lambda$, and substituting in $W^2(R) = W_0^2\left[1 + (\lambda^2 R^2/\pi^2 W_0^4)\right]$, we can see that for $\lambda R >> \pi W_0^2$ (for a 0.25-m spot radius and 1-μm wavelength, this would occur at a range $R >> 3\times 10^5$ m, or 300 km) this expression reduces to $W^2(R) \cong \lambda^2 R^2/\pi^2 W_0^2$, and the power collected by the retroreflector becomes

$$I(0,0,R)A_{rt} \cong \frac{2PA_{rt}\pi W_o^2}{\lambda^2 R^2} \tag{4}$$

This is also the power reflected back towards the transmitter by the retroreflector, assuming its aperture is normal to the received field direction.

The intensity of the retroreflected field at the ground-based transmitter can be determined by starting with the power captured by the retroreflector and calculating the intensity of the associated plane wave propagating through its aperture. If the aperture is circular, then the intensity in the far field is proportional to the two-dimensional Fourier transform of the aperture:

$$I(x,y,z) = I_0\left[\frac{2J_1\left(\dfrac{\pi D_{rt}\sqrt{x^2+y^2}}{\lambda z^*}\right)}{\dfrac{\pi D_{rt}\sqrt{x^2+y^2}}{\lambda z^*}}\right]^2 \tag{5}$$

where D_{rt} is the diameter of the retroreflector and z^* is the distance traveled by the reflected field. The divergence of the diffracted beam in the far field (half angle) is known to be $\theta_0 = 1.22(\lambda/D_{rt})$. The on-axis intensity a distance R^* from the retroreflector is simply

$$I_0 = I_{rt}\left(\frac{\pi D_{rt}^2}{4\lambda R^*}\right)^2 \tag{6}$$

where the intensity of the reflected field at the retroreflector is $I_{rt} \equiv 2P\pi W_o^2/\lambda^2 R^2$. Finally, the power of the reflected field passing through a receiving aperture of area A_{rec} a distance R^* from the retroreflector can be expressed as

$$P_{rec} = \frac{2PA_{rec}\pi W_0^2}{\lambda^2 R^2}\left(\frac{A_{rt}}{\lambda R^*}\right)^2 \tag{7}$$

For the special case where the receiver is located near the transmitter, so that $R^* = R$, the received power becomes

$$P_{rec} = \frac{2PA_{rec}A_w A_{rt}^2}{\lambda^4 R^4} \tag{8}$$

where we have defined the effective area of the Gaussian beam waist as $A_w \equiv \pi W_0^2$. Note that Eq. (8) is dimensionally correct and exhibits the "inverse R^4" dependence characteristic of retroreflected fields.

The effect of radial spacecraft velocity on the retroreflected pulses can be determined in a manner analogous to the transmitted pulse case described in Subsection II.A. In the reference frame of the

spacecraft, the relative radial velocity between the ground-based transmitter and the spacecraft causes an effective compression in the arrival times of the pulses of exactly the same magnitude as derived above for the case of pulses transmitted from a moving spacecraft towards the receiver.

We can view the retroreflected pulse train as being transmitted by the spacecraft, but with pulse spacing of T^* seconds instead of T seconds, where $T^* = T(1 - [v_r/c])$, so that the pulses are transmitted at times $0, T^*, 2T^*, \cdots$, in the time frame of the spacecraft. Suppose now the distance to the spacecraft when the first pulse arrives at the receiver on the ground is R^* meters, which was covered by the first pulse in R^*/c seconds. The second pulse is fired T^* seconds after the first pulse, during which time the spacecraft has moved a distance $v_r T^*$ towards the receiver. The second pulse covers a shorter distance to the receiver, namely $R^* - v_r T^*$ meters, which takes $(R^* - v_r T^*)/c$ seconds, so the second pulse arrives at the receiver $T^* - v_r T^*/c$ after the first pulse. Expressing this in terms of the time intervals at which the pulses were originally launched, the second pulse arrives $T - (v_r/c)T - (v_r/c)T + (v_r/c)^2 T \cong T(1 - [2v_r/c])$ seconds later, where we have ignored the quadratic term since it is negligible for practical cases of interest. Similarly, the third pulse arrives $2T(1 - [2v_r/c])$ seconds after the first, whereas the Nth pulse arrives $(N-1)T(1 - [2v_r/c])$ seconds after the first pulse. Therefore, the Doppler compression on retroreflected pulses is exactly twice that experienced by pulses originally transmitted from a spacecraft with T-second spacing, suggesting that experiments and demonstrations using retroreflected pulses could provide useful information about the performance of direct laser communications links.

III. Optical PPM Receiver Design

A block diagram of the optical receiver is shown in Fig. 5, where the major functional blocks are identified and their interconnection specified. The front end consists of a detector assembly, which can be either a conventional avalanche photodiode detector (APD) or a photomultiplier tube (PMT), selectable by the user. The output of each detector is connected to conditioning circuits via a user-selectable switch for amplifying the detected signal and establishing a suitably high signal-to-noise ratio (SNR). The analog signals are filtered by low-pass analog filters, digitized to 8 bits at a rate determined by the system clock (selected to be 200-MHz nominal, yielding 5-ns samples), and serve as input to the slot synchronization, PPM detection, and ranging assemblies.

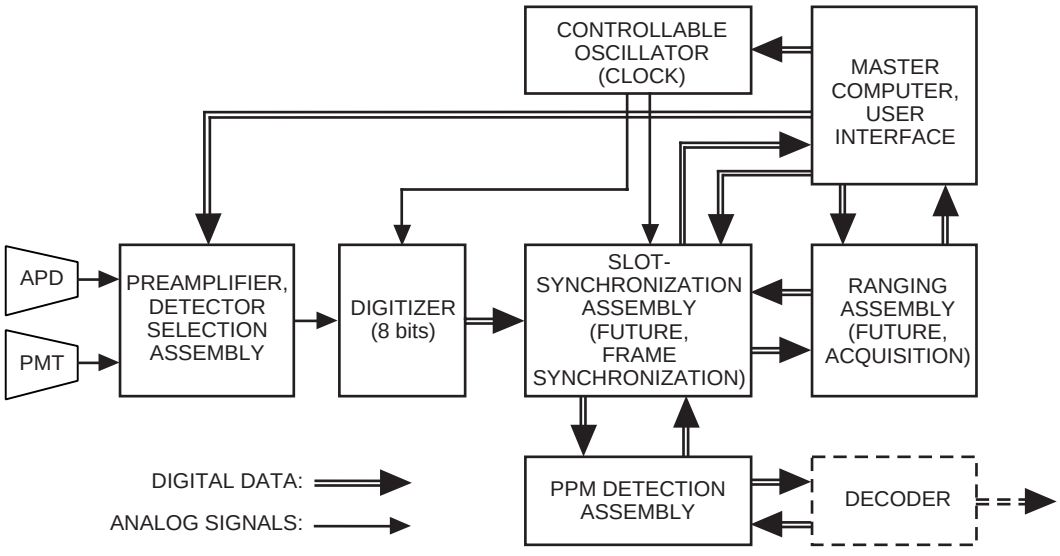

Fig. 5. Optical receiver functional block diagram.

The output of APDs used for communications has been modeled and evaluated in previous articles [8,9]. These detectors do not have sufficient gain (typically 100 to 1000) to fully overcome thermal noise in the electronic circuits and in addition suffer from excess noise due to a random gain mechanism, but nevertheless they provide greater sensitivity than positive-intrinsic-negative (PIN) photodiodes in optical direct detection applications. Photomultiplier tubes, on the other hand, generally provide gains of several million, thereby easily overcoming thermal noise and enabling the detection of individual photons. Randomness in the gain can be mitigated by setting a threshold high enough to discriminate against thermal noise, but much lower than the average height of photon-generated pulses; in this way, nearly every absorbed photon can be counted. However, this requires very high electrical and processing bandwidths, which cannot be implemented in the current design. Therefore, in this receiver, the PMT outputs will also be integrated and sampled, much as the APD outputs. The combination of low-pass filtering followed by sampling at the clock rate is designed to approximate an integrate-and-dump filter by means of simple time-invariant analog circuits; however, for simplicity, the outputs will be treated as true integrations in the following analysis. The integrated pulse stream can be modeled in terms of a count accumulator function, $N(t)$, as shown in Fig. 6.

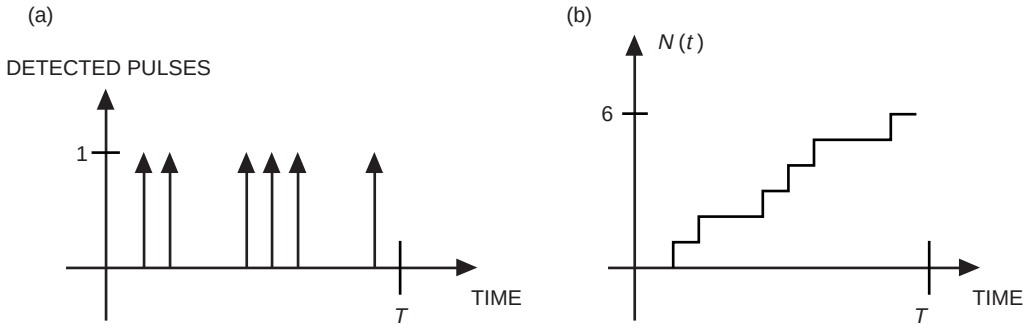

Fig. 6. The integrated pulse stream: (a) detected PMT pulses and (b) the associated count accumulator function *N* (*t*).

A. PPM Signal Format

A description of the received PPM signal generated by a Q-switched laser, and its temporal relation to the receiver clock prior to synchronization, is illustrated in Fig. 2. The subslot is the smallest time interval over which integration and sampling can be carried out, in this case corresponding to 5 ns (determined by digital hardware limits and computational constraints). In order to establish slot synchronization, a minimum of two subslots will be needed, but there is no conceptual upper limit on the number of subslots per slot provided that number is even. It is assumed that the transmitter clock also runs at 200 MHz, generating 5-ns clock edges for controlling the firing of the transmitter laser. The PPM format consists of mapping each sequence of L data bits into one of $M = 2^L$ slot locations according to a predetermined look-up table, and firing the laser so as to place the optical pulse in the correct slot relative to the previous pulse. After firing, there is a recovery time during which the laser cannot be fired, giving rise to a dead time following each pulse; this interval is approximately 15 μs for current laboratory lasers, but can be reduced to 5 to 10 μs with careful design. With no loss in generality, we assume that there is a dead time of exactly N slots following the M signal slots, for a total of $M + N$ frame slots.

The transmitted optical pulse propagates to the receiver over free space at a rate of approximately 3 ns per meter. In the laboratory, the distance between the transmitter and the receiver is typically on the order of meters, whereas in field tests it can be tens of kilometers or more. The finite propagation time gives rise to a delay that is generally not known with great accuracy and, in addition, may be varying with time. This relative delay is represented in Fig. 2 by the symbol Δ and can be thought of as a shift between the origin of the time axes at the two locations; the main purpose of slot synchronization is to establish a one-to-one correspondence between the receiver clock and the transmitter clock following

pulse propagation through the physical channel, so that measurements can be carried out at the receiver over time intervals that contain the entire laser pulse after it arrives. In addition, the slot boundary corresponding to the start of the PPM symbol has to be determined before PPM symbol detection or decoding can take place.

B. Generation of Subslot Observables

To simplify the analysis, we assume a photon-counting model for the PMT output, keeping in mind that our implementation will likely suffer additional losses. Because the photon-counting model is easiest to understand, we adopt it for analysis when the PMT is used to detect the signal. When the APD is used, the models developed in [8] will be employed.

First consider the generation of subslot observables at the output of a PMT. Pulsed optical fields produced by Q-switched lasers can be modeled as Poisson-distributed pulses at the output of a photon-counting detector: the greater the intensity, the greater the average number of pulses (or detected photons). Since the pulses generated by a physical PMT are random in amplitude, these counts can be obtained in practice by setting a threshold to distinguish the pulses from circuit noise, effectively yielding an integer count each time the pulse exceeds threshold. This model can be extended to the case of multimode background radiation entering the receiver along with the signal, which also gives rise to Poisson-distributed counts with average intensity proportional to the sum of signal and background intensities [9].

The processing needed to optimally detect the laser pulse, and hence decode the PPM symbols with the smallest possible probability of error, is to count the number of detected photons in every slot and select the slot with the greatest photo count as the slot most likely to contain the laser pulse [10]. Mathematically, this can be expressed in terms of the slot count accumulator function $N_i(t)$ for the ith T-second slot, as described earlier, which is initially set to zero and increments by one each time a photon is detected, but is reset to zero at the end of each slot.

Since the photo counts obtained from a subslot are Poisson distributed, and the average number of photons over the entire slot is equal to the sum of the average subslot counts, it follows that counting the number of photons in each subslot and adding the results is equivalent to obtaining the full slot count; hence there is no loss in detection optimality by processing subslot observables instead of the slot observables. Defining the subslot count accumulator function analogously to the slot count accumulator function, except now resetting it at the subslot boundary instead of the slot boundary, the subslot counts can be represented by the random integers N_{11}, N_{12}, N_{21}, and N_{22}, where $N_{11} \equiv N_1(T/2), N_{12} \equiv N_1(T)$, and so on, for the case of two subslots per slot. Since these observables are integers, no further quantization is necessary with this ideal model.

For the case of APD detectors, the output statistics are well modeled by Webb plus Gaussian statistics, which account for both the random gain of the APD and the effects of additive circuit noise [8]. It has been shown in [10] that for this case the optimum strategy for detecting PPM pulses is to integrate the APD output over each slot and select according to the largest observable among all valid signal slots. Since the subslot observables can be modeled as Webb plus Gaussian random variables, it follows that the slot variables are also Webb plus Gaussian. Therefore, we can again rely on subslot observables to build up the detection statistics, with no loss in optimality.

With $x(t)$ denoting the output of the APD, the statistics for the case of two subslots per slot are denoted by the continuous random variables $X_{11}, X_{12}, X_{21}, X_{22}$, where now $X_{11} \equiv \int_0^{T/2} x(t)dt$, and so on. The values of these continuous variables are first quantized to 8 bits, then supplied to the digital signal processing assemblies. Having defined the subslot observables for both photon-counting PMT and APD detectors, we now proceed to describe the processing performed by the slot synchronization, PPM detection, and ranging subassemblies.

IV. PPM Detection Assembly

The block diagram of Fig. 5 indicates that the slot synchronization assembly receives the quantized observables first; however, it is actually easier to describe the slot synchronization algorithm employed in this receiver by first examining the PPM detection assembly. We begin by defining the optimum detection statistic for the ith slot as $D_i = Y_{i1} + Y_{i2}$, where $Y_{ij} = N_{ij}$ for the photon counting PMT and $Y_{ij} = X_{ij}$ for the APD detector. As a simple example, consider the photon-counting detection of a laser pulse in the absence of interfering background radiation.

A. Photon-Counting Detection in the Absence of Background

The probability that the detection statistic D takes on the value N, $N = 0, 1, 2, \cdots$, is accurately described by the Poisson distribution

$$p_D(N) = \frac{\bar{N}^N}{N!} \exp(-\bar{N}) \tag{9}$$

where $\bar{N}$ represents the average number of signal photons per slot (or laser pulse). If one or more signal photons are observed, the receiver decodes the PPM symbol correctly. Even if no photons whatsoever are absorbed by the PMT detector, the optimum receiver makes a random choice among the M signal slots, choosing the correct slot one out of M times on the average. The probability of correct detection is, therefore,

$$P(C) = \big(1 - \exp(-\bar{N})\big) + \frac{\exp(-\bar{N})}{M} = 1 - \frac{M-1}{M} \exp(-\bar{N}) \tag{10}$$

which yields the average probability of symbol error

$$P(SE) = 1 - P(C) = \frac{M-1}{M} \exp(-\bar{N}) \tag{11}$$

Note that, in the absence of background light, the probability of committing an error is an exponentially decreasing function of the average signal energy (measured in units of photons). Therefore, average error probabilities of 0.01 to 0.001 can be achieved with very small received signal energy, namely with an average of 4 to 7 photons, even for large values of M. Therefore, under normal operating conditions, it is reasonable to expect to observe several photons in the subslots closest to the arrival time of the laser pulse, even before slot synchronization is established. The subslot containing the greatest number of observed photons can be taken as the starting point for establishing slot synchronization.

B. Photon-Counting Detection in the Presence of Background

When significant amounts of background light enter the receiver along with the signal, it is possible for the the receiver to commit an error even if one or more signal photons are detected, because a noise slot may occasionally produce a greater count than the signal slot. The probability of correct detection for maximum-likelihood detection of PPM symbols in the presence of noise has been derived in [9] and shown to be

$$P_M(C) = \left\{ \sum_{r=0}^{M-1} \left(\frac{1}{r+1} \right) \binom{M-1}{r} \sum_{k=1}^{\infty} \frac{\left(\bar{N}_s + \bar{N}_b\right)^k}{k!} e^{-\left(\bar{N}_s + \bar{N}_b\right)} \left[\frac{\bar{N}_b^k}{k!} e^{-\bar{N}_b} \right]^r \left[\sum_{j=0}^{k-1} \frac{\bar{N}_b^j}{j!} e^{-\bar{N}_b} \right]^{M-1-r} \right\}$$

$$+ M^{-1} e^{-\left(\bar{N}_s + M\bar{N}_b\right)} \tag{12}$$

where $\bar{N}_s$ and $\bar{N}_b$ are the average counts due to signal and background, respectively, and r represents the number of noise-slot counts equal to the maximal signal-slot count in the PPM symbol. The symbol-error probability is given by $P(SE) = 1 - P(C)$. In the case of equalities, the maximum-likelihood receiver makes an arbitrary choice among the slots with the biggest counts, of which there are $(r + 1)$ including the signal slot, and guesses correctly by pure chance $1/(r+1)$ times on the average. The probability that the count random variable equals k in the signal slot is $p_q(k|H_q) = [(\bar{N}_s + \bar{N}_b)^k/k!]e^{-(\bar{N}_s+\bar{N}_b)}$, and the probability that it equals k among one of the noise slots is $p_q(k|H_q) = (\bar{N}_b^k/k!)e^{-\bar{N}_b}$. Detailed descriptions of PPM decoder performance under various background conditions can be found in [9].

C. APD Detection of PPM Signals

The average gain of an APD detector is not sufficient to overcome thermal noise in the transimpedance preamplifier; therefore, it cannot be used to attain shot-noise-limited performance. As shown in [8], the APD output current is well approximated by the Webb density, to which we must also add a Gaussian-distributed random component to account for thermal noise within the electronics. The optimum processing for detecting ideal PPM signals with an APD detector in the presence of Gaussian noise of thermal origin has been shown to be integration over each slot to obtain the detection statistic, followed by selection of the slot with the greatest integrated value as the signal slot [10], which for our interest is equivalent to integrating over each subslot and adding the results to obtain the slot detection statistic. The expression for the probability of symbol error is complicated; however, it was shown in [8] that, for error probabilities on the order of 10^{-3}, a much simpler Gaussian approximation yields good results. Using the simpler Gaussian approximation, the probability that the decision statistic D takes on the value X, which is now modeled as a continuous random variable defined over the real line, is

$$p_D(X) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[\frac{-(X-\mu)^2}{2\sigma^2}\right] \tag{13}$$

where both the mean value μ and the variance σ^2 depend on the total optical energy absorbed by the detector, the surface leakage current, the device parameters k (ionization ratio), and average gain. The random nature of the gain also contributes an excess variance term that depends on the average gain and the ionization ratio of the device, as described in [8]. Using the Gaussian approximation, the expression for the error probability is similar to the expression for any orthogonal signal set (of which PPM is an example) detected in the presence of additive Gaussian noise, except in this case both the mean and the variance depend on the signal:

$$P(SE) = 1 - \frac{1}{\sqrt{2\pi}\sigma_{sb}} \int_{-\infty}^{\infty} dy \exp\left[\frac{-(y-\mu_{sb})^2}{2\sigma_{sb}^2}\right] \left[\frac{1}{\sqrt{2\pi}\sigma_b} \int_{-\infty}^{y} dx \exp\left[\frac{-(x-\mu_b)^2}{2\sigma_b^2}\right]\right]^{(M-1)} \tag{14}$$

The symbol-error performance of 256-PPM signals with both photon-counting and APD detection is shown in Fig. 7, without background radiation and with an average of 20 background photons per slot (the intermediate background case). Note that with photon counting an average of 7 (no background) and 40 (with background) signal photons are needed to achieve a symbol-error probability of 10^{-3}, but with APD detection, assuming nominal APD parameters as in [12], the corresponding numbers are close to an average of 550 signal photons for both cases. The reason is that the performance of the APD detector is dominated by thermal noise, so that small additional increases in noise due to the detected background light hardly matter.

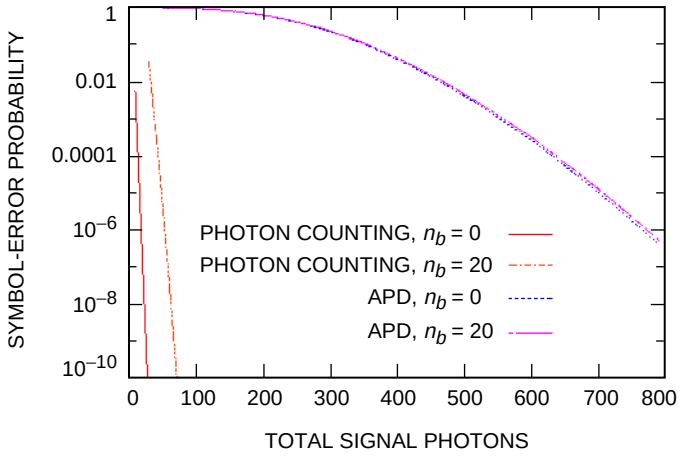

Fig. 7. Symbol-error probabilities for 256 PPM signals, with no background, and with an average of 20 background photons per slot.

V. Slot Synchronization Assembly

Before slot synchronization is established, the receiver is cognizant of time only in terms of its own clock cycles. While it is assumed that the transmitter and receiver clocks run at approximately the same rate, time-varying propagation delays and small differences between the two clock rates have to be taken into account. The receiver initiates the slot-synchronization cycle by starting to count at an arbitrary clock edge, designating that instant as the start of the first subslot. Assuming photon counting in the absence of background light, the receiver measures the subslot accumulator function for each of $2(M + N)$ subsequent subslots and records the location of the subslot containing the greatest count. In general, the counts will not be evenly distributed over two slots, but even if the counts were evenly distributed by chance, the receiver may arbitrarily assign the first subslot as the one containing the pulse. Note that it is possible (although not very likely) that no count whatsoever is observed over the entire frame, in which case the acquisition process has to be repeated until at least one count is observed.

After finding a subslot with a maximal count, the acquisition algorithm defines that subslot as the first half of a valid slot and forms the first error signal $e_1 = N_{11} - N_{12}$. The synchronization assembly then repeats this process for a predetermined number of frames, K, and forms the sum-error signal

$$E = \sum_k e_k \tag{15}$$

which is the synchronization statistic that will be used for adjusting the receiver clock. If most of the laser pulse falls within the first subslot, the average error signal tends to be positive; otherwise it is negative (with high probability). If positive, the clock is advanced by a small amount; if negative, it is delayed. Obtaining a value close to zero for the error signal implies that the slots are essentially synchronized with the received laser pulses by chance; hence, no additional delays need to be applied.

Note that whereas the subslot observables are Poisson-distributed random variables, their difference is not. The mean of the difference is the difference of the means, and the variance is the sum of the variances:

$$e_i = \bar{N}_{i1} - \bar{N}_{i2} \tag{16a}$$

$$\mathrm{var}(e_i) = \bar{N}_{i1} + \bar{N}_{i2} \tag{16b}$$

where $\bar{N}_{ij}, j = 1, 2$, refers to the average count in the jth subslot of the ith slot. The density for the sum of K subslot differences can be obtained directly by noting that the resulting random variable is equivalent to the sum over the second subslot subtracted from the sum over the first subslot over K slots. Letting S_1 be the sum over the first subslot for K slots and S_2 the sum over the second subslot for K slots, the difference statistic can now be expressed as

$$E = \sum_{i=1}^{K} e_i = \sum_{i=1}^{K} (N_{i1} - N_{i2}) = \sum_{i=1}^{K} N_{i1} - \sum_{i=1}^{K} N_{i2} \equiv S_1 - S_2 \tag{17}$$

The partial sums S_1 and S_2 are themselves Poisson random variables with mean and variance given by

$$\bar{S}_j = \sum_{i=1}^{K} \bar{N}_{ij}, \quad j = 1, 2 \tag{18a}$$

$$\mathrm{var}(S_j) = \sum_{i=1}^{K} \bar{N}_{i,j}, \quad j = 1, 2 \tag{18b}$$

This leads to the mean and variance of the K-sum differences as

$$\bar{E} = \bar{S}_1 - \bar{S}_2 \tag{19a}$$

$$\mathrm{var}(E) = \mathrm{var}(S_1) + \mathrm{var}(S_2) \tag{19b}$$

The probability density for the difference of two Poisson random variables has been derived by Pratt in [11]. Substituting the variables defined above yields the density

$$P(E = j|\text{signal slot}) = \left(\frac{\bar{S}_1}{\bar{S}_2}\right)^{-j/2} \exp\left[-(\bar{S}_1 + \bar{S}_2)\right] I_{|j|}\left[2\sqrt{\bar{S}_1 \bar{S}_2}\right] \tag{20}$$

where $I_{|j|}(x)$ is the modified Bessel function of order $|j|$. This density, together with the second-order statistics defined in Eq. (19), completely characterizes the error signal used to update the PPM slot-synchronization system with photon-counting detection, provided the signal slot has been identified correctly. However, as we have shown in Subsection IV.A, the signal slot is identified correctly most of the time under normal operating conditions, since deep-space communications links typically are designed to operate with uncoded symbol-error probabilities of approximately 10^{-3}. The only effect of an occasional detection error is to reduce the expected value and variance of the partial sums, since noise slots have less optical energy than signal-plus-noise slots.

The analysis for the case of APD detection parallels the derivation for the photon-counting case, except that Poisson random variables are now replaced by Gaussian random variables. Again assuming that the synchronization assembly uses only signal slots to obtain the error signal (almost always true

under nominal operating conditions), the mean of the error signal is the difference of the means, and the variance of the error signal is the sum of the variances of the subslot observables.

The average value of the error signal depends on the time delay Δ between the slot boundaries defined by the receiver and the relative delay of the received pulses over a K-frame averaging interval. If the receiver is well synchronized, then the error signal should be zero; otherwise a positive or negative error signal is developed that is used by the synchronization assembly to recenter the pulses within the slots defined by the receiver. The magnitude of the error signal typically increases as the pulse drifts away from zero but reaches a limit and actually starts decreasing as the pulse begins to drift outside the slot. The function defined by the average error signal plotted in terms of the delay offset is called an "S-curve" in the synchronization literature. An example of a measured S-curve obtained with an APD responding to PPM pulses from a doubled Nd:YAG laser is shown in Fig. 8 (reprinted from [12] by permission of the author).

Closed-loop performance of the linearized synchronization assembly depends on the slope of the S-curve at zero and on the total loop gain and any additional filtering before applying the K-sum error signal to the receiver clock.

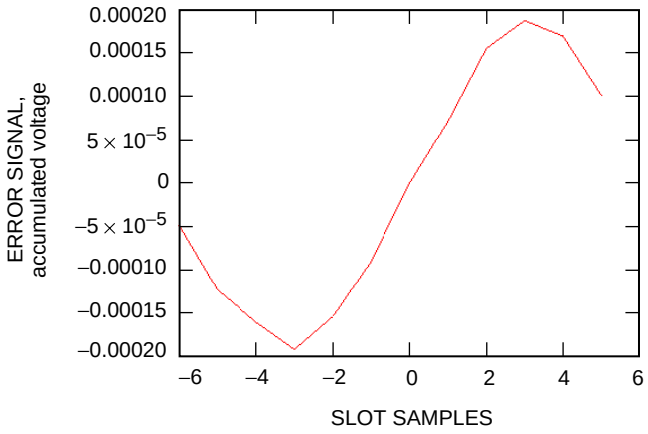

Fig. 8. Measured S-curve obtained with an APD detector (from [12]).

VI. Ranging Assembly

The use of narrow optical pulses in PPM systems enables greater information throughput since the narrower the pulse the more slots that can be packed in a given time interval—hence the greater the information content of each PPM word. An additional advantage of this modulation scheme is that narrow pulses enable precise time resolution since their arrival time can be determined with an accuracy proportional to the electrical bandwidth of the pulse. The ability to determine the arrival time of pulses with a high level of accuracy can in turn be used to determine changes in the distance to the transmitter since light travels with known constant velocity in vacuum. However, it is precisely this change in distance that the synchronization assembly is measuring as it adjusts the delay of the receiver clock. Therefore, it seems natural to attempt to extend the capabilities of the receiver to ranging by keeping track of the applied delays and converting the accumulated delay to distance.

The basic idea is to improve the estimate of the distance between the transmitter and the receiver or between the receiver and a retroreflector by measuring the time of flight of the laser pulse. For the case of a retroreflector, the transmitter and receiver are co-located, and the firing of the laser pulse can be assumed to be known, since it can be determined with great accuracy simply by diverting a small fraction

of the transmitted light to a local detector. For the case of a distant transmitter, the change in distance can be related to the change in delay measured by the synchronization assembly. If the distance to either the retroreflector or the transmitter is known a priori with enough accuracy to resolve frame ambiguities, and the delay applied to the receiver clock by the slot-synchronization assembly is calibrated, then a distance resolution corresponding to a small fraction of a subslot should be possible, just as a small time delay can be measured with greater than subslot accuracy near the zero of the S-curve. For example, the linear portion of the S-curve near zero can be calibrated in distance, and the time-averaged error signal can be supplied to the ranging assembly before it is applied to the receiver clock. The delays applied by the slot-synchronization assembly over time need to be recorded if the distance changes with time. The limit in distance resolution appears to be the length of time the error signals can be integrated, which usually is determined by external factors such as rate of change of trajectory.

VII. Hardware Design and Implementation

The PPM analog signal chain (ASC) design architecture is derived from consideration of (1) detector response, (2) laser pulse shape/width, (3) frame rate, (4) analog-to-digital (ADC) domain, resolution, sample rate, and aperture, and (5) link dynamics.

A. Analog Receiver Front End

Detector response determines overall ASC bandwidth and gain requirements. The laser pulse convolved with the detector response defines an "optimal" filter response for the ASC to minimize noise bandwidth. The PPM frame rate establishes the lower bounds of ASC bandwidth. The ADC parameters establish the upper bounds of the ASC bandwidth and affect gain considerations. The ADC sample aperture is assumed small enough not to significantly reduce the digitization bandwidth from the Nyquist limit, and the sample aperture error is folded into the effective resolution parameter. Link dynamics must be taken into account. Received signal fluctuations may be the result of inherent intensity variations in Q-switched laser output, additional fluctuations due to pointing jitter at the laser transmitter, and atmospheric turbulence-induced fluctuations (which average out for large aperture receivers).

Figure 9 depicts the generic PPM ASC topology. The post-amplifier block maps the detector voltage output (range) to the ADC input (domain) with the goal of minimal noise addition. The optimal filter block matches the ASC transfer response to the expected pulse response to minimize the processed-noise bandwidth. The minimal implementation of the optimal filter block would be a bandpass filter, with a low-frequency limit set by the frame rate and the high-frequency limit set by Nyquist sampling considerations. The clamp amplifier block protects the ADC from signal excursions outside of its nominal conversion domain.

The APD ASC design is driven by the large post-amplifier gain required to match the expected small signal amplitude to the ADC domain. Three design topologies have been considered for the APD post-amplifier, namely, (1) fixed gain, (2) linear gain, and (3) logarithmic approximation. The fixed-gain solution is depicted in Fig. 10, where A_v is the amplifier voltage gain. A multistage amplifier

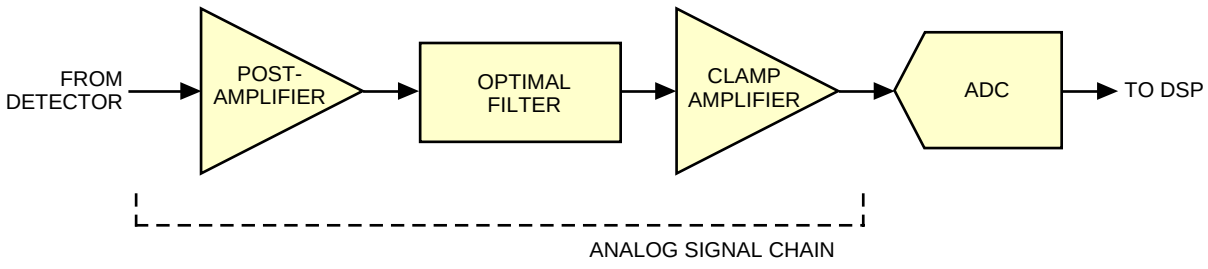

Fig. 9. Generic PPM analog signal-chain architecture.

chain is required due to gain–bandwidth (GBW) product limitations of available low-noise amplifier gain stages. The major defect of this topology is that all of the receiver dynamic range must be provided by increased resolution in the ADC. A linear multiplier stage may be added after the first amplifier stage to accommodate larger link power fluctuations, as depicted in Fig. 11. The gain-control signal comes either from a manual control or from the band-limited output of an AGC loop. The variable-gain topology accommodates a larger input-signal dynamic range with the disadvantage of limited response bandwidth and at the expense of additional electronics noise. The final post-amplifier topology considered eliminates the need for a gain-control loop by approximating a logarithmic response through summing the output from a series of domain-limited amplifiers. This topology is depicted in Fig. 12. Figure 13 compares the transfer function of this topology with a true logarithm response.

Although the logarithmic approximation post-amplifier is the most complicated ASC design considered here, it is presently selected for the PPM receiver design as it has the potential to maximize the efficiency of data acquisition during field tests by minimizing the effects of link dynamics.

An auxiliary design factor to consider is the PMT anode current limit, as this relates to the maximum frame rate and link dynamics (maximum signal level). Starting with an upper bound on the frame rate at the nominal signal level,

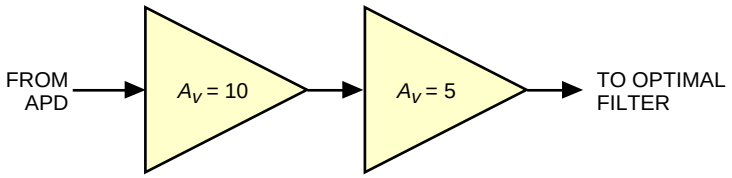

Fig. 10. Fixed-gain APD post-amplifier topology.

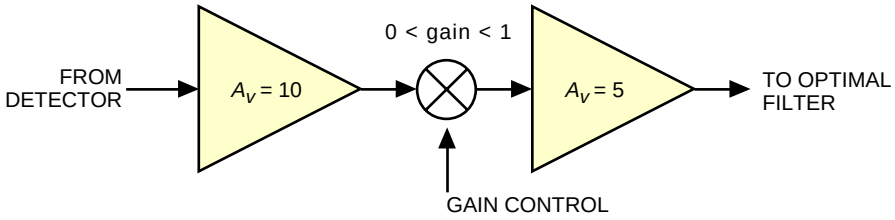

Fig. 11. Variable-gain APD post-amplifier topology.

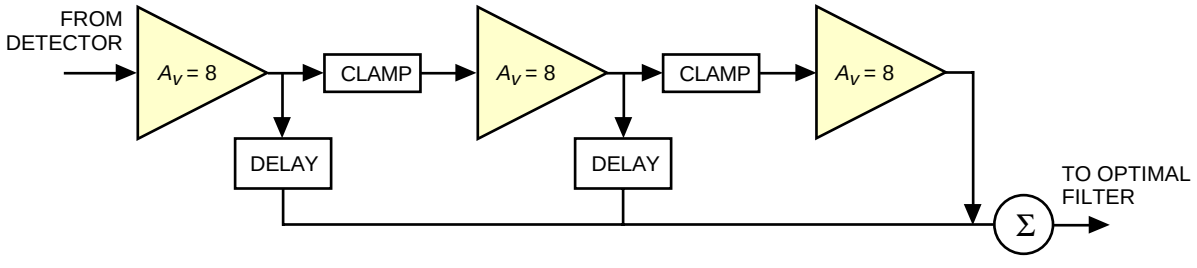

Fig. 12. Logarithmic approximation for the APD post-amplifier topology.

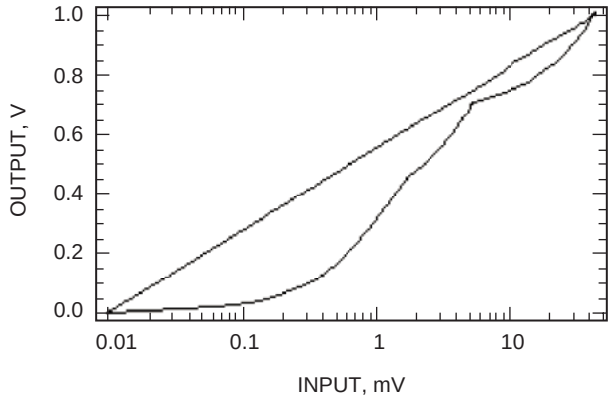

Fig. 13. The transfer function for the proposed logarithmic amplifier compared with an ideal logarithmic response (straight line).

$$PRF_{\max} = \frac{I_{\text{anode limit}}}{\bar{N}_{\text{pulse}} G} \tag{21}$$

where $PRF_{\max}$ is the frame-rate upper bound, $I_{\text{anode limit}}$ is the PMT anode current limit, and G is the PMT gain. The available link "headroom" in db can then be defined as

$$10 \log_{10} \frac{PRF_{\max}}{PRF_{\text{nom}}} \tag{22}$$

where PRF_{nom} is the nominal frame rate.

The optimal filter module should ideally pass only those frequency/phase components that fit the detector response to the signal laser pulse. This minimizes the system bandwidth and, therefore, also minimizes the noise bandwidth to maximize the detector signal-to-noise ratio. However, using numerically simulated Q-switched laser pulses as well as actual laser pulse shapes recorded with detectors in the laboratory, the optimal matched filter provided only 0.1 dB of improvement in signal-to-noise ratio over a simple bandpass filter, due to the limited 100-MHz analog bandwidth. Therefore, a simple bandpass filter extending from 100 kHz to 100 MHz shall be sufficient for this receiver.

The final clamp amplifier to limit the ADC domain is simply required to have a 100-MHz bandwidth and unity gain. It should have a moderately low equivalent input noise and ideally also will be able to directly drive the selected ADC input impedance requirement.

B. Digital Signal Processing Assembly

The digital signal processing assembly (DSPA) implements the slot-synchronization algorithms described in Section IV and supports operation under scenarios described in Sections I and III. The DSPA will initially support PPM frame widths of 256 slots with minimum subslot widths of 5 ns (200 MS/s sampling) and a minimum of 2 subslots per slot for a slot duration of 10 ns. Frame synchronization will be achieved by precalibration of known (constant) data. Support for smaller frame sizes of 128 and 64 slots will be added after initial receiver performance is validated. A block diagram of the digital assembly with external interfaces is shown in Fig. 14.

In addition to identifying the slot number of the largest pulse (the sum of subslot values) as described in Section IV and storing the amplitude of this pulse, the DSPA will output raw slot amplitudes with a slot

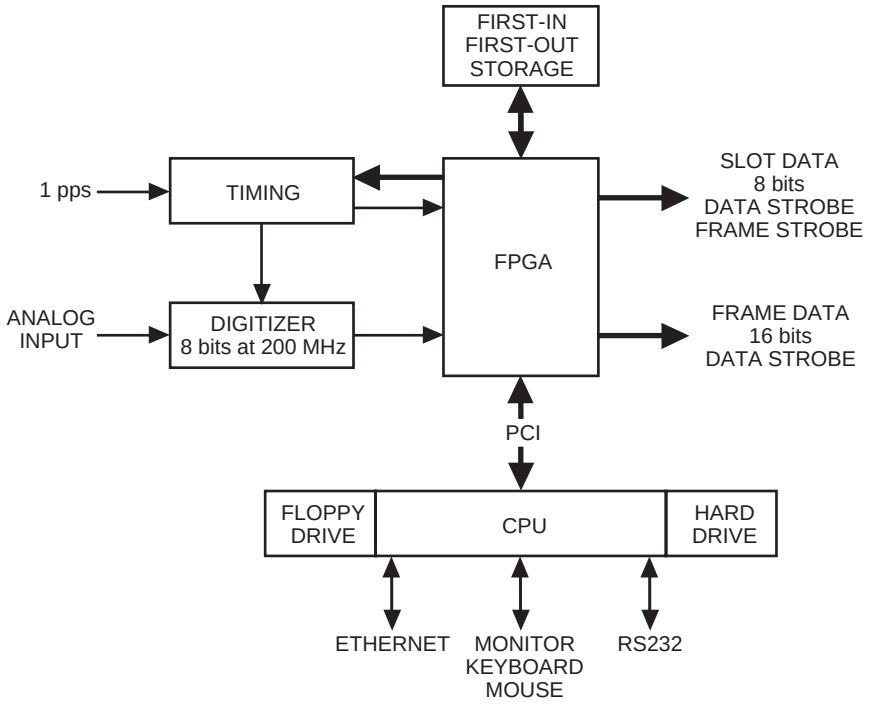

Fig. 14. Block diagram of the DSPA.

and frame strobe to a connector to allow for more extensive signal processing. A limited number of these samples will be captured and written to internal disk for diagnostic purposes. The decoded bytes and their associated frame amplitude will be output on a connector along with a strobe. The same information is recorded on disk. The 200-MHz clock will be generated in the timing block from a crystal voltage-controlled oscillator (VCO) whose frequency can be varied by at least ±50 pulses per million (ppm) to support the relative motion of the transmitter and receiver that will occur under expected operating conditions. The 1 pulse per second (pps) source provides an external timing reference for experiment synchronization and internal time tagging. The digitizer is an 8-bit analog-to-digital converter running at the VCO rate of (nominally) 200 MHz. The field-programmable gate array (FPGA) does most of the work in implementing the algorithms of Section IV. It is presently selected as a Xilinx Virtex II one-million-gate programmable device. This will allow the receiver to be highly reconfigurable to support alternate PPM frame sizes and synchronization algorithms. The central processing unit (CPU) is a standard general-purpose single-board PC that will communicate with the FPGA via the peripheral component interconnect (PCI) bus. The CPU will also implement a user interface to configure slot synchronization, frame synchronization, and data recording parameters. The CPU will also control the frequency of the VCO in order to achieve and maintain slot synchronization based on subslot sums provided by the FPGA.

VIII. Conclusions

In this article, the conceptual design of a prototype optical PPM receiver suitable for laboratory and field demonstrations is presented. The key functions of the receiver have been identified, the required subsystems described, and preliminary analysis of expected performance discussed. A hardware implementation of this receiver concept is currently under way. When completed, the receiver will enable testing, evaluation, and comparison of detectors (such as APD and PMT) and receiver functions critical to deep-space optical communications. Analog front-ends capable of dealing with atmospheric-turbulence-induced fade characteristics of the received laser signal will be demonstrated, and novel PPM symbol detection and slot-synchronization algorithms will be implemented and tested with this receiver. Design flexibility

has been incorporated into the architecture through the use of modular functional blocks employing field-programmable gate arrays, providing for easy expansion of receiver capabilities to demonstrate additional desirable features such as ranging and decoding of encoded PPM symbols. It is believed that the knowledge and experience gained through this research and development effort will contribute substantially towards future implementation of optical communications within the DSN.

Acknowledgment

The authors would like to thank Meera Srinivasan of the Communications Systems and Research Section for generating and plotting the photon-counting and APD detection data of Fig. 7 and also for generating Fig. 8.

References

[1] C.-C. Chen, "Figure of Merit for Direct-Detection Optical Channels," *The Telecommunications and Data Acquisition Progress Report 42-109, January–March 1992*, Jet Propulsion Laboratory, Pasadena, California, pp. 136–151, May 15, 1992.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-109/109L.PDF

[2] H. Plaessmann, S. A. Rae, J. J. Alonis, D. L. Vecht, and W. M. Grossman, "Multipass Diode-Pumped Solid-State Optical Amplifier," *Optics Letters*, vol. 18, pp. 1420–1422, 1993.

[3] D. L. Robinson and M. D. Rayman, "Calculation of Laser Cavity Dumping for Optical Communications," *The Telecommunications and Data Acquisition Progress Report 42-95, July–September 1988*, Jet Propulsion Laboratory, Pasadena, California, pp. 174–179, November 15, 1988.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-95/95Q.PDF

[4] J. Churnside, "Aperture Averaging of Optical Scintillations in the Turbulent Atmosphere," *Applied Optics*, vol. 30, pp. 1982–1994, 1991.

[5] K. E. Wilson and A. Biswas, "Effect of Aperture Averaging on a 570 Mbps 42 km Horizontal Path Optical Link," *Proceedings SPIE, Atmospheric Propagation and Remote Sensing IV*, edited by C. Dainty, Orlando, Florida, pp. 244–253, 1995.

[6] A. Biswas and S. Lee, "Ground-to-Ground Optical Communications Demonstration," *The Telecommunications and Mission Operations Progress Report 42-141, January–March 2000*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–31, May 15, 2000.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-141/141G.pdf

[7] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, New York: John Wiley & Sons, 1991.

[8] M. Srinivasan and V. Vilnrotter, "Symbol-Error Probabilities for Pulse-Position Modulation Signaling With an Avalanche Photodiode Receiver and Gaussian Thermal Noise," *The Telecommunications and Mission Operations Progress Report 42-134, April–June 1998*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–11, August 15, 1998.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-134/134E.pdf

[9] R. M. Gagliardi and S. Karp, *Optical Communications*, New York: John Wiley & Sons, 1976.

[10] V. Vilnrotter, M. Simon, and M. Srinivasan, "Maximum Likelihood Detection of PPM Signals Governed by Arbitrary Point Process Plus Additive Gaussian Noise," *IEE Electronics Letters*, vol. 35, no. 14, pp. 1132–1133, July 1999.

[11] W. K. Pratt, *Laser Communications Systems*, New York: John Wiley & Sons, 1969.

[12] M. Srinivasan, J. Hamkins, B. Madden-Woods, A. Biswas, and J. Beebe, "Laboratory Characterization of Silicon Avalanche Photodiodes (APDs) for Pulse-Position Modulation (PPM) Detection," *The InterPlanetary Network Progress Report 42-146, April–June 2001*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–14, August 15, 2001.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-146/146F.pdf