

# Channel Uncertainty in Ultra Wideband Communication Systems

Dana Porrat *Member, IEEE*, David N. C. Tse *Member, IEEE* and Şerban Nacu

**Abstract**—Channel uncertainty limits the achievable data rates of certain ultra wide band systems due to the need to estimate the channel. The use of bursty duty-cycled transmission reduces the channel uncertainty because the receiver has to estimate the channel only when transmission takes place, but the maximum amount of burstiness and hence the possible reduction of channel uncertainty both depend on the spectral efficiency of the modulation scheme used. This general principle is demonstrated by comparing the channel conditions that allow duty-cycled direct-sequence spread spectrum (DSSS) and pulse position modulation (PPM) to achieve the AWGN capacity in the wideband limit. We show that duty-cycled DSSS systems achieve the wideband capacity as long as the number of independently faded resolvable paths increases *sub-linearly* with the bandwidth, while duty-cycled PPM systems can achieve the wideband capacity only if the number of paths increases *sub-logarithmically*. The difference is due to the fact that DSSS is spectrally more efficient than PPM and hence allows more bursty transmission.

**Keywords:** Channel Uncertainty, Direct Sequence Spread Spectrum Flash Signaling, Pulse Position Modulation, Spectral Efficiency, Wideband Communications,

## I. INTRODUCTION

This work discusses the achievable data rates of systems with very wide bandwidths. Considering communications with an average power constraint, the capacity of the multipath channel in the limit of infinite bandwidth is identical to the capacity of the additive white Gaussian noise (AWGN) channel  $C_{\text{AWGN}} = P/N_0 \log e$ , where  $P$  is the average received power (due to the transmitted signal) and  $N_0/2$  is the received noise double sided spectral density. *Kennedy* [1] and *Gallager* [2] proved this for Rayleigh fading channels using FSK signals with duty cycle transmission; *Telatar & Tse* [3] extended the proof for arbitrary multipath channels.

When using spreading modulations, *Médard & Gallager* [4] show that direct sequence spread spectrum signals, when transmitted continuously (no duty cycle) over fading channels that have a very large number of channel paths, approach zero data rate in the limit of infinite bandwidth. A similar result was shown by *Subramanian & Hajek* [5]. *Telatar & Tse* [3] show that over multipath channels, the data rate in the limit

of infinite bandwidth is inversely proportional to the number of channel paths.

This work is motivated by a recent surge in interest in ultra wide band systems, where spreading signals are often desired. We show that under suitable conditions, spreading signals can achieve AWGN capacity on multipath channels in the limit of infinite bandwidth, if they are used with duty cycle, i.e. shutting off for a fraction of the time to save energy for higher powered transmission for the rest of the time. In other words, peakiness in time is sufficient to achieve AWGN capacity, and the transmitted signal does not have to be peaky in frequency as well. We analyze direct sequence spread spectrum (DSSS) and pulse position modulation (PPM) signals, and show that when the scaling of the number of channel paths is not too rapid, these signals achieve the capacity in the limit as the bandwidth grows.

Our results can be seen as a middle ground between two previous results: 1. FSK with duty cycle achieves AWGN capacity for any number of channel paths and 2. direct sequence spread spectrum signals with continuous transmission (no duty cycle) have zero throughput in the limit, if the number of channel paths increases with the bandwidth.

The effect of duty cycle can be intuitively understood in terms of the channel uncertainty a communication system faces, that can be quantified by the parameter:

$$\text{SNR}_{\text{est}} = \frac{2PT_c}{N_0\theta L}$$

where  $L$  is the number of independent resolvable paths,  $T_c$  is the coherence time and  $\theta$  is the duty cycle parameter or the fraction of time used for transmission. The quantity  $P/\theta \cdot T_c$  is the total received energy over a coherence time period *during transmission*, and so the ratio  $\text{SNR}_{\text{est}}$  can be interpreted as the effective SNR per resolvable path for channel estimation. A necessary condition for a communication system to achieve the AWGN channel capacity in the limit of infinite bandwidth is that the channel estimation in the limit is perfect, and this requires

$$\text{SNR}_{\text{est}} \rightarrow \infty$$

In systems with bounded average received power, the duty cycle parameter must diminish in order to balance the increase in the number of resolvable paths  $L$  as the bandwidth  $W$  scales, so that the overall channel uncertainty diminishes. The more rapid the increase of  $L$  with  $W$ , the more peaky the transmission has to be.

In an AWGN channel, the spectral efficiency of a communication system is not important to achieve the capacity in the wideband limit, since the number of degrees of freedom is

Presented at the 41<sup>st</sup> Allerton Conference on Communication, Control and Computing, October 2003.

Dana Porrat is with the Hebrew University, dporrat@cs.huji.ac.il. David Tse is with the University of California at Berkeley, dtse@eecs.berkeley.edu. Şerban Nacu is with Ecole Normale Supérieure in Paris, France, serban@stanford.edu

This work was funded in part by the Army Research Office (ARO) under grant #DAAD19-01-1-0477, via the University of Southern California.

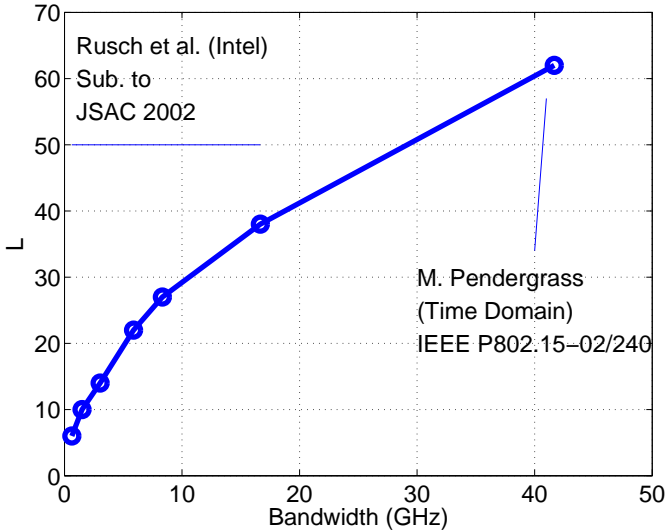


Fig. 1. Number of significant channel paths vs. bandwidth. The paths accounting for 60–90 percent of the energy were counted in two separate measurement campaigns. The Intel data were taken from [6] and the Time Domain data from [7].

not a scarce commodity. This is not the case when low duty cycled transmission is required to reduce channel uncertainty. Spectrally efficient modulation schemes pack more bits into the short on-periods and allow more peaky transmissions (or smaller  $\theta$ ). The difference between the wideband capacities of DSSS and PPM schemes comes about because of their different spectral efficiencies. PPM is an orthogonal modulation, so the number of bits it can transmit per unit time increases only logarithmically with the bandwidth, whereas the number of bits a DSSS transmitter can send per unit time increases linearly. Thus, a lower duty cycle can be used with DSSS and it can tolerate a faster increase of the number of resolvable paths  $L$  than PPM. Note that, in contrast, both PPM and DSSS achieve the channel capacity in the limit of infinite bandwidth over the AWGN channel (as well as over the multipath fading channel where the channel is known to the receiver).

Our main results make these intuitions precise. In the limit of infinite bandwidth, DSSS systems where the receiver knows the path delays achieve AWGN capacity if the number of channel paths is sub-linear in the bandwidth, i.e.  $\frac{L}{W} \rightarrow 0$ , and the system uses an appropriate duty cycle. PPM systems too can achieve AWGN capacity in the limit of infinite bandwidth, but this is possible for a smaller number of channel paths. A PPM system with a receiver that knows the path delays achieves AWGN capacity if  $\frac{L}{\log W} \rightarrow 0$ . PPM systems with lower bounded symbol time have zero throughput if  $\frac{L}{\log W} \rightarrow \infty$ . In systems where the receiver does not know the path gains or delays, we show that DSSS systems can achieve AWGN capacity if  $\frac{L}{W/\log W} \rightarrow 0$  as the bandwidth increases. Measurements of the number of channel paths vs. bandwidth in Figure 1 show an increase of the number of channel paths that appears to be sub-linear.

An alternative approach to communications over multipath fading channels is via orthogonal frequency division multiplexing (OFDM): the spectrum is divided into many narrow,

flatly faded bands and independent signaling is performed separately on each band. The theory of capacity per unit cost [8] tells us that the wideband capacity can be achieved by an on-off scheme per band. Since this scheme does not rely on coding between the bands, it achieves capacity regardless of the joint statistics of the channel gains of the bands. This means the scheme can achieve capacity no matter how the number of resolvable paths scales with the bandwidth. However, its complexity for an ultra wideband system can be prohibitive because of the huge number of sub-carriers. We show that simpler signaling schemes such as PPM can achieve capacity as long as the number of paths scales slowly with the band. Such schemes implicitly exploit the correlation between the channel gains across frequency to improve channel estimation.

In related work, Verdú [9] also pointed out the crucial role of peakiness in time and a connection is also made between the spectral efficiency of a modulation scheme and its data rate over a large bandwidth. The theory there analyzes the capacity of channels with *fixed* channel uncertainty as the average received power per degree of freedom goes to zero. Each such channel can be viewed as a narrowband, flat fading sub-channel of a wideband multipath channel, and thus the framework is suitable for the OFDM system discussed above. The infinite bandwidth AWGN capacity is always approached in the limit by the modulation schemes considered there, and the issue is the *rate of convergence* to that limit. This rate of convergence depends on the spectral efficiency of the modulation scheme. In contrast, the issue we considered here is *whether* a modulation scheme approaches the AWGN capacity at all.

Analysis of capacity of wideband fading channels is also offered by Zheng *et al* [10], [11]. The focus of this work is on the impact of the coherence time on performance. It is shown there that if the coherence time increases fast enough with the bandwidth, then the AWGN capacity can be approached with a rate of convergence as fast as in the coherent case where the channel is perfectly known at the receiver. In contrast, we consider here effectively a dual situation where the coherence time is fixed with the bandwidth, but the *coherence bandwidth* increases with bandwidth when the number of independent resolvable paths increase sub-linearly with the bandwidth.

The outline of the paper is as follows. After presenting the channel model and the signals in Section II, a summary of the results is presented in Section III and discussion is brought in Section IV. Section V presents a bound on the channel uncertainty penalty related to path delays. Sections VI and VII then present bounds on the data rates of DSSS and PPM system where the receiver knows the channel path delays, with conclusions in Section VIII.

## II. CHANNEL MODEL AND SIGNAL MODULATIONS

The natural model for an ultra wide band channel is real, because there is no obvious ‘carrier frequency’ that defines the phase of complex quantities. The channel is composed of  $\tilde{L}$  paths:

$$\tilde{Y}(t) = \sum_{m=1}^{\tilde{L}} \tilde{A}_m(t)X(t - \tau_m(t)) + Z(t) \quad (1)$$

where each path has an amplitude  $\tilde{A}_m(t)$  and a delay  $\tau_m(t)$ ,  $X(t)$  is the transmitted signal,  $Y(t)$  the received signal and  $Z(t)$  is noise. We assume a block fading model: the channel remains constant over coherence periods that last  $T_c$ , and changes independently between coherence periods. The paths' delays  $\{\tau_m\}_{m=1}^L$  are in the range  $[0, T_d]$ , where  $T_d$  is the delay spread. The channel is assumed under-spread, so the delay spread  $T_d$  is much smaller than the coherence period, and signal spillover from one coherence period to the next is negligible.

In order to facilitate a discrete signal model we introduce a chip shaping filter at the transmitter and a corresponding chip matched filter at the receiver, followed by sampling at rate  $W$  that equals the system bandwidth. The chip shaping filter has a rectangular response with duration  $1/W$ . The effective channel model after sampling is discrete, with path amplitudes  $\{A_l\}$ . We assume uncorrelated path amplitudes in the continuous channel (1), the appendix shows that the sampled amplitudes  $\{A_l\}$  at the matched filter output are also uncorrelated. The non-zero taps at the filter output are named 'resolvable paths', their amplitudes are approximately uncorrelated if a raised-cosine or another realistic chip filter is used.

The number of resolvable paths is  $L$  and their delays  $\{D_l/W\}_{l=1}^L$  are multiples of the system time resolution, where  $\{D_l\}_{l=1}^L$  are integers between 0 and  $\lfloor WT_d \rfloor - 1$ . We assume that the  $L$  delays are uniformly distributed over the  $\binom{\lfloor WT_d \rfloor}{L}$  possibilities of combinations of  $L$  values out of  $\lfloor WT_d \rfloor$  positions. Given the system's finite bandwidth, the received signal is approximated within each coherence period by

$$Y(t) = \sum_{l=1}^L A_l X(t - D_l/W) + Z(t)$$

The received signal  $Y(t)$  is discretized by sampling at a rate  $1/W$ . Considering a double sided noise density  $\frac{N_0}{2}$  the discretized and normalized signal is given by

$$Y_i = \sqrt{\frac{\mathcal{E}}{K_c}} \sum_{l=1}^L G_l X_{i-D_l} + Z_i \quad i = 0, \dots, \lfloor T_c W \rfloor - 1 \quad (2)$$

with  $\mathcal{E} = \frac{2PT_c}{N_0\theta} = \text{LSNR}_{\text{est}}$ ,  $P$  is the average received power and  $K_c = \lfloor T_c W \rfloor$ . The noise  $\{Z_i\}$  is standard white Gaussian noise, i.e.  $Z_i \sim N(0, 1)$ . The normalization requires that the path gains  $\{G_l\}$  and the transmitted signal  $\{X_i\}$  are scaled so that the average power of  $\sum G_l X_{i-D_l}$  equals  $\theta$ . We assume that  $\{G_l\}$  are independent and zero mean, thus the scaling requirement is  $E[G_l^2] = 1/L$  and  $X_i^2$  averages (over time and ensemble) to  $\theta$ . The time average is required because the signals we suggest are not stationary. In practical systems, where the channel response is dispersive, the bandwidth of the received signal is larger than the bandwidth of the transmitted signal. We do not consider this difference in our channel model.

In order to avoid complications at the edge of the coherence periods, we approximate the channel using a cyclic difference over  $K_c$  instead of a simple difference:  $(n) \equiv n \bmod K_c$ . The difference is negligible as the delay spread is much smaller

than the coherence time.

$$Y_i = \sqrt{\frac{\mathcal{E}}{K_c}} \sum_{l=1}^L G_l X_{(i-D_l)} + Z_i \quad i = 0, \dots, \lfloor T_c W \rfloor - 1 \quad (3)$$

Note that when  $X$  is a PPM signal described in Section II-B (that includes a  $T_d$  guard time ensuring a silence period between symbols) the circular formulation (3) is identical to the original model (2).

The systems we consider do not use channel information at the transmitter, and the receiver knows the deterministic features of the channel, namely the coherence time  $T_c$ , the delay spread  $T_d$  and the number of paths  $L$ . The receiver sometimes has additional knowledge, about the random features of the channel, in particular the path delays  $D$ . The assumptions on receiver knowledge of the channel are explicitly stated throughout the paper.

The assumptions we make of IID paths that compose the channel ensure that the channel uncertainty is linearly dependent on the number of paths  $L$ . Our results essentially compare the communication data rate to the information (number of bits) needed to describe the channel. Complex channel models, with non-uniform paths, require more delicate analysis to evaluate the channel uncertainty penalty.

#### A. Direct Sequence Spread Spectrum Signals

Each transmitted symbol contains a random series of  $K_c$  IID Gaussian values  $\{r_i\}_{i=0}^{K_c-1}$  with zero mean, and an average energy constraint is satisfied:

$$\frac{1}{Q} \sum_{j=1}^Q E[X_j^2] \Big|_{Q \gg K_c} = \theta E \left[ \frac{1}{K_c} \sum_{i=0}^{K_c-1} r_i^2 \right] = \theta$$

where  $0 < \theta \leq 1$  is the duty cycle parameter or the fraction of time used for transmission. The symbol  $r$  is used for the transmitted signal during active transmission periods, while  $X$  represents all of the transmitted signal, that includes silent periods due to duty cycle. The time average over  $X_j$  is needed because the duty cycle mechanism makes the signal non-stationary.

The duty cycle, with parameter  $\theta$ , is used over coherence periods: of each interval  $\frac{T_c}{\theta}$ , one period of  $T_c$  is used for transmission and in the rest of the time the transmitter is silent (Figure 2). We define the sample autocorrelation of the signal

$$C(m, n) \equiv \frac{1}{K_c} \sum_{i=0}^{K_c-1} r_{i-m} r_{i-n} \quad \forall m, n$$

Edge conditions are settled by assuming that each symbol follows a similar and independent one.

We also consider another type of DSSS signals, where  $r$  is composed of pseudo-random sequences of  $K_c$  values. The empirical autocorrelation of the input is bounded and the signal has a delta-like autocorrelation:

$$|C(m, n) - \delta(n, m)| \leq \frac{d}{K_c} \quad (4)$$

where  $d$  does not depend on the bandwidth.

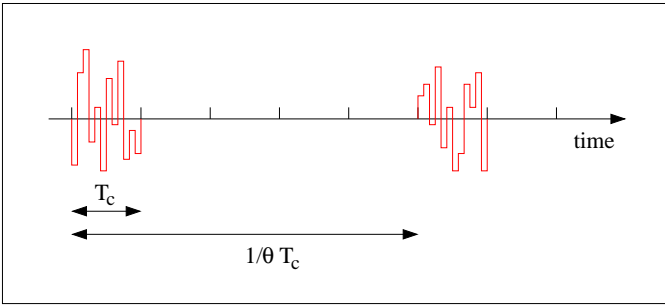


Fig. 2. Direct sequence spread spectrum with duty cycle over coherence periods. The duty cycle parameter  $\theta$  equals the fraction of time used for transmission. The receiver is aware of the active periods of transmission.

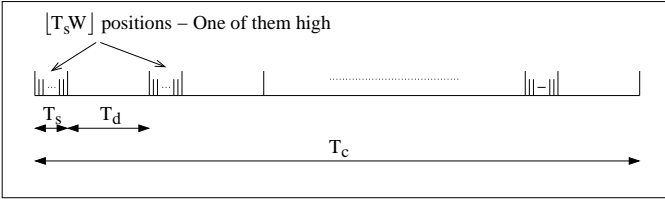


Fig. 3. PPM symbol timing.

### B. PPM Signals

The signals defined in this section are used to calculate lower bounds on the data rates of PPM systems (Section VII-A). The upper bound on PPM performance holds for a wider family of signals, defined in Section VII-B.

Signaling is done over periods  $T_s$  long, with  $[T_s W]$  positions in each symbol, as illustrated in Figure 3. A guard time of  $T_d$  is taken between the symbols, so the symbol period is  $T_s + T_d$ . The symbol time  $T_s$  is typically in the order of the delay spread or smaller. It does not play a significant role in the results. The transmitted signal over a single coherence period is

$$r_i = \begin{cases} \sqrt{W(T_s + T_d)} & \text{one position of each group} \\ & \text{of } [T_s W], \text{ with} \\ & n[(T_s + T_d)W] \leq i \\ & n[(T_s + T_d)W] + [T_s W] - 1 \geq i \\ & n = 0, 1, \dots, \left\lfloor \frac{T_c}{T_s + T_d} \right\rfloor - 1 \\ 0 & \text{other positions} \\ i = 0, 1, \dots, [T_c W] - 1 \end{cases}$$

The number of symbols transmitted over a single coherence period is  $N = \frac{T_c}{T_s + T_d}$ . We assume  $N$  is a whole number, this assumption does not alter the results we prove here. The duty cycle parameter  $0 < \theta \leq 1$  is used over coherence periods: of each period  $\frac{T_c}{\theta}$ , one period of  $T_c$  is used for transmission and in the rest of the time the transmitter is silent.

### III. SUMMARY OF THE RESULTS

A direct sequence spread spectrum system with receiver knowledge of the path delays, can tolerate a sub-linear increase of the number of paths with the bandwidth, and achieve

AWGN capacity (Section VI-A.1). Conversely, if the number of paths increases linearly with the bandwidth, the data rate is sometimes penalized (Sections VI-B).

*Theorem 1: Part 1:* DSSS systems with duty cycle where the receiver knows the path delays (but not necessarily their gains) achieve  $C_{\text{DSSS}} \rightarrow C_{\text{AWGN}}$  as  $W \rightarrow \infty$  if  $\frac{L}{W} \rightarrow 0$ .

*Part 2:* DSSS systems with duty cycle achieve  $C_{\text{DSSS}} \rightarrow C_{\text{AWGN}}$  as  $W \rightarrow \infty$  if  $\frac{L \log W}{W} \rightarrow 0$ . No knowledge of the channel is required.

*Converse 1:* DSSS systems that use pseudo-random sequences that satisfy (4) with duty cycle, where the path gains are unknown to the receiver and uniformly bounded by  $|G_l| \leq \frac{B}{\sqrt{L}}$  with a constant  $B$ , achieve  $C_{\text{DSSS}} < C_{\text{AWGN}}$  in the limit  $W \rightarrow \infty$  if  $\frac{L}{W} \rightarrow \alpha$  and  $\alpha > 0$ . This bound holds whether the receiver knows the path delays or it does not.

The proof is presented in Section VI.

*Theorem 2:* PPM systems with duty cycle, where the receiver knows the path delays, achieve  $C_{\text{PPM}} \rightarrow C_{\text{AWGN}}$  as  $W \rightarrow \infty$  if  $\frac{L}{\log W} \rightarrow 0$  and the path gains  $\{G_l\}_{l=1}^L$  satisfy (A)  $\max_{1 \leq i \leq L} |G_i| \rightarrow 0$  in probability as  $L \rightarrow \infty$ , and (B)  $E_G = \sum_{i=1}^L G_i^2 \rightarrow 1$  in probability as  $L \rightarrow \infty$ . Note that if the gains are IID Gaussians with zero mean then the above condition holds.

*Converse 2:* PPM systems with a non-vanishing symbol time, transmitting over a channel with Gaussian path gains that are unknown to the receiver, achieve  $C_{\text{PPM}} \rightarrow 0$  as  $W \rightarrow \infty$  if  $\frac{L}{\log W} \rightarrow \infty$ . This result holds whether the receiver knows the path delays or it does not.

The proof is presented in Section VII.

### IV. DISCUSSION

This section presents bounds on the data rates of direct sequence spread spectrum and PPM systems for different channels, computed in Sections V, VI, and VII. The channel and system parameters were chosen to represent a realistic low SNR ultra wide band system. For the figures with fixed bandwidth we use: Bandwidth  $W=20$  GHz,  $\frac{P}{N_0}=53$  dB (SNR=-50 dB at  $W=20$  GHz), coherence period  $T_c=0.1$  msec, delay spread  $T_d=200$  nsec, PPM symbol time  $T_s=800$  nsec with guard time of 200 nsec between symbols, and  $B^2 d = 1$ . The constant  $B$  is defined in the converse to Theorem 1, it is used to characterize channel gains, and  $d$  is defined in (4). For the figures with a fixed number of paths we use  $L=100$ .

#### A. The Advantage of Duty Cycle

Figure 4 shows the increase of data rate given by the usage of coherence period duty cycle, for DSSS systems. The figure compares the upper bound on DSSS throughput, where duty cycle is not used (bottom graph) to the lower bound on throughput when optimal duty cycle is used. Both bounds decrease as the number of paths  $L$  increases because the channel uncertainty increases as  $L$  increases, and so does the penalty on the data rate.

#### B. The Duty Cycle Parameter

Figure 5 shows the lower bound on the data rate of a direct sequence spread spectrum system for different duty

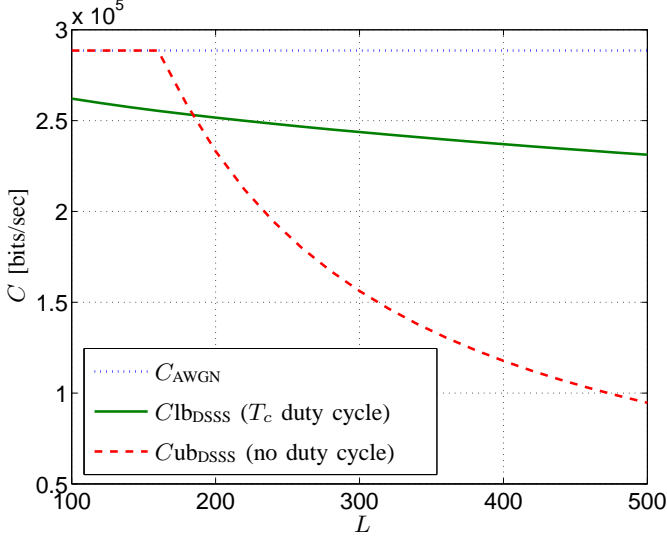


Fig. 4. DSSS throughput bounds, the receiver does not know the channel gains nor the delays, vs. the number of channel paths. This plot contrasts an upper bound on DSSS data rate, when duty cycle is not used (bottom graph, from (18)) with a lower bound on the data rate when coherence period duty cycle is used (top graph, from (12)).  $C_{AWGN}$  is shown for reference.

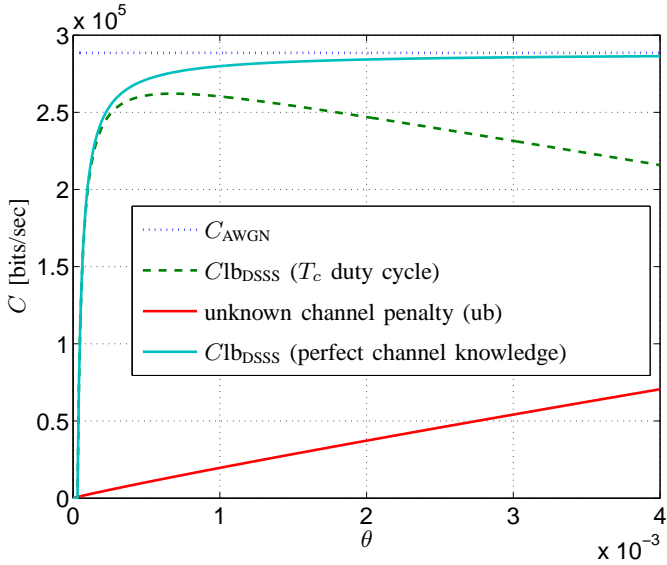


Fig. 5. DSSS throughput vs. duty cycle parameter. The bottom curve shows an upper bound on the channel uncertainty penalty (calculated as the sum of (6) and (5)). The dashed curve shows a lower bound on the system throughput (12), where the receiver does not know the channel path gains nor the delays. The top (full) curve shows the throughput of a system with perfect channel knowledge at the receiver (10).  $C_{AWGN}$  is shown at the dotted curve for reference. The system throughput (dashed curve) is the difference between the data rate of a system with perfect channel knowledge at the receiver (top curve) and the channel uncertainty penalty (bottom curve). The throughput is maximized when the spectral efficiency is balanced against the channel uncertainty penalty.

cycle parameter values. The bound is a difference between the data rate of a system with perfect channel knowledge at the receiver and the channel uncertainty penalty (gain penalty and delay penalty). The data rate of a system with perfect channel knowledge equals the channel capacity ( $C_{AWGN}$ ) in the limit of infinite bandwidth, it is lower when the bandwidth is finite.

The channel uncertainty penalty is small for low values of duty cycle parameter, because the channel is used less often as  $\theta$  decreases. However, the data rate of a system with perfect channel knowledge is severely reduced if  $\theta$  is too low. In this case, transmission occurs with high energy per symbol, where the direct sequence spread spectrum modulation is no longer spectrally efficient, so the data rate with perfect channel knowledge is reduced. Figure 5 shows that the duty cycle parameter must be chosen to balance the channel uncertainty penalty (that is large for large  $\theta$ ) and the spectral efficiency of the selected modulation, that increases with  $\theta$ .

### C. Spectral Efficiency

Figure 6 contrasts the achievable data rates of DSSS and PPM systems, when both use duty cycle on coherence periods with optimal duty cycle parameters. Direct sequence spread spectrum achieves higher data rates because it has a higher spectral efficiency, thus it can pack more bits into each transmission period of length  $T_c$ . By packing bits efficiently a DSSS system is able to use a small duty cycle parameter. In contrast, PPM is less efficient in packing bits into its transmission periods, and is thus forced to transmit more often (it has a larger duty cycle parameter, Figure 7). The PPM system is therefore forced to handle a larger number of channel realizations (per unit time), so it suffers a higher penalty for estimating the channel parameters.

Spectral efficiency, in units of  $\left[\frac{\text{bits}}{\text{sec Hz}}\right]$ , measures the number of bits that can be communicated per unit time per unit bandwidth. The number of bits per DSSS symbol depends linearly on the bandwidth, thus its spectral efficiency does not depend on the bandwidth. The number of bits per PPM symbol (with a fixed symbol time) depends logarithmically on the bandwidth, because PPM is an orthogonal modulation. Thus, the PPM spectral efficiency depends on the bandwidth via  $\frac{\log W}{W}$  and is much lower than the DSSS spectral efficiency if the bandwidth is large.

## V. CHANNEL UNCERTAINTY PENALTY

The mutual information between the transmitted and the received signals is decomposed into the mutual information where the channel is known and a penalty term:

$$I(X; Y) = I(X; Y|D, G) - I(X; D, G|Y)$$

where  $I(X; D, G|Y)$  is the channel uncertainty penalty. This can be decomposed:

$$I(X; D, G|Y) = I(X; D|Y) + I(X; G|D, Y)$$

where  $I(X; D|Y)$  is the *delay uncertainty penalty* and  $I(X; G|D, Y)$  is the *gain uncertainty penalty* when the delays are known.

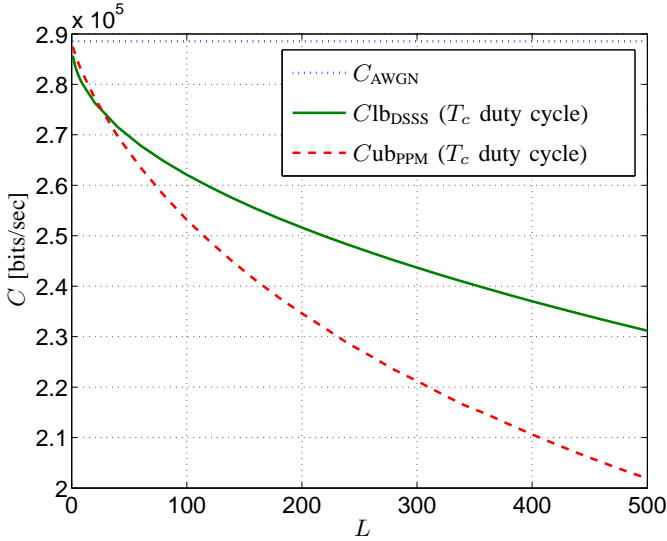


Fig. 6. DSSS and PPM throughput bounds. The DSSS lower bound (12) is calculated without channel knowledge at the receiver. The PPM upper bound (26) is calculated with a receiver that knows the channel delays but not the gains, coherence period duty cycle is used.  $C_{AWGN}$  is shown for reference.

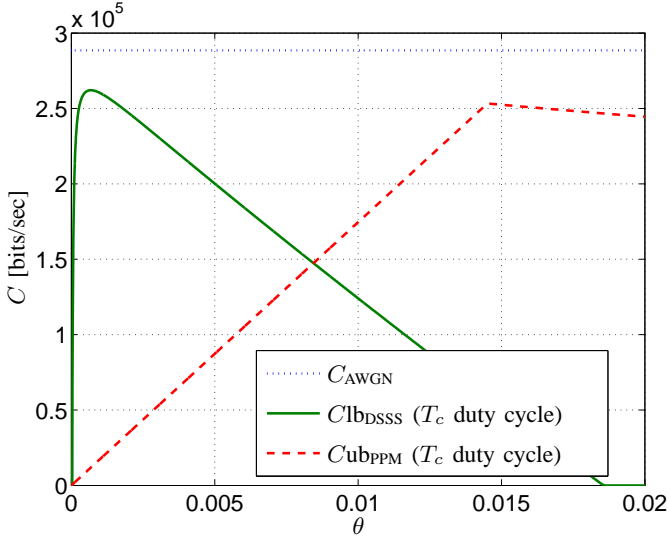


Fig. 7. Throughput bounds vs. duty cycle parameter, the receiver knows the channel path delays but not the gains. DSSS throughput lower bound (12) is maximized at a lower duty cycle parameter than the PPM upper bound (26).

The delay uncertainty penalty is upper bounded by the entropy of the path delays. The  $L$  path delays are uniformly distributed over the  $\binom{WT_d}{L}$  possibilities of combinations of  $L$  values out of  $WT_d$  positions spanning the delay spread.

$$I(X; D|Y) \leq H(D) \leq \frac{\theta L}{T_c} \log_2(WT_d) \quad \left[ \frac{\text{bits}}{\text{sec}} \right] \quad (5)$$

and the gain uncertainty penalty when the delays are known is upper bounded (for DSSS signaling) by

$$I(X; G|Y, D) \leq I(G; Y|X, D) \leq \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{\mathcal{E}}{L} \right) \quad \left[ \frac{\text{bits}}{\text{sec}} \right] \quad (6)$$

see (11) for the derivation.

## VI. SPREAD SPECTRUM BOUNDS

### A. When is the Channel Capacity Achieved?

We start with a result in the case of known path delays (Theorem 1 Part 1) that shows that the channel capacity is achieved if the number of paths is sub-linear with the bandwidth. A second result is then given for the case of unknown channel, (Theorem 1 Part 2), that shows that the channel capacity is still achievable, but at simpler environments, with a smaller number of channel paths.

#### 1) Path Delays Known to Receiver:

*Theorem 1: Part 1:* Direct sequence spread spectrum systems with duty cycle, where the receiver knows the path delays, achieve  $C_{DSSS} \rightarrow C_{AWGN}$  as  $W \rightarrow \infty$  if  $\frac{L}{W} \rightarrow 0$ .

*Proof:* The proof is based on a lower bound on the mutual information.

*Proposition 3:* DSSS systems where the receiver knows the channel path delays achieve

$$I(X; Y|D) \text{ [b/s]} \geq C_{AWGN} \quad (7)$$

$$- \min_{0 < \theta \leq 1} \left\{ \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{P}{N_0} \frac{2T_c}{\theta L} \right) + \frac{3P^2}{N_0^2 \theta W} \log_2 e \right\}$$

*Discussion of Proposition 3:* The channel uncertainty penalty (due to path gains) has two parts, the first

$$\frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{P}{N_0} \frac{2T_c}{\theta L} \right)$$

is the penalty due to the unknown gains, it increases as the number of paths  $L$  or the duty cycle parameter  $\theta$  increase.

The second part of the penalty

$$\frac{3P^2}{N_0^2 \theta W} \log_2 e$$

is due to the limitation on spectral efficiency of the spread spectrum modulation. It penalizes the system for using a too small duty cycle parameter, where the system concentrates too much energy on each transmitted symbol. Mathematically, this term is the quadratic term in the series approximating the mutual information, that is logarithmic. The first (linear) term in this series equals  $C_{AWGN}$  in the limit. The balance between the two penalties is shown in Figure 5.

Looking at the limit of infinite bandwidth, the data rate converges to the AWGN capacity if  $\theta W \rightarrow \infty$  and  $\theta L \rightarrow 0$ . If  $L$  is sub-linear in  $W$ , these two requirements can be met simultaneously, that is, there exists a duty cycle parameter  $\theta$ , that depends on the bandwidth, such that  $\theta L \rightarrow 0$  and  $\theta W \rightarrow \infty$ . Thus, the proof of Theorem 1 Part 1 follows from Proposition 3.

*Proof of Proposition 3:* The proof of (7) follows Theorem 3 of [3], with a real channel instead of the complex channel used there. We start from

$$I(X; Y|D) = I(X; Y|G, D) - I(X; G|Y, D)$$

where the second summand is the gain uncertainty penalty with known delays. By lower bounding this penalty:

$$I(X; G|Y, D) = I(G; Y|X, D) - I(G; Y|D) \leq I(Y; G|X, D) \quad (8)$$

we get

$$I(X; Y|D) \geq I(Y; X|G, D) - I(Y; G|X, D) \quad (9)$$

The first part of (9):

$$I(Y; X|G, D) = \frac{1}{2} E_{G,D} \log \det \left( I + \frac{\mathcal{E}}{WT_c} AA^* \right)$$

where  $A$  is a  $K_c \times K_c$  matrix,  $A_{im} = G_l$  if  $m = (i - D_l)$  and zero otherwise, and  $\mathcal{E} = \frac{2PT_c}{N_0\theta}$ . The eigenvalues of  $AA^*$  are  $\left| F \left( \frac{k}{K_c} \right) \right|^2$ ,  $k = 0, 1, \dots, K_c - 1$ , and

$$F(f) = \sum_{l=1}^L G_l \exp(2\pi j D_l f)$$

For large  $L$  ( $\gg 1$ )  $F(f)$  is complex Gaussian with independent real and imaginary parts that may have different variances (for small  $f$ ), but  $E_G |F(f)|^2 = 1$  and  $E_G |F(f)|^4 \leq 3$ . Although  $\{D_l\}_1^L$  are dependent (as they have  $L$  different values), this dependence is weak and the central limit theorem applies to  $F(f)$ .

$$\begin{aligned} I(Y; X|G, D) &= \frac{1}{2} E_{G,D} \left[ \sum_{k=0}^{WT_c-1} \log \left( 1 + \frac{\mathcal{E}}{WT_c} \left| F \left( \frac{k}{WT_c} \right) \right|^2 \right) \right] \\ &\geq \frac{\log e}{2} E_{G,D} \left[ \sum_{k=0}^{WT_c-1} \frac{\mathcal{E}}{WT_c} \left| F \left( \frac{k}{WT_c} \right) \right|^2 - \frac{1}{2} \frac{\mathcal{E}^2}{W^2 T_c^2} \left| F \left( \frac{k}{WT_c} \right) \right|^4 \right] \\ &\geq \frac{\mathcal{E}}{2} \log e - \frac{3\mathcal{E}^2}{4WT_c} \log e \\ &= \frac{PT_c}{N_0\theta} \log e - \frac{3P^2 T_c}{N_0^2 \theta^2 W} \log e \end{aligned}$$

and in [bits/sec] units:

$$I(Y; X|G, D) \geq \frac{P}{N_0} \log_2 e - \frac{3P^2}{N_0^2 \theta W} \log_2 e \quad (10)$$

where we used  $\log(1+x) \geq \left(x - \frac{x^2}{2}\right) \log e$ , that is valid for  $x \geq 0$ .

The second part of (9):

$$I(Y; G|X, D) \leq \frac{1}{2} E_{X,D} \log \det \left( I + \frac{\mathcal{E}}{WT_c} B\Lambda B^* \right)$$

where  $B_{im} = X_{(i-m)}$  and  $\Lambda = \frac{1}{L} I$ . The upper bound is tight for Gaussian channel gains. Following [3] we get an upper bound

$$I(Y; G|X, D) \leq \frac{L}{2} \log \left( 1 + \frac{\mathcal{E}}{L} \right) \quad (11)$$

Rewriting (9):

$$\begin{aligned} I(X; Y|D) &\geq I(Y; X|G, D) - I(Y; G|X, D) \\ &\geq \frac{\mathcal{E}}{2} \log e - \frac{3\mathcal{E}^2}{4WT_c} \log e - \frac{L}{2} \log \left( 1 + \frac{\mathcal{E}}{L} \right) \end{aligned}$$

in [bits/sec]:

$$\begin{aligned} I(X; Y|D) \text{ [b/s]} &= \frac{\theta I(X; Y|D)}{T_c} \\ &\geq \frac{\theta \mathcal{E}}{2T_c} \log_2 e - \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{\mathcal{E}}{L} \right) \\ &\quad - \frac{3\theta \mathcal{E}^2}{4K_c T_c} \log_2 e \\ &= \frac{P}{N_0} \log_2 e - \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{P}{N_0} \frac{2T_c}{\theta L} \right) \\ &\quad - \frac{3P^2}{N_0^2 \theta W} \log_2 e \end{aligned}$$

The bound is valid for any  $\theta$ , and we choose its maximal value:

$$\begin{aligned} I(X; Y|D) \text{ [b/s]} &\geq \max_{0 < \theta \leq 1} \left\{ \frac{P}{N_0} \log_2 e - \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{P}{N_0} \frac{2T_c}{\theta L} \right) - \frac{3P^2}{N_0^2 \theta W} \log_2 e \right\} \\ &= C_{\text{AWGN}} - \min_{0 < \theta \leq 1} \left\{ \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{P}{N_0} \frac{2T_c}{\theta L} \right) + \frac{3P^2}{N_0^2 \theta W} \log_2 e \right\} \end{aligned}$$

## 2) Path Delays Unknown to Receiver:

*Theorem 1: Part 2:* DSSS systems with duty cycle achieve  $C_{\text{DSSS}} \rightarrow C_{\text{AWGN}}$  as  $W \rightarrow \infty$  if  $\frac{L \log W}{W} \rightarrow 0$ .

*Proof:* The proof is based on Proposition 3 and equation (5) that relates the mutual information in the case of channel knowledge of the path delays with the mutual information in the general case. With no channel knowledge at the receiver (path delays and gains unknown), we get:

$$\begin{aligned} I(X; Y) \text{ [b/s]} &\geq C_{\text{AWGN}} \quad (12) \\ &\quad - \min_{0 < \theta \leq 1} \left\{ \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{P}{N_0} \frac{2T_c}{\theta L} \right) + \frac{3P^2}{N_0^2 \theta W} \log_2 e + \frac{\theta L}{T_c} \log_2 (WT_d) \right\} \end{aligned}$$

The third penalty term describes the penalty due to path delays, from (5). This term is a bound on the penalty, that depends linearly on the number of path delays per unit time.

At the limit of infinite bandwidth the bound equals the AWGN capacity if

- $\theta W \rightarrow \infty$
- $\theta L \rightarrow 0$
- $\theta L \log W \rightarrow 0$

The second condition may be dropped, as the third is stronger. These conditions can be met simultaneously, that is, there exists a duty cycle parameter  $\theta$  that depends on the bandwidth and satisfies the conditions, if  $L \log W$  is sub-linear

in  $W$  namely  $\frac{L \log W}{W} \rightarrow 0$ . Assuming this condition holds, a possible choice for  $\theta$  is  $\frac{1}{\sqrt{WL \log W}}$  so

$$\theta L \log W = \sqrt{\frac{L \log W}{W}}$$

that diminishes, while  $\theta W = \sqrt{\frac{W}{L \log W}}$  that diverges. ■

### B. When is the Channel Capacity Not Achieved?

An additional assumption on gains is used in this section: the gains are uniformly upper bounded by  $|G_l| \leq \frac{B}{\sqrt{L}}$ , this is a technical condition that follows [3].

*Converse to Theorem 1:* DSSS systems that use pseudo-random sequences that satisfy (4) with duty cycle, where the path gains are unknown to the receiver and uniformly bounded by  $|G_l| \leq \frac{B}{\sqrt{L}}$  with a constant  $B$ , achieve  $C_{\text{DSSS}} < C_{\text{AWGN}}$  in the limit  $W \rightarrow \infty$  if  $\frac{L}{W} \rightarrow \alpha$  and  $\alpha > 0$ . This bound holds whether the receiver knows the path delays or it does not.

*Proof:* We first note that the mutual information in a system where the receiver knows the path delays upper bounds the mutual information in the general case:

$$I(X; Y) = I(X; Y|D) - I(D; X|Y) \leq I(X; Y|D)$$

So we only need to prove the theorem regarding the conditional mutual information, where the receiver knows the path delays.

The proof is based on the following upper bound on the mutual information.

*Proposition 4:* Pseudo-random DSSS systems with duty cycle parameter  $\theta$  operating over a channel where  $L/W \xrightarrow{W \rightarrow \infty} \alpha$  for  $\alpha > 0$  achieve

$$I(X; Y|D) \text{ [b/s]} \leq \frac{1}{2} W \theta \log_2 \left( 1 + \frac{2P}{N_0 W \theta} \right) \quad (13)$$

If the duty cycle is chosen so that  $\theta L \rightarrow \infty$  as  $W \rightarrow \infty$ , then a second upper bound holds in the limit:

$$I(X; Y|D) \text{ [b/s]} \leq C_{\text{AWGN}} \frac{T_d B^2 d}{T_c} \quad (14)$$

where  $d$  is defined near (4).

*Discussion of Proposition 4:* We now look at two different possibilities regarding the duty cycle parameter  $\theta$ :

- $\theta W < \infty$ . In this case the bound (13) is strictly lower than  $C_{\text{AWGN}}$ .
- $\theta W \xrightarrow{W \rightarrow \infty} \infty$ . Using our assumption on the number of channel paths we get  $\theta L \rightarrow \infty$ , so the second bound (14) becomes relevant. In situations where  $T_d \ll T_c$ , this bound is significantly lower than the AWGN capacity.

If the number of paths is sub-linear in  $W$ , the duty cycle can be chosen so that  $\theta W \rightarrow \infty$  and  $\theta L \rightarrow 0$  and the bounds in (14) become irrelevant. In this case the upper bound converges to  $C_{\text{AWGN}}$  in the limit of infinite bandwidth.

To summarize the behavior of the bound in the limit, in the case of a linear increase of the number of paths with the bandwidth, the upper bound is lower than the AWGN capacity in normal operating conditions ( $T_d \ll T_c$ ). The upper bound

equals  $C_{\text{AWGN}}$  in the case of a sub-linear increase of the number of paths with the bandwidth.

*Proof of Proposition 4* We start with a simple bound:

$$I(X; Y|D) \text{ [b/s]} \leq \frac{1}{2} W \theta \log_2 \left( 1 + \frac{2P}{N_0 W \theta} \right)$$

This is the capacity of an AWGN channel used with duty cycle  $\theta$ . It upper bounds the data rate for systems with channel knowledge at the receiver. In order to achieve the capacity at the limit of infinite bandwidth, the duty cycle parameter must be chosen so that  $\theta W \rightarrow \infty$ .

To prove (14) we follow Theorem 2 of [3], that gives an upper bound on the mutual information.

$$I(X; Y|D) \leq E_G \log \left( E_H \exp \left[ \mathcal{E} \sum_{l=1}^L H_l G_l \right] \right) + \mathcal{E} \frac{B^2}{L} E_D \left[ \sum_{l=1}^L \sum_{m=1}^L E_X |C(D_l, D_m) - \delta_{lm}| \right] \log_2 e \quad (15)$$

where  $G$  and  $H$  are identically distributed and independent,  $|H_l| \leq \frac{B}{\sqrt{L}}$ .

The first part of (15):

$$\begin{aligned} & \log E_H \exp \left[ \mathcal{E} \sum_{l=1}^L H_l G_l \right] \\ &= \log E_{|H|} E_\psi \exp \left[ \mathcal{E} \sum_{l=1}^L e^{-j\psi_l} |H_l| G_l \right] \end{aligned}$$

The phases  $\{\psi_l\}$  equal 0 or  $\pi$  with probability 1/2.

$$E_\psi \exp(ae^{j\psi_l}) = \frac{e^a + e^{-a}}{2} \approx 1 + \frac{a^2}{2} \quad \text{for } a \ll 1 \quad (16)$$

$$\begin{aligned} & E_{|H|} E_\psi \exp \left[ \mathcal{E} \sum_{l=1}^L e^{-j\psi_l} |H_l| G_l \right] \\ & \leq E_{|H|} \prod_{l=1}^L (1 + \mathcal{E}^2 / 2 |H_l|^2 |G_l|^2) \end{aligned}$$

and the condition in (16) is  $\frac{\mathcal{E} B^2}{2L} \ll 1$ , which holds if  $\theta L \rightarrow \infty$ . We thus bound the first part of (15):

$$\begin{aligned} & E_G \log \left( E_H \exp \left[ \mathcal{E} \sum_{l=1}^L H_l G_l \right] \right) \\ & \leq E_G \log E_{|H|} \prod_{l=1}^L (1 + \mathcal{E}^2 / 2 |H_l|^2 |G_l|^2) \end{aligned}$$

Using Jensen's inequality:

$$\begin{aligned} & E_G \log \left( E_H \exp \left[ \mathcal{E} \sum_{l=1}^L H_l G_l \right] \right) \\ & \leq \sum_{l=1}^L \log (1 + E_{|G_l|, |H_l|} \mathcal{E}^2 / 2 |H_l|^2 |G_l|^2) \\ & = \sum_{l=1}^L \log \left( 1 + \frac{\mathcal{E}^2}{2L^2} \right) \\ & = L \log \left( 1 + \frac{\mathcal{E}^2}{2L^2} \right) \end{aligned}$$

with  $\mathcal{E} = \frac{2PT_c}{N_0\theta}$ :

$$\begin{aligned} & E_G \log \left( E_H \exp \left[ \mathcal{E} \sum_{l=1}^L H_l G_l \right] \right) \\ & \leq L \log \left( 1 + \frac{2P^2 T_c^2}{N_0^2 \theta^2 L^2} \right) \\ & \leq \frac{2P^2 T_c^2}{N_0^2 \theta^2 L} \log e \end{aligned}$$

The second part of (15):

$$\begin{aligned} & \mathcal{E} \frac{B^2}{L} E_D \left[ \sum_{l=1}^L \sum_{m=1}^L E_X |C(D_l, D_m) - \delta_{lm}| \right] \log_2 e \\ & \leq \mathcal{E} B^2 L \frac{d}{K_c} \log_2 e \\ & = \frac{2PL}{N_0 \theta W} B^2 d \log_2 e \\ & \leq \frac{2PT_d}{N_0 \theta} B^2 d \log_2 e \end{aligned} \quad (17)$$

so (17) is valid for pseudo-random chip sequences that satisfy (4).

Putting the two parts back into (15):

$$I(X; Y|D) \leq \frac{2P^2 T_c^2}{N_0^2 \theta^2 L} \log e + \frac{2PT_d}{N_0 \theta} B^2 d \log_2 e$$

In units of [bits/sec]:

$$I(X; Y|D) \text{ [b/s]} \leq \frac{2P^2 T_c}{N_0^2 \theta L} \log_2 e + \frac{2PT_d}{N_0 T_c} B^2 d \log_2 e \quad (18)$$

Using  $\theta L \rightarrow \infty$  we get

$$I(X; Y|D) \text{ [b/s]} \leq \frac{2PT_d}{N_0 T_c} B^2 d \log_2 e = C_{\text{AWGN}} \frac{T_d B^2 d}{T_c}$$

## VII. PPM BOUNDS

### A. When is the Channel Capacity Achieved? (Path Delays Known to Receiver)

*Theorem 2:* PPM systems with duty cycle, where the receiver knows the path delays, achieve  $C_{\text{PPM}} \rightarrow C_{\text{AWGN}}$  as  $W \rightarrow \infty$  if  $\frac{L}{\log W} \rightarrow 0$  and the path gains  $\{G_l\}_{l=1}^L$  satisfy (A)  $\max_{1 \leq i \leq L} |G_i| \rightarrow 0$  in probability as  $L \rightarrow \infty$ , and (B)  $E_G = \sum_{i=1}^L G_i^2 \rightarrow 1$  in probability as  $L \rightarrow \infty$ . Note that if the gains are IID Gaussians with zero mean then the above conditions hold.

*Discussion of the Proof:* The proof starts by breaking the mutual information in two parts:

$$I(X; Y|D) \geq I(X; Y|G, D) - I(Y; G|X, D) \quad (19)$$

The maximal data rate achievable by systems that use the PPM signals defined in Section II-B is the maximum of (19) over the duty cycle parameter  $\theta$ .

$$\begin{aligned} C_{\text{PPM}} &= \max_{\theta} I(X; Y|D) \\ &\geq \max_{\theta} [I(X; Y|G, D) - I(Y; G|X, D)] \end{aligned} \quad (20)$$

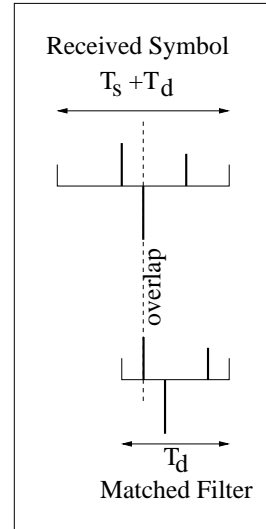


Fig. 8. A received symbol and the matched filter showing a partial overlap.

The first part of (20) describes the throughput of a system that knows the channel perfectly. Section VII-A.2 shows that it approaches  $C_{\text{AWGN}}$  if the duty cycle parameter  $\theta$  is chosen appropriately. This result is shown by demonstrating that the probability of error diminishes as the bandwidth increases while the data rate is as close to  $C_{\text{AWGN}}$  as desired. Our analysis shows not only that the probability of error diminishes, it also demonstrates a specific reception technique, showing achievability of the capacity in the limit. Section VII-A.1 calculates the penalty on the system for its channel gain knowledge and shows that it diminishes for our choice of duty cycle parameter  $\theta$ .

The receiver we use in the analysis of probability of error is based on a matched filter, it is derived from the optimal (maximum likelihood) receiver in the case of a channel with a single path. PPM signals are composed of *orthogonal* symbols. When a PPM signal is transmitted over an impulsive (single path) channel, the orthogonality of the symbols is maintained at the receiver side.

Considering a multipath channel, the received symbol values are no longer orthogonal, so the matched filter receiver is no longer optimal. The non-orthogonality of the received symbols has an adverse effect on receiver performance. As the number of channel paths increases, the received symbols become increasingly non-orthogonal, and the receiver performance is degraded. The matched filter receiver can sustain a growth of the number of channel paths (as the bandwidth increases), but this growth must not be too rapid. To put it more formally, our system achieves the channel capacity in the limit of infinite bandwidth, if the number of paths obeys  $\frac{L^6}{W} \rightarrow 0$ .

For each possible transmitted symbol value, the receiver matches the signal with a filter that has  $L$  fingers at the right delays (Figure 8). The receiver is essentially of the rake type, it constitutes a matched filter to the received signal, that depends on the (known) channel. The receiver uses a threshold parameter  $A = \alpha \sqrt{\mathcal{E}/N}$  for deciding on the transmitted value. If one output only is above  $A$ , the input is guessed according

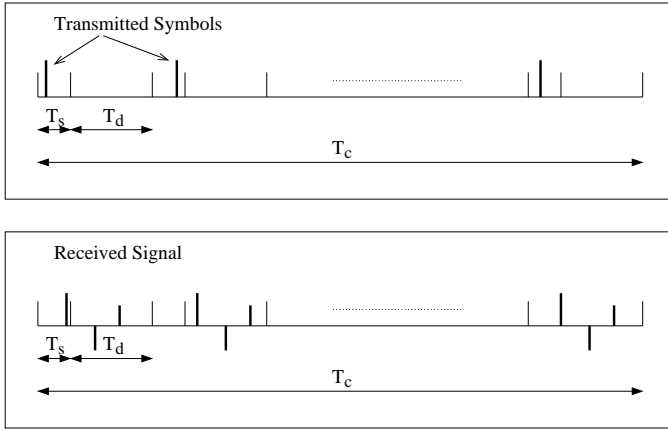


Fig. 9. The transmitted PPM signal, with guard time of  $T_d$  between symbols, and the received signal.

to this output. If none of the outputs pass  $A$ , or there are two or more that do, an error is declared.

We calculate an upper bound on the probability of error of this system, and show that it converges to zero as the bandwidth increases, if the number of channel paths does not increase too rapidly, namely  $\frac{L^6}{W} \rightarrow 0$ , and the duty cycle parameter is chosen properly.

The second part of (20) describes a penalty due to unknown path gains, it is analyzed separately in Section VII-A.1, the upper bound calculated there does not depend on the coding used by the transmitter, and it diminishes as the bandwidth increases for our choice of duty cycle parameter.

We summarize here the conditions for convergence of (20), to bring the conclusion of the following lengthy calculation: The system uses duty cycle with parameter  $\theta$  over coherence periods; The first part of (20) converges to  $C_{AWGN}$  in the limit of infinite bandwidth, if the following conditions take place (end of Section VII-A.2):

- $\theta \log L \rightarrow 0$
- $\frac{L^6}{W} \rightarrow 0$
- $\theta \log W \sim \text{const}$

The second part of (20) contains the penalty for channel (gain) uncertainty; it converges to zero if  $\theta L \rightarrow 0$  (Section VII-A.1). These conditions can exist simultaneously if  $\frac{L}{\log W} \rightarrow 0$ .

*Proof of Theorem 2:* We calculate bounds on the two parts of (19).

1) *Upper Bound on  $I(Y; G|X, D)$ :* The position of the signal fingers is known as  $X$  and  $D$  are known.

$$I(Y; G|X, D) = h(Y|X, D) - h(Y|X, G, D) \quad (21)$$

The first part of (21) is upper bounded by the differential entropy in the case of Gaussian path gains. Considering the signals during a coherence period ( $N$  transmitted symbols), the discretized received signal is composed of  $NM_r$  values (chips), where  $M_r = W(T_s + T_d)$ . Given  $X$  and  $D$ , it is known which chips contain signal and which contain only noise (Figure 9). The  $NM_r$  received values are distributed as follows:

- $N(M_r - L)$  are IID Gaussian  $\sim N(0, 1)$ .

- $NL$  values are divided into groups of size  $N$ . Each group is independent of the other groups, it has zero mean and its correlation matrix is

$$\Lambda = \frac{2P(T_s + T_d)}{\theta N_0 L} + \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

The differential entropy (in bits per coherence time) is bounded by

$$h(Y|X, D) \leq \frac{N(M_r - L)}{2} \log_2(2\pi e) + \frac{L}{2} \log_2((2\pi e)^N |\Lambda|)$$

The determinant  $|\Lambda|$  is the product of the eigenvalues of  $\Lambda$ : 1 with multiplicity  $N - 1$  and  $1 + \frac{\mathcal{E}}{L}$  with multiplicity one

$$|\Lambda| = 1 + \frac{\mathcal{E}}{L}$$

The second part of (21) is given by

$$h(Y|X, D, G) = \frac{NM_r}{2} \log_2(2\pi e)$$

Combining both parts, and translating to units of [bits/sec]:

$$\begin{aligned} I(Y; G|X, D) \text{ [b/s]} &\leq \frac{\theta L}{2T_c} \log_2 |\Lambda| \\ &= \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{\mathcal{E}}{L} \right) \\ &= \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{2P(T_s + T_d)N}{\theta N_0 L} \right) \end{aligned} \quad (22)$$

The bound (22) converges to zero as  $\theta L \rightarrow 0$ .

2) *Lower Bound on  $\max_{\theta} I(Y; X|G, D)$ :* This bound holds if the path gains  $\{G_l\}_{l=1}^L$  satisfy (A)  $\max_{1 \leq i \leq L} |G_i| \rightarrow 0$  in probability as  $L \rightarrow \infty$ , and (B)  $E_G = \sum_{i=1}^L G_i^2 \rightarrow 1$  in probability as  $L \rightarrow \infty$ . We first show that the Gaussian distribution  $G_l \sim N(0, 1/L)$  satisfies these conditions. Condition (B) follows easily from the law of large numbers. To prove that condition (A) holds, we use the following well-known tail estimate for a standard normal  $Z$  and any  $x > 0$ :

$$P(Z > x) \leq \frac{1}{\sqrt{2\pi}x} \exp(-x^2/2) \quad (23)$$

Using  $\beta = 2\sqrt{\frac{\log L}{L}}$  we get

$$\begin{aligned} P\left(\max_{1 \leq l \leq L} |G_l| > \beta\right) &\leq LP(|G_1| > \beta) \\ &\leq LP(|Z| > \beta\sqrt{L}) \\ &\leq \frac{L\sqrt{L}}{\sqrt{2\pi \log L}} \exp(-2 \log L) \\ &\rightarrow 0 \quad \text{as } L \rightarrow \infty \end{aligned}$$

Clearly  $\beta \rightarrow 0$  as  $L \rightarrow \infty$ .

*Analysis of the Signals in the Receiver:* For every symbol, the receiver calculates

$$s_i = \sum_{j=1}^L G_j Y_{i+D_j-1} \quad i = 1, \dots, WT_s$$

Assuming that  $x_1$  was transmitted the desired output is Gaussian with

$$\begin{aligned} E[s_1] &= \sqrt{\mathcal{E}/N} E_G \\ \sigma_{s_1}^2 &= E_G \end{aligned}$$

Note that the expectation here is taken with deterministic  $\{G_j\}$ .

There are up to  $L^2 - L$  Gaussian overlap terms that contain part of the signal (Figure 8). Each of these overlap terms can be described by a set OL of pairs of path gains, each index  $I_{o,o}$  (between 1 and  $\lceil T_d L \rceil$ ) indicates a path in the channel response.

$$\text{OL} = \{(I_{1,1}, I_{1,2}), \dots, (I_{|\text{OL}|,1}, I_{|\text{OL}|,2})\}$$

The number of terms in the set may vary in the range  $1 \leq |\text{OL}| \leq L - 1$  and the indices take values between 1 and  $\lceil T_d W \rceil$ . The pairs of indices are composed of two different indices, because the case where the filter position corresponds to the actual signal position is already accounted for in  $s_1$ .

Given the overlap positions, or the set OL, the overlap terms are Gaussian with

$$\begin{aligned} E[s_{\text{overlap}}|\text{OL}] &= \sum_{i=1}^{|\text{OL}|} G_{I_{i,1}} G_{I_{i,2}} \sqrt{\mathcal{E}/N} \\ E[s_{\text{overlap}}^2|\text{OL}] &\leq \sum_{i=1}^{|\text{OL}|} G_{I_{i,1}}^2 G_{I_{i,2}}^2 \mathcal{E}/N + E_G \end{aligned}$$

Assuming a small number of paths, the probability that there are two or more overlaps (the set OL has more than one element) converges to zero as the bandwidth increases to infinity, see the proof of this convergence in Section VII-A.3. The assumption on the number of paths is satisfied if  $\frac{L^4}{W} \rightarrow 0$ . Additional conditions on  $L$ , stemming from other parts of the proof, will require an even slower increase of  $L$  with the bandwidth.

In addition, there are up to  $WT_s - 1$  Gaussian noise terms:

$$\begin{aligned} E[s_{\text{noise}}] &= 0 \\ E[s_{\text{noise}}^2] &= E_G \end{aligned}$$

For each possible transmitted symbol value, the receiver compares  $\{s_i\}_{i=1}^{WT_s}$  to a threshold  $A = \alpha \sqrt{\mathcal{E}/N} = \alpha \sqrt{\frac{2PT_c}{N_0 \theta N}}$  where  $\alpha \in (0, 1)$ . If one output only is above  $A$ , the input is guessed according to this output. If none of the outputs pass  $A$ , or there are two or more that do, an error is declared.

There are three types of error events, and the error probability is upper bounded using the union bound:

$$\begin{aligned} P(\text{error}) &\leq P(s_1 \leq A) + (L^2 - L)P(s_{\text{overlap}} \geq A) \\ &\quad + (WT_s - 1)P(s_{\text{noise}} \geq A) \\ &\leq P(s_1 \leq A) + L^2 P(s_{\text{overlap}} \geq A) \\ &\quad + WT_s P(s_{\text{noise}} \geq A) \end{aligned} \quad (24)$$

The first probability is bounded using the Chebyshev inequality, and the second and third using the normal tail estimate.

*First Error Event:* Recall  $s_1$  has expectation  $\sqrt{\mathcal{E}/N} E_G$  and variance  $E_G$ . From the Chebyshev inequality,

$$\begin{aligned} P(s_1 \leq A) &\leq \frac{\sigma_{s_1}^2}{(E[s_1] - A)^2} \\ &= \frac{E_G}{(E_G - \alpha)^2 \mathcal{E}/N} \end{aligned}$$

Since  $\alpha < 1$  and  $E_G \rightarrow 1$  in probability, for large  $L$  the ratio  $E_G/(E_G - \alpha)^2$  is bounded. Since  $\mathcal{E}/N \rightarrow \infty$ , the probability converges to 0.

*Second Error Event:* The probability that an overlap term exceeds the threshold is expressed as a sum over the  $L - 1$  possibilities of the number of overlap positions:

$$\begin{aligned} L^2 P(s_{\text{overlap}} \geq A) \\ &= L^2 \sum_{i=1}^{L-1} P(|\text{OL}| = i) P(s_{\text{overlap}} \geq A | |\text{OL}| = i) \end{aligned}$$

Section VII-A.3 shows that if the number of paths is such that  $\frac{L^4}{W} \rightarrow 0$ , then the probability of overlap at more than one position diminishes as  $W \rightarrow \infty$ .

$$P(|\text{OL}| > 1) \rightarrow 0$$

In order to ensure that the overlap terms with more than one overlap position are insignificant in the calculation of the probability of error, we require  $\frac{L^6}{W} \rightarrow 0$ , and then get in the limit of large bandwidth

$$L^2 P(s_{\text{overlap}} \geq A) \rightarrow L^2 P(s_{\text{overlap}} \geq A | |\text{OL}| = 1)$$

The condition  $|\text{OL}| = 1$  is omitted in the remainder of the calculation.

Recall that in the single overlap case  $s_{\text{overlap}}$  is normal with mean  $\mu = G_l G_m \sqrt{\mathcal{E}/N}$  and variance  $E_G = \sum G_i^2$ . Hence  $(s_{\text{overlap}} - \mu)/\sqrt{E_G}$  is a standard normal. By assumption,  $\max |G_i| \rightarrow 0$  and  $E_G \rightarrow 1$  in probability, so for  $L$  large we can assume  $\mu \leq \alpha/2 \sqrt{\mathcal{E}/N} = A/2$  and  $E_G \leq 4$ . Then

$$\begin{aligned} P(s_{\text{overlap}} \geq A) \\ &= P((s_{\text{overlap}} - \mu)/\sqrt{E_G} \geq (A - \mu)/\sqrt{E_G}) \\ &\leq P((s_{\text{overlap}} - \mu)/\sqrt{E_G} \geq A/4) = P(Z \geq A/4) \end{aligned}$$

where  $Z$  stands for a standard normal. Using the normal tail estimate (23), we obtain

$$\begin{aligned} L^2 P(s_{\text{overlap}} \geq A) &\leq L^2 P(Z \geq A/4) \\ &\leq \exp(2 \ln L - A^2/32 - \ln A/2) \end{aligned}$$

In order to ensure convergence to zero of this probability, it is enough to have  $(\ln L)/A^2 \rightarrow 0$ . Recalling  $A = \alpha \sqrt{\mathcal{E}/N} = \alpha \sqrt{\frac{2PT_c}{N_0 \theta N}}$ , we get an equivalent condition:  $\theta \ln L \rightarrow 0$ .

*Third Error Event:* Recall  $s_{\text{noise}}$  is normal with mean 0 and variance  $E_G$ . The third probability in (24) is upper bounded using the normal tail estimate (23) for the standard normal  $Z = s_{\text{noise}}/\sqrt{E_G}$ :

$$P(s_{\text{noise}} \geq A) \leq (2\pi)^{-1/2} (\sqrt{E_G}/A) \exp(-A^2/(2E_G))$$

$$WT_s P(s_{\text{noise}} \geq A) \leq \exp [\ln(WT_s) - A^2/(2E_G) - \ln A]$$

The data rate (in bits/sec) is

$$R = \frac{\theta N}{T_c} \log_2(WT_s) = \frac{\theta N}{T_c} \log_2 e \ln(WT_s)$$

and the capacity  $C_{\text{AWGN}} = P/N_0 \log_2 e$ , so

$$(C_{\text{AWGN}}/R) \ln(WT_s) = (PT_c)/(\theta NN_0) = \mathcal{E}/2N$$

Since  $A^2 = \alpha^2 \mathcal{E}/N$ , we obtain the bound

$$WT_s P(s_{\text{noise}} \geq A) \leq \exp [\ln(WT_s) - (\alpha^2/E_G)(C_{\text{AWGN}}/R) \ln(WT_s) - \ln \frac{A}{\sqrt{E_G}}]$$

Since  $E_G \rightarrow 1$ , the bound converges to 0 as long as  $\alpha^2 > R/C_{\text{AWGN}}$ . This can be achieved for any data rate below the AWGN capacity.

*Achieving Capacity:* So far we have assumed  $\alpha$  is a constant smaller than 1, so the communication system can achieve any rate below  $C_{\text{AWGN}}$ . To achieve asymptotically  $C_{\text{AWGN}}$ , the parameter  $\alpha$  must approach 1 as the bandwidth increases, and the following conditions need to be satisfied:

- $(E_G - \alpha)^2 \log W \rightarrow \infty$  in probability (first error event)
- $(\alpha^2/E_G)(C_{\text{AWGN}}/R) \geq 1$  with probability  $\rightarrow 1$  (third error event)

The exact choice of  $\alpha$  depends on the rate at which  $E_G$  converges to 1.

*Summary of the Bound:* The system uses IID symbols, a duty cycle  $\theta$  and a threshold  $A = \alpha \sqrt{\mathcal{E}/N} = \alpha \sqrt{\frac{2PT_c}{N_0 \theta N}}$  where  $\alpha \in (0, 1)$ .

We calculated an upper bound on the error probability

$$P(\text{error}) \leq \text{upper bound}(W, L, \frac{P}{N_0}, \alpha, \theta)$$

that converges to zero as  $W \rightarrow \infty$  if

- $\frac{L^6}{W} \rightarrow 0$  (second error event)
- $\theta \log W \sim \text{const}$  (to ensure positive rate  $R$ )
- $\theta \log L \rightarrow 0$  (second error event)
- $\theta L \rightarrow 0$  (penalty for unknown gains)

If  $\frac{L}{\log W} \rightarrow 0$  these the conditions can be realized simultaneously, namely it is possible to choose a duty cycle parameter  $\theta \sim \frac{1}{\log W}$  that satisfies all the conditions.

Note that if the path delays are not known, the additional penalty (5) increases as  $\theta L \log W \sim L$ , which diverges, so the above proof is not useful.

3) *Estimation of Number of Overlap Terms:* The number of possible path positions is  $L_m = \lfloor WT_d \rfloor$ . We assume that, over one coherence period, the  $L$  delays are chosen uniformly at random among the  $\binom{L_m}{L}$  possibilities. We prove that if the number of paths grows slowly enough, then with probability converging to 1 there will be at most one overlap between the set of delays and any of its translations.

*Definition 1:* For any set  $S \subset \mathbf{Z}$  and any integer  $t \in \mathbf{Z}$ , we denote by  $S + t$  the translation of  $S$  by  $t$ :  $S + t = \{s + t | s \in S\}$ .

$S$  corresponds to the received symbol when  $x_1$  is transmitted, and  $S + t$  corresponds to  $x_{t+1}$ .

For integers  $1 \leq L \leq L_m$  pick a random subset  $S \subset \{1, \dots, L_m\}$  uniformly among all subsets of  $\{1, \dots, L_m\}$  with  $L$  elements. Let  $\mathbf{P}_{L_m, L}$  be the law of  $S$ ; when there is no ambiguity we drop the subscripts and refer to it as  $\mathbf{P}$ .

*Theorem 5:* Assume  $L^4/L_m \rightarrow 0$  as  $L_m \rightarrow \infty$ , and a set  $S$  is chosen according to  $\mathbf{P}_{L_m, L}$ . Then  $\mathbf{P}_{L_m, L}(|S \cap (S + t)| > 1 \text{ for some } t \neq 0) \rightarrow 0$ .

Note that  $t$  can take both positive and negative values. We emphasize that the theorem says that with high probability, **none** of the translates will have more than one overlap. The proof requires the following

*Lemma 6:* Fix  $t \neq 0$ , and let  $A$  be a set such that  $A \cup (A - t) \subset \{1, \dots, L_m\}$ , and  $|A| \leq L$ . Then

$$\mathbf{P}(A \subset S \cap (S + t)) = [L]_a / [L_m]_a \leq (L/L_m)^a \quad (25)$$

where  $a = |A \cup (A - t)|$  and  $[x]_a = x(x-1)\dots(x-a+1)$ .

*Proof:* Clearly  $A \subset S \cap (S + t)$  is equivalent to  $A \subset S, (A - t) \subset S$ , hence  $S$  has to contain  $A \cup (A - t)$ . Hence  $a$  elements of  $S$  are fixed, while the remaining  $L - a$  ones can be chosen in  $\binom{L_m - a}{L - a}$  ways. The total number of subsets of  $\{1, \dots, L_m\}$  with  $L$  elements is  $\binom{L_m}{L}$ , hence

$$\mathbf{P}(A \subset S \cap (S + t)) = \binom{L_m - a}{L - a} / \binom{L_m}{L} = [L]_a / [L_m]_a$$

The inequality (25) follows easily.  $\blacksquare$

**Note.** If  $A \cup (A - t)$  is not a subset of  $\{1, \dots, L_m\}$ , or if  $|A| > L$ , then the probability (25) is 0.

We are ready to obtain estimates. Fix  $t > 0$ . If  $|S \cap (S + t)| \geq 2$ , then  $S \cap (S + t)$  must contain  $A = \{i, j\}$  for some  $t+1 \leq i < j \leq L_m$ . There are  $\binom{L_m - t}{2}$  such sets  $A$ . Exactly  $L_m - 2t$  of them (namely  $\{i, i + t\}$  for  $t+1 \leq i \leq L_m - t$ ) have  $|A \cup (A - t)| = 3$ ; all the others have  $|A \cup (A - t)| = 4$ . Hence

$$\begin{aligned} & \mathbf{P}(|S \cap (S + t)| \geq 2) \\ & \leq \sum_{t+1 \leq i < j \leq L_m} \mathbf{P}(\{i, j\} \subset S \cap (S + t)) \\ & \leq (L_m - 2t)(L/L_m)^3 + \binom{L_m - t}{2} (L/L_m)^4 \\ & \leq \frac{1}{L_m^2} (L^3 + L^4) \end{aligned}$$

The same estimate holds for  $t < 0$ . Hence

$$\begin{aligned} & \mathbf{P}(|S \cap (S + t)| > 1 \text{ for some } t \neq 0) \\ & \leq \sum_{\substack{-L_m \leq t \leq L_m \\ t \neq 0}} \mathbf{P}(|S \cap (S + t)| > 1) \\ & \leq 2L_m \frac{1}{L_m^2} (L^3 + L^4) \\ & \leq 4L^4/L_m \end{aligned}$$

and the proof of Theorems 5 and 2 is complete.  $\blacksquare$

### B. When is the Channel Capacity Not Achieved?

*Converse to Theorem 2:* PPM systems with a lower bounded symbol time transmitting over a channel with Gaussian path gains that are unknown to the receiver, achieve  $C_{\text{PPM}} \rightarrow 0$  as  $W \rightarrow \infty$  if  $\frac{L}{\log W} \rightarrow \infty$ . This result holds whether the receiver knows the path delays or it does not.

The signals we consider are PPM with symbol time that may depend on the bandwidth, but cannot exceed the coherence period of the channel and cannot diminish (by assumption). The symbol time is divided into positions separated by  $\frac{1}{W}$ . Guard time may be used, no restriction is imposed over it, we use  $T_{\text{sympb}}$  to denote the overall symbol time, that includes the guard time. The signal transmitted over one coherence period is of the form:

$$r_i = \begin{cases} \sqrt{WT_{\text{sympb}}} & \text{one position of each group} \\ & \text{of } \lfloor T_{\text{sympb}}W \rfloor, \text{ with} \\ & n \lfloor T_{\text{sympb}}W \rfloor \leq i \\ & i \leq n \lfloor T_{\text{sympb}}W \rfloor + \lfloor T_{\text{sympb}}W \rfloor - 1 \\ & n = 0, 1, \dots, \lfloor \frac{T_c}{T_{\text{sympb}}} \rfloor - 1 \\ 0 & \text{other positions} \\ i = 0, 1, \dots, \lfloor T_c W \rfloor - 1 \end{cases}$$

The number of symbols transmitted over a single coherence period is  $N = \frac{T_c}{T_{\text{sympb}}}$ . We assume that  $N$  is a whole number, this assumption does not alter the result we prove here. Duty cycle or any other form of infrequent transmission may be used over any time period. We analyze systems that use duty cycle over coherence periods, because this choice yields the highest data rate that serves as an upper bound. The channel is composed of  $L$  paths with independent and identically distributed Gaussian gains, and delays in the range  $[0, T_d)$ .

Edge effects between coherence periods are not considered, they may add a complication to the analysis, without contributing to the understanding of the problem or the solution.

*Outline of the Proof of The Converse to Theorem 2:* The mutual information of the transmitted and received signals is upper bounded by the mutual information when the receiver knows the path delays. This, in turn, is upper bounded in two ways: the first is the PPM transmitted bit rate, and the second is based on the performance of a simple PPM system with no inter-symbol interference. The proof is based on the conditions where the upper bound we calculate on the mutual information diminishes as the bandwidth increases.

*Proof:* We first point out that the mutual information of a system can only increase if the receiver is given information on the path delays:

$$I(X; Y) \leq I(X; Y|D)$$

We calculate an upper bound on PPM mutual information with a real Gaussian multipath channel, in [bits/sec]:

*Proposition 7:*

$$\begin{aligned} I(X; Y|D) \text{ [b/s]} &\leq \max_{0 < \theta \leq 1} \min \{ I_1(\theta), I_2(\theta) \} & (26) \\ I_1(\theta) \text{ [b/s]} &\equiv \frac{\theta \log_2(WT_{\text{sympb}})}{T_{\text{sympb}}} \\ I_2(\theta) \text{ [b/s]} &\equiv \frac{\theta W (T_d + T_{\text{sympb}})}{2T_{\text{sympb}}} \times \\ &\log_2 \left( 1 + \frac{2PT_{\text{sympb}}}{\theta N_0 W (T_d + T_{\text{sympb}})} \right) \\ &- \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{2PT_c}{\theta N_0 L} \right) & (27) \end{aligned}$$

*Discussion of Proposition 7:* The first part of the bound,  $I_1(\theta)$ , is an upper bound on the PPM bit rate for an uncoded system, it is a trivial upper bound on the mutual information.  $\theta$  is the fraction of time used for transmission, and the bound (26) is maximized over the choice of its value. The second part,  $I_2(\theta)$  depends on the number of channel paths  $L$ .

Using Proposition 7, the converse to Theorem 2 follows simply: The bound (26) is positive in the limit  $W \rightarrow \infty$  if both its parts are positive. We note that the symbol time  $T_{\text{sympb}}$  is lower bounded by a constant that does not depend on the bandwidth. The first part,  $I_1(\theta)$ , is positive if the parameter  $\theta$  is chosen so that  $\theta \log(WT_{\text{sympb}}) > 0$ . The second part  $I_2(\theta)$  is positive in the limit of infinite bandwidth if  $\theta L < \infty$ . If the environment is such that  $\frac{L}{\log W} \rightarrow \infty$ , the two conditions involving  $\theta$  cannot be met simultaneously by any choice of fractional transmission parameter. In this case, the bound (26) is zero in the limit of infinite bandwidth.

*Proof of Proposition 7:* The first part of (26) follows simply from the fact that  $I_1(\theta)$  is an upper bound on the transmitted data rate. For any choice of fractional transmission parameter  $\theta$ :

$$I(X; Y|D) \text{ [b/s]} \leq I_1(\theta) \text{ [b/s]}$$

The second part of (26) is proven by comparing the mutual information of our system, with that of a hypothetical system that is easier to analyze. The conditional mutual information  $I(X; Y|D)$  is upper bounded using a hypothetical system that transmits the same symbols as the system we analyze, and receives them without inter-symbol interference (ISI). This is possible, for example, by using many pairs of transmitters and receivers, each pair transmitting symbols with long silence periods between them. The transmitter-receiver pairs are located in such a way that each receiver can ‘hear’ only its designated transmitter. This hypothetical system operates over a channel identical to the one of the original system. The difference between the original system and the hypothetical system is apparent in the number of different noise samples they face, the hypothetical system receives more noise, it processes  $WT_c = WT_c \frac{T_{\text{sympb}} + T_d}{T_{\text{sympb}}}$  positions per coherence period. In spite of this difference, the hypothetical system can achieve higher data rates and its mutual information is an upper bound on the mutual information in the original system. We use  $\tilde{X}$  and  $\tilde{Y}$  to indicate the transmitted and received signals of this system.

$$I(X; Y|D) \text{ [b/s]} \leq I(\tilde{X}; \tilde{Y}|D) \text{ [b/s]}$$

We now prove that for any choice of  $\theta$

$$I(\tilde{X}; \tilde{Y}|D) \text{ [b/s]} \leq I_2(\theta) \text{ [b/s]}$$

Each received symbol in the no-ISI system is composed of  $W(T_{\text{symp}} + T_d)$  chips,  $L$  of them corresponding to the channel paths. All output positions have additive Gaussian noise of variance one. The mutual information is given by

$$I(\tilde{X}; \tilde{Y}|D) \text{ [b/s]} = \frac{\theta}{T_c} \left[ H(\tilde{Y}|D) - H(\tilde{Y}|\tilde{X}, D) \right] \quad (28)$$

We start with the first part of (28):

$$H(\tilde{Y}|D) \leq \sum_{i=1}^{W\tilde{T}_c} H(\tilde{Y}_i|D) \leq \sum_{i=1}^{W\tilde{T}_c} H(\tilde{Y}_i)$$

$$\tilde{Y}_i \sim \begin{cases} N\left(0, 1 + \frac{2PT_{\text{symp}}}{\theta N_0 L}\right) & \text{prob } p_i \\ N(0, 1) & \text{prob } 1 - p_i \end{cases}$$

$$i = 1, \dots, W\tilde{T}_c$$

$p_i$ , the probability of receiving signal energy in the  $i^{\text{th}}$  position, depends on the distribution of transmitted symbols, but there are exactly  $L$  positions in the received symbol that contain a path, thus

$$\sum_{i=1}^{W(T_d+T_{\text{symp}})} p_i = L$$

and each probability value satisfies  $0 \leq p_i \leq 1$ .

$$\begin{aligned} \sigma_{\tilde{Y}_i}^2 &= E[\tilde{Y}_i^2] = 1 - p_i + p_i \left(1 + \frac{2PT_{\text{symp}}}{\theta N_0 L}\right) \\ &= 1 + p_i \frac{2PT_{\text{symp}}}{\theta N_0 L} \end{aligned}$$

$$\begin{aligned} H(\tilde{Y}_i) &\leq \frac{1}{2} \log(2\pi e \sigma_{\tilde{Y}_i}^2) \\ &= \frac{1}{2} \log\left(2\pi e \left(1 + p_i \frac{2PT_{\text{symp}}}{\theta N_0 L}\right)\right) \end{aligned}$$

$$\begin{aligned} H(\tilde{Y}|D) &\leq \sum_{i=1}^{W\tilde{T}_c} H(\tilde{Y}_i) \\ &\leq \sum_{T_c/T_{\text{symp}} \text{ symbols}} \\ &\quad \sum_{i=1}^{W(T_d+T_{\text{symp}})} \frac{1}{2} \log\left(2\pi e \left(1 + p_i \frac{2PT_{\text{symp}}}{\theta N_0 L}\right)\right) \\ &= \frac{W\tilde{T}_c}{2} \log(2\pi e) \\ &\quad + \sum_{\text{symbols}} \sum_{i=1}^{W(T_d+T_{\text{symp}})} \frac{1}{2} \log\left(1 + p_i \frac{2PT_{\text{symp}}}{\theta N_0 L}\right) \\ &= \frac{W\tilde{T}_c}{2} \log(2\pi e) \\ &\quad + \frac{T_c}{2T_{\text{symp}}} \sum_{i=1}^{W(T_d+T_{\text{symp}})} \log\left(1 + p_i \frac{2PT_{\text{symp}}}{\theta N_0 L}\right) \end{aligned}$$

Using the concavity of the log we get

$$\begin{aligned} H(\tilde{Y}|D) &\leq \frac{W\tilde{T}_c}{2} \log(2\pi e) \\ &\quad + \frac{WT_c(T_d + T_{\text{symp}})}{2T_{\text{symp}}} \times \\ &\quad \log\left(1 + \frac{2PT_{\text{symp}}}{\theta N_0 W(T_d + T_{\text{symp}})}\right) \end{aligned}$$

Now for the second part of (28). For  $N$  transmitted symbols, the  $W\tilde{T}_c$  received values are distributed as follows, when the input  $\tilde{X}$  and the delays  $D$  are known:

- $W\tilde{T}_c - NL$  positions are IID Gaussians  $\sim N(0, 1)$ . The receiver knows which positions contain only noise, and which have signal as well.
- The number of positions with some signal is  $NL$ . These values are divided into groups of size  $N$ , each corresponding to a single path. Each group (at known positions) is independent of the other groups and its distribution is  $\sim N(0, \Lambda)$  where

$$\Lambda = \frac{2PT_{\text{symp}}}{\theta N_0 L} + \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

$$\begin{aligned} -H(\tilde{Y}|\tilde{X}, D) &= -\frac{W\tilde{T}_c - NL}{2} \log(2\pi e) \\ &\quad - \frac{L}{2} \log((2\pi e)^N |\Lambda|) \end{aligned} \quad (29)$$

The determinant  $|\Lambda|$  is the product of the eigenvalues of  $\Lambda$ : 1 with multiplicity  $N - 1$  and  $\left(1 + \frac{2PT_{\text{symp}}N}{\theta N_0 L}\right)$  with multiplicity one

$$|\Lambda| = 1 + \frac{2PT_{\text{symp}}N}{\theta N_0 L} = 1 + \frac{2PT_c}{\theta N_0 L}$$

Simple manipulation yields

$$\begin{aligned} -H(\tilde{Y}|\tilde{X}, D) &= -\frac{W\tilde{T}_c}{2} \log(2\pi e) \\ &\quad - \frac{L}{2} \log\left(1 + \frac{2PT_c}{\theta N_0 L}\right) \end{aligned}$$

We note that this expression depends on  $\text{SNR}_{\text{est}}$ ; It is similar to expressions in [3], Section III.C.

Combining the two parts into (28) we get

$$\begin{aligned}
I(X; Y|D) \text{ [b/s]} &\leq I(\tilde{X}; \tilde{Y}|D) \text{ [b/s]} \\
&= \left[ H(\tilde{Y}|D) - H(\tilde{Y}|\tilde{X}, D) \right] \frac{\theta}{T_c} \\
&\leq \left[ \frac{W\tilde{T}_c}{2} \log_2(2\pi e) \right. \\
&\quad + \frac{WT_c(T_d + T_{\text{symp}})}{2T_{\text{symp}}} \times \\
&\quad \left. \log_2 \left( 1 + \frac{2PT_{\text{symp}}}{\theta N_0 W (T_d + T_{\text{symp}})} \right) \right. \\
&\quad \left. - \frac{W\tilde{T}_c}{2} \log_2(2\pi e) - \right. \\
&\quad \left. \frac{L}{2} \log_2 \left( 1 + \frac{2PT_c}{\theta N_0 L} \right) \right] \frac{\theta}{T_c} \\
&= \frac{\theta W (T_d + T_{\text{symp}})}{2T_{\text{symp}}} \times \\
&\quad \log_2 \left( 1 + \frac{2PT_{\text{symp}}}{\theta N_0 W (T_d + T_{\text{symp}})} \right) \\
&\quad - \frac{\theta L}{2T_c} \log_2 \left( 1 + \frac{2PT_c}{\theta N_0 L} \right) \\
&= I_2(\theta)
\end{aligned}$$

Note that in the infinite bandwidth limit, if  $\theta L \rightarrow \infty$  than also  $\theta W \rightarrow \infty$  and  $I_2(\theta)$  converges to zero:

$$\begin{aligned}
I_2(\theta) &\longrightarrow \\
&\frac{\theta}{2T_{\text{symp}}} \sum_{i=1}^{W(T_d+T_{\text{symp}})} p_i \frac{2PT_{\text{symp}}}{\theta N_0 L} \log_2 e \\
&\quad - \frac{\theta L}{2T_c} \frac{2PT_c}{\theta N_0 L} \log_2 e \\
&= 0
\end{aligned}$$

### VIII. CONCLUSION

This paper revealed the importance of the rate of increase of the apparent number of paths as the bandwidth increases. This rate of increase determines the behavior of channel uncertainty as bandwidth increases, thus determining the spectral efficiency required of a communication system, in order to achieve channel capacity in the limit of large bandwidth. Practical mobile communication systems handle channel uncertainty by occasional estimation of the channel, usually assisted by pilot (training) signals. The rate of repeating the pilot signal is determined by the rate of channel variation and the required accuracy of estimation. This paper showed that systems with a low spectral efficiency, and in particular PPM systems, have to sacrifice a significant portion of their data-rate in order to obtain channel estimation. When operating over difficult channels, namely channels where the number of paths increases quickly with the bandwidth, a PPM system has to devote an increasing fraction of its resources to channel estimation, until in the limit of infinite bandwidth it cannot communicate.

The effect of channel uncertainty depends on the data rate of the communication system, it is pronounced in systems that attempt to approach the channel capacity. Intuitively, a slow system has ‘spare’ resources that are not used for communication and can be exploited for channel estimation, making the system robust to some channel uncertainty. The effective signal to noise ratio  $\text{SNR}_{\text{est}} = \frac{2PT_c}{N_0\theta L}$  clarifies the effect of channel uncertainty on systems that attempt to approach capacity, as it compares the channel uncertainty per unit time  $\frac{\theta L}{T_c}$  to the capacity in the limit of infinite bandwidth that is proportional to  $\frac{P}{N_0}$ . A communication system can achieve the channel capacity in the limit of infinite bandwidth only if the channel uncertainty per unit time becomes insignificant relative to the capacity or

$$\text{SNR}_{\text{est}} \rightarrow \infty$$

In systems with bounded average received power, the duty cycle parameter must diminish in order to balance the increase in the number of channel components  $L$ , and let the overall channel uncertainty diminish.

The key to understanding wideband operation over a radio channel is the behavior of channel uncertainty as bandwidth increases. The channel model we used is a simplistic one, where the channel paths are statistically similar so each represents the same amount of channel uncertainty. Realistic channel models are more complex, so it is harder to characterize their channel uncertainty and its dependence on system bandwidth. Wideband radio channel models show a spread of the impulse response energy over a number of paths, that may not be of equal amplitude. The dependence of channel uncertainty on the bandwidth depends on the rate of increase of the number of significant channel paths as the bandwidth increases, and we expect similar results to ours, where the number of significant paths replaces the number of paths  $L$  we used.

### ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their careful reading of our paper and helpful suggestions.

### APPENDIX

This appendix shows that a system that operates over a channel composed of paths with uncorrelated amplitudes, and uses a rectangular chip filter and a matched filter as discussed in Section II, has an effective channel model where the resolvable paths are uncorrelated.

Define

- $p(t)$  the rectangular chip shaping pulse,
- $R_p(\tau) = \int p(t)p(t+\tau)dt$  the autocorrelation function of  $p(t)$ ,
- $h(\tau)$  the channel impulse response,
- $P_h(\tau) = E[|h(\tau)|^2]$  the delay power spectrum of the channel, and
- $g(\tau) = p(\tau) \star h(\tau) \star p(-\tau) = h(\tau) \star R_p(\tau)$  the concatenated system consisting of the shaping pulse in the transmitter side, the channel and the matched filter in the receiver.

Assuming that the channel paths are uncorrelated we have

$$E[h(\tau)h(\tau + \Delta\tau)] = P_h(\tau)\delta(\Delta\tau)$$

where  $\delta(\bullet)$  denotes the Dirac impulse. The concatenated system  $g(\tau)$  has the following correlation function:

$$R_g(\tau_1, \tau_2) = E[g(\tau_1)g(\tau_2)] = \int P_h(\tau)R_p(\tau_1 - \tau)R_p(\tau_2 - \tau)d\tau$$

The equivalent discrete time channel response results by sampling the concatenated response  $g(\tau)$  with sampling rate  $W$ . Define the equivalent path amplitudes as  $A_l = g(l/W)$ . Then

$$E[A_l A_{l'}] = \left[ \int P_h(\tau)R_p^2(l/W - \tau) d\tau \right] \delta_{l-l'}$$

where  $\delta_{\bullet}$  denotes the Kronecker delta sequence. Thus, the resolvable path amplitudes are uncorrelated. To a good approximation this property can be also assumed to hold if  $p(\tau)$  is a raised-cosine pulse.

## REFERENCES

- [1] R. S. Kennedy, *Fading Dispersive Communication Channels*. Wiley & Sons, 1969.
- [2] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley & Sons, 1968, section 8.6.
- [3] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1384–1400, July 2000.
- [4] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Transactions on Information Theory*, vol. 2002, no. 4, pp. 840–852, April 2002.
- [5] V. G. Subramanian and B. Hajek, "Broad-band fading channels: Signal burstiness and capacity," *IEEE Transactions on Information Theory*, vol. 48, no. 4, pp. 809–827, April 2002.
- [6] L. Rusch, C. Prettie, D. Cheung, Q. Li, and M. Ho, "Characterization of UWB propagation from 2 to 8 GHz in a residential environment," *IEEE Journal on Selected Areas in Communications*.
- [7] M. Pendergrass, "Empirically based statistical ultra-wideband model," *IEEE 802.15 Working Group for Wireless Area Networks (WPANs)*, July 2002, contribution 02/240, available at [http://grouper.ieee.org/groups/802/15/pub/2002/Jul02/02240r1p802-15\\_SG3a-Empirically-Based\\_UWB\\_Channel\\_Model.ppt](http://grouper.ieee.org/groups/802/15/pub/2002/Jul02/02240r1p802-15_SG3a-Empirically-Based_UWB_Channel_Model.ppt).
- [8] S. Verdú, "On channel capacity per unit cost," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 1019–1030, September 1990.
- [9] —, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.
- [10] L. Zheng, D. Tse, and M. Médard, "Channel coherence in the low snr regime," in *International Symposium on Information Theory (ISIT)*. IEEE, June-July 2004, p. 413.
- [11] —, "On the costs of channel state information," in *Information Theory Workshop (ITW)*. IEEE, October 2004, pp. 423–427.

**David N. C. Tse** received the B.A.Sc. degree in systems design engineering from University of Waterloo, Canada in 1989, and the M.S. and Ph.D. degrees in electrical engineering from Massachusetts Institute of Technology in 1991 and 1994 respectively. From 1994 to 1995, he was a postdoctoral member of technical staff at A.T.&T. Bell Laboratories. Since 1995, he has been at the Department of Electrical Engineering and Computer Sciences in the University of California at Berkeley, where he is currently a Professor. He received a 1967 NSERC 4-year graduate fellowship from the government of Canada in 1989, a NSF CAREER award in 1998, the Best Paper Awards at the Infocom 1998 and Infocom 2001 conferences, the Erlang Prize in 2000 from the INFORMS Applied Probability Society, the IEEE Communications and Information Theory Society Joint Paper Award in 2001, and the Information Theory Society Paper Award in 2003. He was the Technical Program co-chair of the International Symposium on Information Theory in 2004, and was an Associate Editor of the IEEE Transactions on Information Theory from 2001 to 2003. He is a coauthor, with Pramod Viswanath, of the text "Fundamentals of Wireless Communication". His research interests are in information theory, wireless communications and networking.

**Șerban Nacu** received his B. A. in mathematics from Harvard University in 1996 and his Ph. D. in statistics from University of California, Berkeley in 2004. During 2004-2006 he was a postdoctoral researcher at Stanford University, and is currently with the Ecole Normale Supérieure in Paris, France. He works in various areas of probability theory and applied statistics.

**Dana Porrat** was born in Jerusalem, Israel. Received the B.Sc. in Electrical Engineering from the Technion - Israel Institute of Technology in 1991, the M.Sc. from Tel Aviv University in 1997 and the Ph.D. from Stanford University in 2002.

In 2005 she returned to Jerusalem and joined the Engineering School of the Hebrew University where she is now a lecturer. Her research interests include wideband radio, channel characterization and modeling and the information theoretic aspects of communications over unknown channels.