Performance Analysis and Channel Capacity for Multiple-Pulse Position Modulation on Multipath Channels

Hyuncheol Park and John R. Barry
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332-0250

Abstract

Although multiple pulse-position modulation performs well on ideal channels, its performance on multipath channels is degraded significantly. In an attempt to quantify the inherent penalty due to multipath dispersion, we evaluate upper bounds for the error probability of each modulation scheme in the presence of intersymbol interference, considering both an unequalized receiver and the optimal maximum-likelihood sequence detection receiver. We also present upper and lower bounds of the channel capacity for multiple pulse-position modulation and its variants, PPM and overlapping PPM. Numerical results show that the PPM-based schemes are significantly more sensitive to multipath dispersion than is on-off keying.

I. Introduction

Non-directed infrared radiation offers several advantages over radio as a medium for indoor wireless networks, including an immense window of unregulated bandwidth, immunity to multipath fading (but not multipath distortion), and a lack of interference from one room to another [1]. But the intense background light of typical indoor environments is a severe impediment, and power-efficient modulation is needed to achieve high data rates or long range.

The wireless infrared channel is accurately modeled by the following baseband AWGN model [2]:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)\,d\tau + n(t), \tag{1}$$

where $x(t)$ represents the instantaneous optical power of the transmitter, $y(t)$ represents the instantaneous current of the receiving photodetector, $h(t)$ represents the multipath impulse response, and $n(t)$ is white Gaussian noise with two-sided power spectral density N_0. The high intensity of the background light makes the Gaussian noise model extremely accurate.

Because $x(t)$ represents optical power, it must satisfy:[1]

$$x(t) \geq 0 \ \text{ and } \ \langle x(t) \rangle \leq P, \tag{2}$$

where P is the average optical power constraint of the transmitter.

A recent paper [3] examined the performance of multiple PPM on the wireless infrared channel assuming no multipath dispersion. Audeh and Kahn [4] examined the effects of intersymbol interference (ISI) on PPM. We extend these results by examining the effects of ISI on multiple PPM and overlapping PPM.

The channel capacity is considered to be a fundamental limit for reliable transmission over a given channel [5]. Georghiades [6] considered the channel capacity for MPPM over a photon counting channel, without intersymbol interference. Hirt [7] calculated the channel capacity for a binary discrete-time Gaussian channel using a Monte Carlo approximation. Shamai [8] derived lower and upper bounds of the channel capacity with i.i.d. input symbols for a scalar discrete-time Gaussian channel. In this paper, we examine the channel capacity of various modulation schemes on the channel (1) under the constraints of (2).

In Sect. II, we develop the general system model for multiple PPM and its variants. In Sect. III, we analyze the performance of the unequalized receiver and the maximum-likelihood sequence detector. In Sect. IV, we present the expression for channel capacity on the ideal channel and bound the channel capacity of PPM-based schemes over ISI channels. We present numerical results in Sect. V.

II. System Model

Consider the system model shown in Fig. 1(a). Information bits with rate R_b b/s enter the encoder, which groups the bits into blocks of length $\log_2 L$ and maps each block to one of L codewords $\boldsymbol{c}_0 \ldots \boldsymbol{c}_{L-1}$, where each codeword is a binary n-tuples of weight w. The set of allowable codewords is what distinguishes the different modulation schemes. When all $\binom{n}{w}$ n-tuples of weight w are valid codewords, the resulting modulation scheme is called multiple PPM (MPPM). When the only valid codewords are those binary n-tuples of weight w in which the w ones are consecutive, the result is overlapping PPM (OPPM). Finally, when w is restricted to unity, both MPPM and OPPM reduce to conventional PPM modulation. Note that the number of codewords L for MPPM, OPPM, and PPM is $\binom{n}{w}$, $n-w+1$, and n, respectively. The output of the encoder is a sequence of codewords $\{\boldsymbol{x}_k\}$ with rate $1/T = R_b/\log_2 L$. This sequence is serialized to produce the binary chip sequence $\{x_j\}$ with rate n/T, where $\boldsymbol{x}_k = [x_{kn},\ x_{kn+1}, \cdots,\ x_{kn+n-1}]^T$. The binary chip sequence drives a transmitter filter with a rectangular pulse shape $p(t)$ of duration T/n and unity height. To satisfy the power constraint of (2), the filter output is multiplied by (nP/w) before the signal is sent across the channel.

As shown in Fig. 1(a), the receiver uses a unit-energy filter $f(t)$ and samples the output at the chip rate n/T producing y_j. The receiver groups the samples y_j into blocks of length n, producing a sequence of observation vectors $\{\boldsymbol{y}_k\}$, where $\boldsymbol{y}_k = [y_{kn}, y_{kn+1}, \cdots, y_{kn+n-1}]^T$. The receiver passes each observation vector through a decision device to form an estimate $\hat{\boldsymbol{x}}_k$ of $\boldsymbol{x}_k$.

Supported by NSF under grant number NCR-9308968.

1. $\langle \cdot \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} (\cdot)\,dt$

To appear, International Symposium on Personal, Indoor and Mobile Radio Communications, Taipei, October 1996.

The equivalent discrete-time channel between transmitted and received chips is:

$$y_j = \sum_{i=-\infty}^{\infty} h_i x_{j-i} + n_j = s_j + n_j, \tag{3}$$

where h_j is the equivalent chip-rate impulse response:

$$h_j = \frac{nP}{w}\Big(p(t) * g(t) * f(t) \Big)\Big|_{t = jT/n}, \tag{4}$$

and where s_j is defined by (3).

We assume that $f(t) * f(-t)$ is a Nyquist pulse, which will be true when $f(t)$ is matched to $p(t)$ or when $f(t)$ is a whitened-matched filter. In this case, the noise samples $\{n_j\}$ will be independent zero-mean Gaussian random variables with variance N_0. As shown in Fig. 1(b), the equivalent vector channel between transmitted codewords $\boldsymbol{x}_k$ and observation vectors $\boldsymbol{y}_k$ is given by:

$$\boldsymbol{y}_k = \sum_{l=0}^{\mu} \mathbf{H}_l \boldsymbol{x}_{k-l} + \boldsymbol{n}_k = \boldsymbol{s}_k + \boldsymbol{n}_k, \tag{5}$$

where the channel impulse response is a Toeplitz sequence $\mathbf{H}_k$, with $[\mathbf{H}_k]_{ij} = h_{kn+i-j}$, $\boldsymbol{s}_k = [s_{kn}, s_{kn+1}, \ldots, s_{kn+n-1}]^T$ is the signal component, $\boldsymbol{n}_k = [n_{kn}, n_{kn+1}, \ldots, n_{kn+n-1}]^T$ is the noise component and μ is the number of memories in vector ISI channel. Throughout this paper we constrain the input symbols $\boldsymbol{x}_k$ to be independent and uniformly distributed over a MPPM or OPPM alphabet.

III. Performance Analysis

III-A. Without Equalization

By definition, the unequalized receiver uses the decision device that would be optimal were there no ISI. In other words, it decides on the codeword $\boldsymbol{c}_l$ that maximizes the correlation:

$$\Lambda_l = \boldsymbol{c}_l{}^T \boldsymbol{y}_k \quad \text{for } l = 0, \ldots, L-1. \tag{6}$$

If we knew that $\boldsymbol{x}_k = \boldsymbol{c}_i$, and if we knew all past and future codewords $\mathbf{X}' = \{\ldots, \boldsymbol{x}_{k-2}, \boldsymbol{x}_{k-1}, \boldsymbol{x}_{k+1}, \boldsymbol{x}_{k+2}, \ldots\}$, then the probability of error $\hat{\boldsymbol{x}}_k \neq \boldsymbol{x}_k$ can be bounded using the union bound:

$$\begin{aligned}\Pr[\text{error} \mid \boldsymbol{x}_k = \boldsymbol{c}_i, \mathbf{X}'] &\le \sum_{j=0, j\neq i}^{L-1} \Pr[\Lambda_i \le \Lambda_j \mid \boldsymbol{x}_k = \boldsymbol{c}_i, \mathbf{X}'] \qquad (7)\\ &= \sum_{j=0, j\neq i}^{L-1} \Pr[(\boldsymbol{c}_i - \boldsymbol{c}_j)^T \boldsymbol{n}_k > (\boldsymbol{c}_i - \boldsymbol{c}_j)^T \boldsymbol{s}_k \mid \boldsymbol{x}_k = \boldsymbol{c}_i, \mathbf{X}'].\end{aligned}$$

But $(\boldsymbol{c}_i - \boldsymbol{c}_j)^T \boldsymbol{n}_k$ is a zero-mean Gaussian random variable with variance $d_{ij}N_0$, where $d_{ij} = d_H(\boldsymbol{c}_i, \boldsymbol{c}_j)$ is the Hamming distance between codewords $\boldsymbol{c}_i$ and $\boldsymbol{c}_j$. Therefore, (7) reduces to:

$$\Pr[\text{error} \mid \boldsymbol{x}_k = \boldsymbol{c}_i, \mathbf{X}'] \le \sum_{j=0, j\neq i}^{L-1} Q\left(\frac{(\boldsymbol{c}_i - \boldsymbol{c}_j)^T \boldsymbol{s}_k}{\sqrt{d_{ij} N_o}} \right). \tag{8}$$

Averaging overall all possible codeword sequences gives:

$$\Pr[\text{error}] \le \frac{1}{L^{M+1}} \sum_{\mathbf{X}'} \sum_{i=0}^{L-1} \sum_{j=0, j\neq i}^{L-1} Q\left(\frac{(\boldsymbol{c}_i - \boldsymbol{c}_j)^T \boldsymbol{s}_k}{\sqrt{d_{ij} N_o}} \right), \tag{9}$$

where the first summation is over all $\mathbf{X}' \in \mathcal{C}^M$, where $\mathcal{C}$ is the set of L valid codewords and $M+1$ is the number of nonzero terms in the impulse response $\{\mathbf{H}_k\}$. Finally, the bit-error probability is:

$$\Pr[\text{bit error}] = 1 - (1 - \Pr[\text{error}])^{1/\log_2 L}. \tag{10}$$

For on-off keying (OOK), the bit stream, symbol stream, and chip stream are all one in the same. By averaging over all possible bit streams $\{\boldsymbol{x}'\}$, the total bit error probability is:

$$\Pr[\text{bit error}] = \frac{1}{2^M} \sum_{\boldsymbol{x}'} Q\left(\frac{h_0/2 - \sum{}' x_n h_{j-n}}{\sqrt{N_o}} \right), \tag{11}$$

where the summation is overall all binary M-tuples $\{\boldsymbol{x}'\}$, where $M+1$ is the number of nonzero terms in h_j.

Hereafter, we present simplified expressions for the symbol error probability for the special case of an ideal channel, without ISI. First, consider multiple PPM: when the channel has no ISI, the expression (9) simplifies to:

$$\Pr[\text{error}] \le \sum_{k=1}^{w} a_k Q\left(\sqrt{\frac{k s^2}{2N_o}} \right), \tag{12}$$

where $a_k = \binom{w}{k}\binom{n-w}{k}$ is the number of codewords with mutual distance $2k$, and $s = (P/w)\sqrt{n \log_2 L / R_b}$. This expression follows from the ISI-free results for the photon counting channel of [6]. When $w = 1$, (i.e., for PPM), (12) simplifies further to:

$$\Pr[\text{error}] \le (L-1) Q\left(\frac{s}{\sqrt{2N_o}} \right), \tag{13}$$

where $s = P\sqrt{L \log_2 L / R_b}$. Next, consider $\binom{n}{w}$ OPPM: when the channel has no ISI, the expression (9) simplifies to:

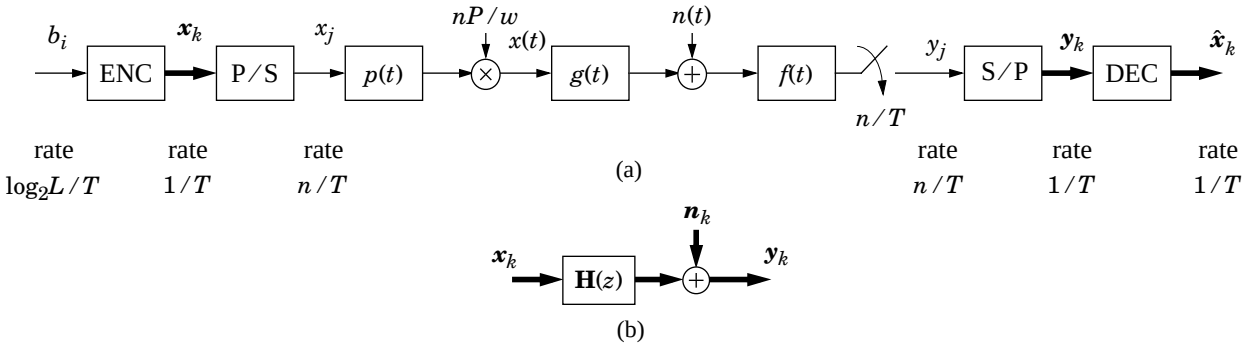

Fig. 1. (a) Block diagram of multiple PPM system; (b) equivalent vector model.

$$\Pr[\text{error}] \le \frac{1}{L}\sum_{k=1}^{w} b_k Q\left(\sqrt{\frac{ks^2}{2N_o}}\right), \tag{14}$$

$$\text{where } b_k = \begin{cases} 2(L-k) & k = 1, 2, \ldots, w-1 \\ (L-w)(L-w+1) & k = w \end{cases}, \tag{15}$$

and $s = (P/w)\sqrt{n \log_2 L / R_b}$. This result uses the photon-counting results of [6]. Finally, consider OOK: when the channel has no ISI, then h_j from (4) reduces to $(2P/\sqrt{R_b})\delta_j$, and so the error probability from (11) simplifies to $Q(P/\sqrt{N_0 R_b})$.

III-B. ML Sequence Detection

The maximum-likelihood (ML) sequence detector for PPM is derived in [9], and it easily generalizes to multiple PPM and overlapping PPM. The receiver filter is the whitened-matched filter, so that h_j is causal and minimum phase. The probability of a symbol (block) error for the ML sequence detector is well approximated at high SNR by:

$$\Pr[\text{error}] \approx Q(d_{min}/2\sqrt{N_0}), \tag{16}$$

where d_{min} is the minimum distance between received sequences:

$$d_{min}^2 = \min_{\{\boldsymbol{e}_k\}} \sum_k \| \sum_m \mathbf{H}_m \boldsymbol{e}_{k-m} \|^2. \tag{17}$$

The above minimization is performed over all possible error sequences $\{\boldsymbol{e}_k\}$ starting at time zero, using an error alphabet of $\{\boldsymbol{u} - \boldsymbol{v}: \boldsymbol{u} \neq \boldsymbol{v}; \boldsymbol{u}, \boldsymbol{v} \in C\}$, where C is the set of valid multiple PPM codewords.

IV. Channel Capacity

IV-A. Memoryless Channel

If the channel is memoryless (ISI-free), so that $\mathbf{H}_0 = \mathbf{I}$, $\mathbf{H}_l = \mathbf{0}$ for all $l \neq 0$ in (5), then the channel capacity (in bits per codeword) under the i.i.d. constraint is [5]:

$$I_{iid} = I(\boldsymbol{x}; \boldsymbol{x} + \boldsymbol{n}) =$$
$$\frac{1}{L}\sum_{l=0}^{L-1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\|\boldsymbol{y}-\boldsymbol{x}_l\|^2/2\sigma^2} \log_2 \left(\frac{e^{-\|\boldsymbol{y}-\boldsymbol{x}_l\|^2/2\sigma^2}}{\frac{1}{L}\sum_{m=0}^{L-1} e^{-\|\boldsymbol{y}-\boldsymbol{x}_m\|^2/2\sigma^2}} \right) d\mathbf{y}. \tag{18}$$

Following [10], we substitute $\boldsymbol{z} = (\boldsymbol{y} - \boldsymbol{x}_l)/\sigma$ and $\boldsymbol{v}_m = \boldsymbol{x}_m/\sigma$, so that (18) reduces to:

$$I_{iid} = \log_2 L - \frac{1}{L}\sum_{l=0}^{L-1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{2\pi^{n/2}} e^{-\|\boldsymbol{z}\|^2/2} \log_2 \left(\sum_{m=0}^{L-1} e^{-(\boldsymbol{v}_l - \boldsymbol{v}_m)\boldsymbol{z}} e^{-\|\boldsymbol{v}_l - \boldsymbol{v}_m\|^2/2} \right) d\boldsymbol{z}. \tag{19}$$

The above equation contains an n dimensional integral and has no simple closed form solution. As a consequence, we shall use the Monte Carlo method to estimate I_{iid} in Sect. V.

IV-B. Multipath Channel

Following [8], we can also represent the channel (5) using matrix notation:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N} = \mathbf{S} + \mathbf{N}, \tag{20}$$

where $\mathbf{Y} = [\boldsymbol{y}_0^T, \boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_{N-1}^T]^T$, $\mathbf{X} = [\boldsymbol{x}_0^T, \boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_{N-1}^T]^T$, $\mathbf{S} = [\boldsymbol{s}_0^T, \boldsymbol{s}_1^T, \ldots, \boldsymbol{s}_{N-1}^T]^T$, and $\mathbf{N} = [\boldsymbol{n}_0^T, \boldsymbol{n}_1^T, \ldots, \boldsymbol{n}_{N-1}^T]^T$ are $Nn \times 1$ column vectors. The two equations (5) and (20) are equivalent as $N \to \infty$, and the rows of $\mathbf{H}$ are specified by circular shifts of $\{\mathbf{H}_i\}$:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{0} & \ldots & \mathbf{0} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_\mu & \ldots & \mathbf{H}_0 \end{bmatrix}. \tag{21}$$

The channel capacity for (20) under the i.i.d. constraint is [5][7]:

$$I_{iid} = \lim_{N\to\infty} \frac{1}{N} I(\mathbf{Y};\mathbf{X}) = \lim_{N\to\infty} \frac{1}{N} (h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X})) \tag{22}$$

$$= \log_2 L - \lim_{N\to\infty} \frac{1}{N} \sum_{l=0}^{L^N-1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{L^N} \frac{1}{2\pi^{Nn/2}} e^{-\|\boldsymbol{z}\|^2/2} \log_2 \left(\sum_{m=0}^{L^N-1} e^{-(\boldsymbol{S}_l - \boldsymbol{S}_m)\boldsymbol{z} - \|\boldsymbol{S}_l - \boldsymbol{S}_m\|^2/2} \right) d\boldsymbol{z} \tag{23}$$

Exact evaluation of this expression is not possible, so we resort to upper and lower bounds. In the following discussion, we present lower and upper bounds for the channel capacity under the i.i.d. constraint. Our presentation is a straightforward generalization of the scalar results of Shamai [8] to the vector channel (5).

Following [8], we represent the entropy of the output vector in (20) using the chain rule:

$$h(\mathbf{Y}) = \sum_{l=0}^{N-1} h(\boldsymbol{y}_l | \boldsymbol{y}_{l-1}, \boldsymbol{y}_{l-2}, \ldots \boldsymbol{y}_0). \tag{24}$$

Since conditioning decreases the entropy:

$$h(\mathbf{Y}) \ge \sum_{l=0}^{N-1} h(\boldsymbol{s}_l + \boldsymbol{n}_l | \boldsymbol{s}_{l-1}, \ldots, \boldsymbol{s}_0, \boldsymbol{n}_{l-1}, \ldots, \boldsymbol{n}_0) = \sum_{l=0}^{N-1} h(\boldsymbol{s}_l + \boldsymbol{n}_l | \boldsymbol{s}_{l-1}, \ldots, \boldsymbol{s}_0), \tag{25}$$

where the last equality follows because $\boldsymbol{s}_l$, $\boldsymbol{n}_l$, $\boldsymbol{n}_{l-1}$, …, $\boldsymbol{n}_0$ are independent. Since $\boldsymbol{s}_l = \sum_j \mathbf{H}_{l-j}\boldsymbol{x}_j$, (25) reduces to:

$$h(\mathbf{Y}) \ge \sum_{l=0}^{N-1} h(\mathbf{H}_0 \boldsymbol{x}_l + \boldsymbol{n}_l). \tag{26}$$

The mutual information between the input and output vectors is:

$$I(\mathbf{X};\mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{N}) \ge \sum_{l=0}^{N-1} \{h(\mathbf{H}_0 \boldsymbol{x}_l + \boldsymbol{n}_l) - h(\boldsymbol{n}_l)\} = \sum_{l=0}^{N-1} I(\mathbf{H}_0 \boldsymbol{x}_l + \boldsymbol{n}_l; \boldsymbol{x}_l). \tag{27}$$

This leads to the following lower bound I_L for the channel capacity under the i.i.d. constraint:

$$\lim_{N\to\infty}\frac{1}{N}I(\mathbf{X};\mathbf{Y}) \ge \lim_{N\to\infty}\frac{1}{N}\sum_{l=0}^{N-1} I(\mathbf{H}_0\boldsymbol{x}_l+\boldsymbol{n}_l;\boldsymbol{x}_l) = I(\mathbf{H}_0\boldsymbol{x}+\boldsymbol{n};\boldsymbol{x}) \tag{28}$$

We can evaluate (28) by replacing $\boldsymbol{x}$ with $\mathbf{H}_0\boldsymbol{x}$ in (18). As pointed out in [9], the lower bound of the channel capacity is equivalent to the mutual information between the input and the output of an error-free block zero-forcing decision feedback equalizer.

We now present an upper bound I_U for the capacity I_{iid} of (23). By the chain rule, we can represent the mutual information between the input and output of (20) as:

$$I(\mathbf{Y};\mathbf{X}) = \sum_{l=0}^{N-1} I(\mathbf{Y};\boldsymbol{x}_l|\boldsymbol{x}_{l-1},\boldsymbol{x}_{l-2},\ldots,\boldsymbol{x}_0) =$$
$$\sum_{l=0}^{N-1}[h(\boldsymbol{x}_l|\boldsymbol{x}_{l-1},\boldsymbol{x}_{l-2},\ldots,\boldsymbol{x}_0) - h(\boldsymbol{x}_l|\boldsymbol{x}_{l-1},\boldsymbol{x}_{l-2},\ldots,\boldsymbol{x}_0,\mathbf{Y})] \tag{29}$$

Since $\{\boldsymbol{x}_k\}$ are i.i.d. and conditioning decreases the entropy:

$$\begin{aligned} I(\mathbf{Y};\mathbf{X}) &\le \sum_{l=0}^{N-1}[h(\boldsymbol{x}_l|\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{l-1},\boldsymbol{x}_{l+1},\boldsymbol{x}_{l+2},\ldots,\boldsymbol{x}_N) - \\ &\quad h(\boldsymbol{x}_l|\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{l-1},\boldsymbol{x}_{l+1},\boldsymbol{x}_{l+2},\ldots,\boldsymbol{x}_N,\mathbf{Y})] \\ &= \sum_{l=0}^{N-1} I(\boldsymbol{x}_l;\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{l-1},\boldsymbol{x}_{l+1},\boldsymbol{x}_{l+2},\ldots,\boldsymbol{x}_N,\mathbf{Y}) \\ &= \sum_{l=0}^{N-1} I(\boldsymbol{x}_l;\mathbf{H}_0\boldsymbol{x}_l+\boldsymbol{n}_0,\mathbf{H}_1\boldsymbol{x}_l+\boldsymbol{n}_1,\ldots,\mathbf{H}_\mu\boldsymbol{x}_l+\boldsymbol{n}_\mu) \\ &= \sum_{l=0}^{N-1} I(\boldsymbol{x}_l;\tilde{\mathbf{Y}}), \end{aligned} \tag{30}$$

where $\tilde{\mathbf{Y}} = \tilde{\mathbf{H}}\boldsymbol{x}_l + \tilde{\mathbf{N}}$, and where $\tilde{\mathbf{H}} = [\mathbf{H}_0{}^T, \mathbf{H}_1{}^T, \ldots, \mathbf{H}_\mu{}^T]^T$ and $\tilde{\mathbf{N}} = [\boldsymbol{n}_0{}^T, \boldsymbol{n}_1{}^T, \ldots, \boldsymbol{n}_\mu{}^T]^T$.

Let $\mathbf{V} = \tilde{\mathbf{H}}^T\tilde{\mathbf{Y}} = \mathbf{B}\boldsymbol{x}_l + \boldsymbol{w}$, where $\tilde{\mathbf{H}}^T$ is a matched filter, so that $\mathbf{B} = \sum_{m=0}^{\mu}\mathbf{H}_m{}^T\mathbf{H}_m$ and $\boldsymbol{w}$ is a zero-mean Gaussian vector with correlation matrix $\mathrm{E}[\boldsymbol{w}\boldsymbol{w}^T] = \sigma^2\mathbf{B}$. Since $\mathbf{B}$ is positive definite, it can be factored into $\mathbf{B} = \Gamma\Gamma^T$ for some matrix Γ. We can whiten the noise by applying Γ^{-1} to $\mathbf{V}$, yielding $\mathbf{Z} = \Gamma^{-1}\mathbf{V} = \Gamma^T\boldsymbol{x}_l + \boldsymbol{n}$, where $\boldsymbol{n}$ has the same distribution as $\boldsymbol{n}_l$, a zero-mean Gaussian vector with correlation matrix $\sigma^2\mathbf{I}$. Since both the matched filter and noise whitener are information lossless, we have $I(\boldsymbol{x}_l;\tilde{\mathbf{Y}}) = I(\boldsymbol{x}_l;\mathbf{V}) = I(\boldsymbol{x}_l;\mathbf{Z})$. Therefore, (30) reduces to:

$$I(\mathbf{Y};\mathbf{X}) \le \sum_{l=0}^{N-1} I(\boldsymbol{x}_l;\Gamma^T\boldsymbol{x}_l+\boldsymbol{n}_l). \tag{31}$$

Finally, taking the limit as $N\to\infty$ yields our upper bound I_U of the channel capacity under the i.i.d constraint:

$$\lim_{N\to\infty}\frac{1}{N}I(\mathbf{X};\mathbf{Y}) \le \lim_{N\to\infty}\frac{1}{N}\sum_{l=0}^{N-1} I(\boldsymbol{x}_l+\boldsymbol{n}_l;\boldsymbol{x}_l) = I(\Gamma^T\boldsymbol{x}+\boldsymbol{n};\boldsymbol{x}) \tag{32}$$

We can evaluate (32) by replacing $\boldsymbol{x}$ with $\Gamma^T\boldsymbol{x}$ in (18). Note that, for the scalar case, the matrix Γ^T reduces to $\sqrt{\sum_m |h_m|^2}$, and the upper bound (32) reduces to the matched filter bound [8].

V. Numerical Results

To generate numerical results, we assume that the underlying continuous-time channel in Fig. 1 has impulse response $g(t) = We^{-Wt}u(t)$, a first-order low-pass filter with a 3-dB bandwidth of W, where $u(t)$ is the unit step function. Observe that the channel has unity d.c. gain. We also assume that the receive filter $f(t)$ is a unit-energy whitened-matched filter.To reduce computational complexity, we truncate the vector channel of (5) to four terms, so that $\boldsymbol{y}_k = \sum_{l=0}^{3}\mathbf{H}_l\boldsymbol{x}_{k-l} + \boldsymbol{n}_k$. This truncation will have no appreciable effect when n is large or when R_b/W is small, although it may not be accurate for small n and large R_b/W. Observe that the channel has unity d.c. gain.

We calculated the optical power required to achieve a 10^{-6} bit-error rate over this ISI channel. The results are summarized in Fig. 2, where the normalized power requirement is plotted versus the bit-rate-to-bandwidth ratio R_b/W. The power requirements are normalized by $P_{OOK} = \sqrt{N_0 R_b}\,Q^{-1}(10^{-6})$, the power required by OOK in the ideal case ($W = \infty$) to achieve a 10^{-6} bit error rate.

In Fig. 2(a) we plot power requirement versus R_b/W when equalization is not used. We see that some modulation schemes

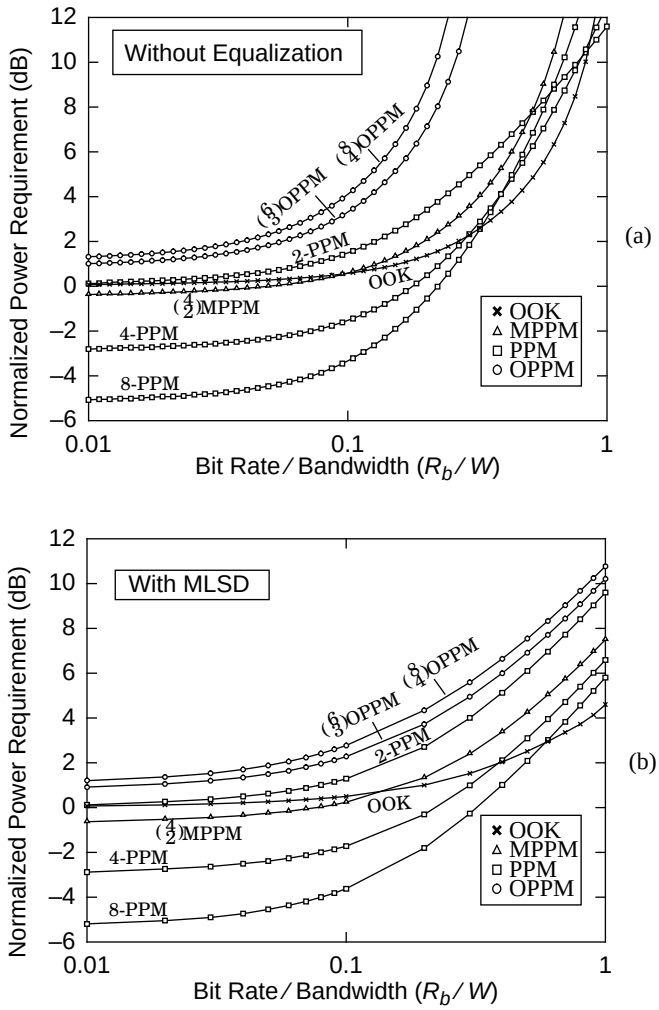

Fig. 2. Power requirement vs bit rate on an ISI channel: (a) Unequalized system (b) ML sequence detection.

are more sensitive to ISI than others. At large bandwidth ($R_b/W < 0.1$), the ISI penalties are small. At one extreme is OOK (represented by the symbol '×'), with a power requirement increasing slowly with decreasing bandwidth. At the other extreme is OPPM, for which the power requirement grows rapidly with decreasing bandwidth. It is thus highly desirable to use signal processing at the receiver to mitigate ISI, either equalization or maximum-likelihood sequence detection.

In contrast to the unequalized results of Fig. 2(a), the results of Fig. 2(b) are based on the maximum-likelihood sequence detector (MLSD). Comparing Fig. 2(a) and Fig. 2(b), we see that MLSD offers significant improvement. The power requirements do not grow as rapidly as in the unequalized case, and the normalized power requirement is always less than 12 dB, even when $R_b/W = 1$. We note that MLSD is much more effective in reducing the power requirement for OOK than for other modulation schemes.

The exact capacity I_{iid} of (23) is difficult to evaluate, but we can form an estimate $\hat{I}_{iid}$ by choosing N suitably large in (23) rather than letting $N \to \infty$, and by using the Monte Carlo method with Q sample vectors to approximate the multiple integral. In Fig. 3, I_{ideal} (19), I_L (28), and I_U (32) are shown for 2 PPM, and compared to the approximate channel capacity $\hat{I}_{iid}$ (based on $N = 6$, $Q = 1000$). Our results show that I_L and I_U are 0.5 dB apart at moderate SNR when $R_b/W = 0.5$.

When SNR = 3.3 dB, the channel capacity is 0.95 bits/codeword when the channel is ideal, but is only 0.18 bits/codeword when $R_b/W = 0.5$. To achieve a capacity of 0.95 bits/codeword using 2 PPM at $R_b/W = 0.5$, the required SNR is 9.5 dB. In contrast, we see from Fig. 2(b) that an uncoded 2-PPM system with MLSD requires an additional 3.3 dB, or 12.8 dB SNR, to achieve 10^{-6} BER at $R_b/W = 0.5$. Thus, the coding gain for a code based on 2-PPM can be at most 3.3 dB in this case. Higher coding gains are possible for higher-order alphabets.

VI. Conclusions

We have examined the performance of multiple PPM and its variants PPM and OPPM on ISI channels with additive white Gaussian noise. We have derived the channel capacity over ISI channels of multiple PPM and its variants PPM, OPPM, and OOK with additive white Gaussian noise, assuming the codewords are independent and uniformly distributed. For numerical comparisons we considered a first-order low-pass filter channel with bandwidth W. The error probability and channel capacity results indicates that, although PPM modulation schemes are extremely power efficient across ISI-free channels, their power efficiency drops dramatically on ISI channels.

VII. References

[1] F. R. Gfeller, U. Bapst, "Wireless In-House Data Communications via Diffuse Infrared Radiation," *Proceedings of the IEEE*, vol. 67, no. 11, pp. 1474-1486, November 1979.

[2] J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental Characterization of Non-Directed Indoor Infrared Links," *IEEE Transactions on Communications*, vol. 43, no. 2-4, pp. 1613-1623, Feb.-Apr. 1995.

[3] H. Park and J. R. Barry, "Modulation Analysis for Wireless Infrared Communication," *Proc. of Intl. Conf. on Comm.*, Seattle, WA, pp. 1182-1186, June 1995.

[4] M. D. Audeh and J. M. Kahn, "Performance Evaluation of L-Pulse Position Modulation on Non-Directed Indoor Infrared Channels," *ICC*, New Orleans, LA, pp. 660-664, May 1994.

[5] R. G. Gallager, *Information Theory and Reliable Communication*, New York, Wiley, 1968.

[6] C. N. Georghiades, "Modulation and Coding for Throughput-Efficient Optical Systems," *IEEE Transactions on Information Theory*, vol. 40, no.5, pp. 1313-1326, September 1994.

[7] W. Hirt, "Capacity and Information Rates of Discrete-Time Channels with Memory," Ph. D. Dissertation, Swiss Federal Institute of Technology, Zürich, 1988.

[8] S. Shamai (Shitz), L. H. Ozarow, and A. D. Wyner, "Information Rates for a Discrete-Time Gaussian Channel with Intersymbol Interference and Stationary Inputs," *IEEE Transactions on Information Theory*, vol. 37, No. 6, 1527-1539, November 1991.

[9] J. R. Barry, "Sequence Detection and Equalization for Pulse-Position Modulation," *Proc. of Intl. Conf. on Comm.*, New Orleans, LA, pp. 1561-1565, May 1994.

[10] R. E. Blahut, *Principles and Practice of Information Theory*, Reading, MA, Addison-Wesley, 1987.

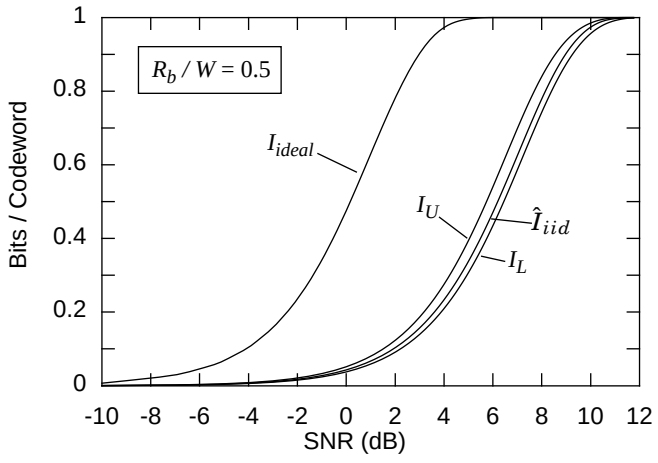

Fig. 3. Bounds on the channel capacity for 2 PPM vs SNR = $P / \sqrt{N_0 R_b}$ for $R_b/W = 0.5$. (The capacity I_{ideal} for the ideal case $R_b/W = 0$ is included for reference.)