Copyright Notice

©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This document was downloaded from Chalmers Publication Library (http://publications.lib.chalmers.se/), where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec. 8.1.9 (http://www.ieee.org/documents/opsmanual.pdf)

(Article begins on next page)
Abstract—In this paper, the design and analysis of a new bandwidth-efficient signalling method over the bandlimited intensity-modulated direct-detection (IM/DD) channel is proposed. The channel can be modeled as a bandlimited channel with nonnegative input and additive white Gaussian noise. Due to the nonnegativity constraint, the methods previously proposed for conventional bandlimited channels cannot be applied here. We propose a method to transmit without intersymbol interference in a narrower bandwidth compared to previous works, by combining Nyquist pulses with a constant bias. In fact, we can transmit with a bandwidth equal to that of coherent transmission. A trade-off between the required average optical power and the bandwidth is investigated. At low bandwidths, the most power-efficient transmission is obtained by either the parametric linear pulse or the so-called “better than Nyquist” pulse, depending on the exact bandwidth.

Index Terms—Optical communications, intensity-modulated direct-detection (IM/DD) systems, indoor diffuse wireless optical communications, Nyquist pulses, short-haul optical fiber links, strictly bandlimited signalling

I. INTRODUCTION

The growing demand for high-speed data transmission systems has introduced new design paradigms for optical communications. The need for low-complexity and cost-effective systems has motivated the usage of affordable optical hardware (e.g., incoherent transmitters, optical intensity modulators, multimode fibers, direct-detection receivers) to design short-haul optical fiber links (e.g., fiber to the home and optical interconnects) [1] and diffuse indoor wireless optical links [2]–[4]. These devices impose three important constraints on signalling design. First, the transmitter only modulates information on the instantaneous intensity of an optical carrier, contrary to conventional coherent channels where the amplitude and phase of the carrier can be modulated [5, Sec. 4.3]. In the receiver, only the optical intensity of the incoming signal will be detected [3]. Due to these limitations, the transmitted signal must be nonnegative. Such transmission is called intensity modulation with direct detection (IM/DD). Second, the peak and average optical power (i.e., the peak and average of the transmitted signal in the electrical domain) must be below a specific limit for eye- and skin-safety concerns [3] and to avoid nonlinearities in the devices [6], [7], whereas in conventional channels, such constraints are usually imposed on the peak and average of the squared electrical signal. Third, the bandwidth is limited due to the nonzero response times of the optoelectronic devices [4], [8] and other limitations (e.g., in short-haul optical fiber links due to the modal dispersion [9] and in diffuse indoor wireless optical links due to multipath distortion [3]). Consequently, the coherent modulation formats and pulse shaping methods designed for conventional electrical channels (i.e., with no nonnegativity constraint on the transmitted signal) cannot be directly applied to IM/DD channels.

Much research has been conducted on determining upper and lower bounds on the capacity of IM/DD channels considering power and bandwidth limitations [10]–[15]. Pulse shaping for the purpose of reducing intersymbol interference (ISI) in conventional channels has been previously investigated in [16]–[19]. In [3], [20]–[25], the performance of various modulation formats in IM/DD channels were studied using rectangular or other time-disjoint (i.e., infinite-bandwidth) pulses. Hralinovic in [26] pioneered in investigating the problem of designing strictly bandlimited pulses for IM/DD channels with nonnegative pulse-amplitude modulation (PAM) schemes. He showed the existence of nonnegative bandlimited Nyquist pulses, which can be used for ISI-free transmission over IM/DD channels, and evaluated the performance of such pulses. He also showed that root-Nyquist pulses with matched filters are not suitable for this purpose. He concluded that transmission is possible with a bandwidth twice the required bandwidth over the corresponding conventional electrical channels. This work was extended to other Nyquist pulses that can introduce a trade-off between bandwidth and average optical power in [8], [27].

In this paper, we present a new signalling method for bandlimited IM/DD channels, in which the transmitted signal becomes nonnegative by the addition of a constant direct-current (DC) bias. Hence, we can transmit ISI-free with a bandwidth equal to the bandwidth in coherent conventional channels, while benefiting from the reduced complexity and cost of IM/DD system. We will evaluate the spectral efficiency and optical power efficiency of binary modulation formats with Nyquist pulses for achieving a specific bit-error-rate (BER) or a specific electrical received energy.

The remainder of the paper is organized as follows. Section II presents the system model. In Section III, we introduce the pulses that have been used extensively for conventional bandlimited channels, as well as the ones that have been suggested for nonnegative bandlimited channels, and a method of computing the required DC bias for a general pulse. Section IV introduces the performance measure and analyzes the
performance of the system under different scenarios. Finally, conclusions are drawn in Section V on the performance of the system.

II. SYSTEM MODEL

In applications such as diffuse indoor wireless optical links and short-haul optical fiber communications, where inexpensive hardware is used, IM/DD is often employed. In such systems, the data is modulated on the optical intensity of the transmitted light using an optical intensity modulator such as a laser diode or a light-emitting diode. This optical intensity is proportional to the transmitted electrical signal. As a result, the transmitted electrical signal must be nonnegative. This is in contrast to conventional electrical channels, where the data is modulated on the amplitude and phase of the carrier [5, Sec. 4.3]. In the receiver, the direct-detection method is used in which the photodetector generates an output which is proportional to the incident received instantaneous power [22].

Another limitation, which is considered for safety purposes, is a constraint on the average optical power, or equivalently, a constraint on the average of the signal in the electrical domain [3, 8, 10, 11, 13]. In this study, we consider the IM/DD transmission system with a strict bandwidth limitation and binary modulation.

Fig. 1 represents the system model for an IM/DD optical transmission system. It can be modeled as an electrical baseband transmission system with additive white Gaussian noise (AWGN) and a nonnegativity constraint on the channel input [2, 3, 8, 28]. Initially, based on the independent and identically distributed information bits \( a_k \in \{0, 1\} \), where \( k \in \mathbb{Z} \) is the discrete time instant, an electrical signal \( I(t) \) is generated. The optical intensity modulator converts the electrical signal to an optical signal with intensity \( x(t) \), which is a linear function of \( I(t) \) [3] and is given by

\[
x(t) = JI(t) = JA \left( \mu + \sum_{k=-\infty}^{\infty} a_k q(t - kT_s) \right),
\]

where \( J \) is the laser conversion factor, \( A \) is a scaling factor that can be adjusted depending on the desired transmitted power, \( \mu \) is the required DC bias, \( q(t) \) is an arbitrary pulse, and \( T_s \) is the symbol duration.

Three requirements are placed on \( x(t) \): it should be nonnegative, bandlimited, and ISI-free. The nonnegativity constraint, \( x(t) \geq 0 \) for all \( t \in \mathbb{R} \), is fulfilled by choosing \( \mu \) in (1) sufficiently large, see Sec. III-C. This DC bias is added equally to each symbol to maintain a strictly bandlimited signal \( x(t) \), in contrast to works like [22], [24], [25] in which the bias is allowed to vary with time. The bandwidth constraint is fulfilled by choosing the pulse \( q(t) \) such that

\[
\mathcal{F}\{q(t)\} = Q(\omega) = 0, \ |\omega| \geq B,
\]

where \( \mathcal{F}\{\cdot\} \) denotes the Fourier transform. The condition of ISI-free transmission, finally, is fulfilled by choosing \( q(t) \) as a Nyquist pulse, see Sec. III-A and III-B. Fig. 2 illustrates an example of the transmitted intensity given by (1).

It is desirable to minimize the average optical power [3, 8, 11, 13]

\[
P_{\text{opt}} = \frac{1}{T_s} \int_{0}^{T_s} \mathbb{E}\{x(t)\} \, dt,
\]

where \( \mathbb{E}\{\cdot\} \) denotes expectation, which for the definition of \( x(t) \) in (1) yields

\[
P_{\text{opt}} = \frac{1}{T_s} \int_{0}^{T_s} JA \left( \mu + \mathbb{E}\{a_k\} \sum_{k=-\infty}^{\infty} q(t - kT_s) \right) \, dt = JA \left( \mu + \frac{\eta}{2} \right),
\]

where

\[
\eta = \frac{1}{T_s} \int_{-\infty}^{\infty} q(t) \, dt.
\]

The optical signal then propagates through the channel and is detected and converted to the electrical signal [3, 11]

\[
y(t) = R h(t) \otimes x(t) + n(t),
\]

where \( R \) is the responsivity of the photodetector, \( \otimes \) is the convolution operator, \( h(t) \) is the channel impulse response, and \( n(t) \) is the noise. Without loss of generality, we assume that \( R = J = 1 \) [3]. Furthermore, the channel is considered to be flat in the bandwidth of interest, i.e., \( h(t) = H(0) \delta(t) \).

We model the optical intensity modulator and photodetector as ideal linear devices, ignoring all impairments except the noise. Since the thermal noise of the receiver and the shot noise induced by ambient light are two major noise sources

---

**Fig. 1**: Baseband system model, where \( a_k \) is the \( k \)-th input bit, \( q(t) \) is an arbitrary pulse, \( \mu \) is the DC bias, \( I(t) \) is the transmitted electrical signal, \( x(t) \) is the optical intensity, \( h(t) \) is the channel impulse response, \( n(t) \) is the noise, \( g(t) \) is the impulse response of the receiver filter, and \( \hat{a}_k \) is an estimate of \( a_k \).
in this setup, which are independent from the signal, \( n(t) \) can be modeled by a zero-mean AWGN with double-sided power spectral density \( N_0/2 \) [3], [5], [13], [29]. Although the input signal to the channel \( x(t) \) must be nonnegative, there is no such constraint on the received signal \( y(t) \) [10].

The received signal, similarly to [8], [26], enters a sampling receiver. Hence, it passes through a filter with impulse response \( g(t) \), resulting in

\[
r(t) = y(t) \otimes g(t),
\]

which is then sampled at the symbol rate. In this paper, the receiver filter is assumed to have rectangular frequency response to limit the power of the noise in the receiver, and is given by

\[
G(\omega) = \begin{cases} 
G(0) & |\omega| < B \\
0 & |\omega| \geq B 
\end{cases}.
\]

Without loss of generality, we assume that \( G(0) = H(0) = 1 \).

The system model introduced in this section is a generalization of the one in [8], which is obtained by setting \( \mu = 0 \) in (1). If \( \mu = 0 \), the pulse \( q(t) \) should be nonnegative to guarantee a nonnegative signal \( x(t) \). In our proposed system model, by introducing the bias \( \mu \), the nonnegativity condition can be fulfilled for a wider selection of pulses \( q(t) \).

### III. BANDBLOCKED NYQUIST PULSES

In order to have ISI-free transmission with a sampling receiver, the pulse \( q(t) \) must satisfy the Nyquist criterion [16]. In other words, for any \( k \in \mathbb{Z} \) [5, Eq. (9.2-11)],

\[
q(kT_s) = \begin{cases} 1 & k = 0 \\
0 & k \neq 0 \end{cases}.
\]

#### A. Traditional Nyquist Pulses

In this section, we consider three Nyquist pulses that have been proposed for the conventional coherent channel. In all cases, the bandwidth can be adjusted via the roll-off factor \( \alpha \) chosen in the range \( 0 \leq \alpha \leq 1 \):

(i) Raised-cosine (RC) pulse which is defined as

\[
q_{RC}(t) = \sin \left( \frac{t}{T_s} \right) \cos \left( \frac{\alpha \pi t}{T_s} \right)
\]

where \( \sin(x) = \sin(\pi x)/\pi x \).

(ii) The parametric linear (PL) pulse of first order defined in [18], which is given by

\[
q_{PL}(t) = \sin \left( \frac{t}{T_s} \right) \sin \left( \frac{\alpha t}{T_s} \right).
\]

(iii) The so-called “better than Nyquist” (BTN) pulse [19], which in [18] was referred to as parametric exponential pulse, given by

\[
q_{BTN}(t) = \sin \left( \frac{t}{T_s} \right) \frac{4/3\pi t \sin \left( \frac{\pi t}{T_s} \right) + 2/3 \cos \left( \frac{2\pi t}{T_s} \right) - \beta^2}{4\pi^2 t^2 + \beta^2},
\]

where \( \beta = 2T_s \ln 2/\alpha \).

All these pulses have a lowpass bandwidth \( B = \pi(1+\alpha)/T_s \) and time average \( \bar{q} = 1 \).

#### B. Nonnegative Nyquist Pulses

In this section, which is motivated by [8], all the three aforementioned constraints should be satisfied by the pulse. As a result, in (1), \( \mu = 0 \) and \( q(t) \geq 0 \) for all \( t \in \mathbb{R} \).

In [8], it has been shown that pulses that satisfy these requirements must be the square of a general Nyquist pulse. This will result in having pulses with bandwidth twice that of the original Nyquist pulses. Three pulses that satisfy these constraints were introduced in [8]:

(i) Squared sinc (S2), which is given by

\[
q_{S2}(t) = \sin^2 \left( \frac{t}{T_s} \right),
\]

has the bandwidth \( B = 2\pi/T_s \) and time average \( \bar{q} = 1 \).

(ii) Squared raised-cosine (SRC), given by

\[
q_{SRC}(t) = q_{RC}^2(t),
\]

requires a larger bandwidth \( B = 2\pi(1+\alpha)/T_s \) compared to squared sinc, and \( \bar{q} = 1 - \frac{\alpha}{2} \).

(iii) Squared double-jump (SDJ), given by

\[
q_{SDJ}(t) = \left[ \left( \frac{1-\alpha}{2} \right) \sin \left( \frac{1-\alpha}{T_s} \right) + \left( \frac{1+\alpha}{2} \right) \sin \left( \frac{1+\alpha}{T_s} \right) \right]^2,
\]

requires the same bandwidth as SRC (i.e., \( B = 2\pi(1+\alpha)/T_s \)), but has a lower \( \bar{q} \) for a given \( \alpha \) compared to the other two pulses \( \bar{q} = 1 - \frac{\alpha}{2} \).

Figs. 2 and 3 depict the normalized transmitted signal \( x(t)/A \) using the RC (7) and SRC (11) pulses, respectively.

#### C. Required DC Bias

Our goal is to find the lowest \( \mu \) that guarantees the nonnegativity of \( x(t) \). From (1) and (4), \( x(t) \geq 0 \), the smallest required DC bias is

\[
\mu = -\min_{\forall a_i, -\infty \leq t \leq \infty} \sum_{i=-\infty}^{\infty} a_i q(t-iT_s)
\]

\[
\leq -\min_{\forall a_i, -\infty \leq t \leq \infty} \sum_{i=-\infty}^{\infty} \left( a_i - \frac{1}{2} \right) q(t-iT_s) + \frac{1}{2} q(t-iT_s)
\]

\[
= \frac{1}{2} \max_{0 \leq t < T_s} \left[ \sum_{i=-\infty}^{\infty} |q(t-iT_s)| - \sum_{i=-\infty}^{\infty} q(t-iT_s) \right].
\]

The notation \( \forall a_i \) in (13) and (14) means that the minimization should be over all \( a_i \in \{0, 1\} \) where \( i = \ldots, -1, 0, 1, 2, \ldots \).

The reason that (15) is minimized over \( 0 \leq t < T_s \) is that \( \sum_{i=-\infty}^{\infty} q(t-iT_s) \) and \( \sum_{i=-\infty}^{\infty} |q(t-iT_s)| \) are periodic functions with period equal to \( T_s \). Since for all pulses defined in Sec. III, \( q(t) \) can rescale with \( T_s \) as \( q(t) = f(t/T_s) \) for some function \( f(t) \), then \( \mu \) is independent of \( T_s \).

Fig. 4 illustrates the required DC bias for various pulses. Due to the fact that by increasing \( \alpha \), the ripples of the pulses decrease, the required DC bias decreases as well. It can be seen
Fig. 2: The normalized transmitted signal $x(t)/A$ using an RC pulse (7) with $\alpha = 0.6$ as $q(t)$. It can be seen that without using the bias $\mu = 0.184$, the raised-cosine pulse would create a signal $x(t)$ that can be negative.

Fig. 3: The normalized transmitted signal $x(t)/A$ using an SRC pulse (11) with $\alpha = 0.6$ as $q(t)$. In this case, the required DC $\mu$ is zero.

that the RC pulse (7) always requires more DC bias. Moreover, the PL (8) and the BTN (9) pulses require approximately the same DC bias. The BTN pulse requires slightly less DC bias in $0.250 \leq \alpha \leq 0.256$, $0.333 \leq \alpha \leq 0.363$, and $0.500 \leq \alpha \leq 0.610$, while the parametric linear pulse is better for all other roll-off factors in the range $0 < \alpha < 1$.

The expression for $\mu$ given in (15) illustrates the reason why the double-jump and sinc pulses are not considered in Sec. III-A. These pulses decay as $1/|t|$. As a result, the summation in (15) does not converge to a finite value. Hence, they require an infinite amount of DC bias to be nonnegative.

**IV. ANALYSIS AND RESULTS**

**A. Received Sequence**

Considering the assumptions mentioned in Sec. II, the received signal (4) is

$$r(t) = (x(t) + n(t)) \otimes g(t)$$

$$= A \left( \mu + \sum_{k=-\infty}^{\infty} a_k q(t - kT_s) \right) \otimes g(t) + z(t)$$

$$= A \left[ \mu + \sum_{k=-\infty}^{\infty} a_k q(t - kT_s) \right] + z(t),$$

where (18) holds since $g(t)$ has a flat frequency response given by (5) over the bandwidth of $q(t)$ given by (2); Therefore, the convolution has no effect on $x(t)$. The noise at the output of the receiver filter, which is given by $z(t) = n(t) \otimes g(t)$, is zero mean additive white Gaussian with variance $\sigma_z^2 = N_0 B/(2\pi)$.

Applying the Nyquist criterion given in (6) to the sampled version of (18), we can write the $i$-th filtered sample as

$$r(iT_s) = A[\mu + a_i] + z(iT_s).$$

The corresponding electrical energy of the noiseless filtered samples will be

$$E_{rec} = A^2 \mathbb{E} \left\{ [\mu + a_i]^2 \right\}.$$

**B. Comparison Between Pulses**

To compare the optical power of various pulses, a criterion called optical power gain is used, which is defined as [8]

$$\Upsilon = 10 \log_{10} \left( \frac{P_{\text{ref}}}{P_{\text{opt}}} \right),$$

where $P_{\text{opt}}$ is the average optical power for a reference system. This reference is chosen to be the squared sinc pulse (10), for
which no bias is needed. Using (3), \( P_{\text{opt}} = \frac{A}{2} \) and

\[
\Upsilon = 10 \log_{10} \left( \frac{A_{\text{ref}}}{A (2\mu + \bar{q})} \right) \tag{21}
\]

where \( A_{\text{ref}} \) is the scaling factor for the reference pulse.

Initially, we compare the pulses in a noise-free setting. Fig. 5 demonstrates the comparison of the optical power gain for various pulses, defined in Sec. III, where the signals are scaled to have equal \( E_{\text{rec}} \) (20). Squared sinc (10), which is used as a baseline for comparison, is shown in the figure with an arrow. It must be noted that the results for \( BT_s \geq 2\pi \) have been driven before in [8, Fig. 4], whereas the results for \( BT_s < 2\pi \) are novel and according to our proposed method.

For the pulses in Sec. III, to have the same \( E_{\text{rec}} \) as the reference system,

\[
A_{\text{ref}} = \sqrt{\mathbb{E}\left\{ (\mu + a_i)^2 \right\}} / \sqrt{\mathbb{E}\{a_i^2\}}.
\]

As a result, (21) can be written as

\[
\Upsilon = 10 \log_{10} \left( \frac{\sqrt{2\mu^2 + 2\mu + 1}}{2\mu + \bar{q}} \right).
\]

The results in Fig. 5 are consistent with [8, Fig. 4], where the nonnegative pulses in Sec. III-B were studied, but it can be seen that when the pulses discussed in Sec. III-A are used, and the nonnegativity constraint is satisfied by adding a DC bias, the transmission is possible over a much narrower bandwidth. However, since the DC bias consumes energy and does not carry information, the optical power gain will be reduced. It must be noted that there is a compromise between bandwidth and optical power gain, due to the fact that \( \mu \) will be reduced by increasing the roll-off factor (see Fig. 4), whereas the required bandwidth increases. The highest optical power gain in all cases will be achieved when the roll-off factor is one. The BTN and the PL pulses have approximately similar optical power gain, and the RC pulse has smaller gain.

It appears from Fig. 5 that the studied pulses become more power-efficient when the bandwidth is increased. A higher bandwidth, however, means that the receiver filter admits more noise, which reduces the receiver performance. In Fig. 6, we therefore compare the average optical power gain of the same pulses when the power is adjusted to yield a constant BER. Similarly to previous case, the squared sinc (10) pulse is used as a baseline for comparison.

To find the optical power gain as a function of BER, we first apply a maximum likelihood detector to (19), which yields the BER [5, Sec. 9.3]

\[
P_{\text{err}} = Q \left( \frac{A}{2 \sqrt{N_0B}} \right),
\]

where

\[
Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp \left( -\frac{x^2}{2} \right) dx
\]

is the Gaussian Q-function. As a result,

\[
A = 2Q^{-1}(P_{\text{err}}) \sqrt{\frac{N_0B}{2\pi}},
\]

and

\[
A_{\text{ref}} = \sqrt{\frac{B_{\text{ref}}}{B}},
\]

where \( B_{\text{ref}} = 2\pi/T_s \) is the bandwidth for the reference pulse. The optical power gain now follows from (21). The dependence on \( \sqrt{B (2\mu + \bar{q})} \) is the reason why by increasing the bandwidth, the gain for SRC (11) slightly decreases whereas for SDJ (12) increases, where \( \mu = 0 \) for both cases.

We observe that for the biased pulses in Sec. III-A, the gain increases by increasing the bandwidth. The reason is that by increasing the roll-off factor, the required bias decreases much faster (see Fig. 4) than the speed of increase in bandwidth. The BTN and PL pulses have approximately similar gain, and the

![Fig. 5: The optical power gain \( \Upsilon \) versus normalized bandwidth \( BT_s/(2\pi) \). The electrical power of noiseless samples for all pulses is equal. The results for \( BT_s < 2\pi \) are novel, while those for \( BT_s \geq 2\pi \) are prior knowledge.](image1)

![Fig. 6: The optical power gain versus normalized bandwidth \( BT_s/(2\pi) \). The BER for all pulses is equal.](image2)
gain of the RC pulse is always smaller than the gain of the other two pulses.

When the roll-off factor is equal to zero (i.e., the normalized bandwidth $BT_\alpha/(2\pi)$ for the biased pulses is equal to 0.5), the pulses discussed in Sec. III-A will become equal to a sinc pulse with bandwidth $\pi/T_\alpha$. As discussed in Sec. III-C, the required DC will be infinite for the sinc pulse. Hence, the gain $\Upsilon$ will asymptotically go to $-\infty$ when $\alpha \to 0$.

V. CONCLUSION

In this work, a pulse shaping method for strictly bandlimited IM/DD systems is presented. Such transmission imposes constraints on the bandwidth and average optical power. Moreover, the transmitted signal must be nonnegative. The approach proposed uses the DC bias as a degree of freedom to satisfy the nonnegativity constraint. This allows us to use Nyquist pulses for ISI-free transmission with narrower bandwidth compared to previous works. Hence, it is possible to transmit with a bandwidth equal to that of ISI-free transmission in conventional coherent channels. This allows us to have the benefit of using affordable hardware over IM/DD channels, while requiring the same bandwidth as conventional channels.

To compare our proposed transmission scheme with previously designed ones, we evaluated analytically the average optical power versus bandwidth in two different scenarios. Of the studied pulses, the “better than Nyquist” pulse is most power efficient for $0.625 \leq BT_\alpha/(2\pi) \leq 0.628$, $0.666 \leq BT_\alpha/(2\pi) \leq 0.682$, and $0.750 \leq BT_\alpha/(2\pi) \leq 0.805$, while the parametric linear pulse is better for all other bandwidths in the range $0.5 < BT_\alpha/(2\pi) < 1$. At $BT_\alpha/(2\pi) \geq 1$, the squared double-jump pulse is the best known, as previously shown in [8].

This work can be a starting point for ISI-free pulse shaping design for transmission within a bandwidth equal to that of coherent conventional channels. Future work can concentrate on designing coding schemes to improve the BER performance and compensate the effect of the DC bias.

REFERENCES


