IPN Progress Report 42-207 • November 15, 2016

Repeat-PPM Super-Symbol Synchronization

Joseph Connelly*

ABSTRACT. — To attain a wider range of data rates in pulse position modulation (PPM) schemes with constrained pulse durations, the sender can *repeat* a PPM symbol multiple times, forming a *super-symbol*. In addition to the slot and symbol synchronization typically required for PPM, the receiver must also properly align the noisy super-symbols.

We present a low-complexity approximation of the maximum-likelihood method for performing *super-symbol synchronization* without use of synchronization sequences. We provide simulation results demonstrating performance advantage when PPM symbols are spread by a pseudo-noise sequence, as opposed to simply repeating. Additionally, the results suggest that this super-symbol synchronization technique requires signal levels below those required for reliable communication. This validates that the PPM spreading approach proposed to CCSDS can work properly as part of the overall scheme.

I. Introduction

The prospective Consultative Committee for Space Data Systems (CCSDS) recommendation for high photon efficiency (HPE) optical communications [1, 2] specifies transmitting data through an optical channel using pulse position modulation (PPM). *Pulse position modulation* is a technique for transmitting digital data, in which a period of time is divided into M possible *slots*, and a binary sequence of length $\log_2 M$ is encoded to a PPM *symbol* by transmitting a single pulse in exactly one of the M time slots, where M is typically taken to be a power of two. In the case of HPE optical communications, this corresponds to emitting a laser pulse in one of the time slots. PPM is known to approach the capacity of an optical channel under certain conditions [3] and is well suited for HPE optical communications, as it requires lower average power than other modulation techniques [4].

---

*Communications Architectures and Research Section.

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

Two significant synchronization problems arise when using PPM: (1) *Slot synchronization*, in which the receiver attempts to determine the boundaries of the PPM slots. (2) *Symbol synchronization*, in which the receiver attempts to determine which slots begin PPM symbols. Slot and symbol synchronization have been extensively studied (e.g., [5, 6, 7, 8, 9]), particularly for the optical channel.[2] In this article, we describe a solution for a different synchronization issue that can arise when PPM symbols are sequentially repeated prior to transmission.

A. Repeat-PPM

Throughout this article, we assume the use of a detector that counts the number of photons received in a duration of time, and we further assume that slot synchronization has been obtained. The output of such a detector is well modeled as a memoryless Poisson process with mean K_b in a noise slot and with mean $K_s + K_b$ in a signal slot [4]. We also assume symbol synchronization has been obtained so that the receiver can correctly group together slots into noisy PPM symbols.

When the slot duration is increased by a factor of N and all other parameters are held constant (e.g., intensity of the laser), both K_s and K_b are increased by a factor of N. However, due to power constraints or other system limitations, the transmitter may not be able to generate a pulse in the entire duration of the desired PPM time slot.

The same slot photon count distribution can be achieved by transmitting a PPM symbol N times and *collapsing* the received *super-symbol* by summing the respective slot photon counts of each corresponding symbol. Since the channel is memoryless, the slot photon count of a collapsed super-symbol is a sum of N independent Poisson random variables, each with the same mean. Hence, when a PPM symbol is transmitted N times, the slot photon count of each collapsed super-symbol is itself a Poisson random variable with mean NK_b in a noise slot and with mean $N(K_s + K_b)$ in a signal slot.

The proposed CCSDS recommendation for HPE optical communication uses repeat-PPM in both the uplink (ground-to-space) and the downlink (space-to-ground) schemes to achieve effective slot durations not otherwise possible due to the available pulse durations of the lasers [1, 2]. This allows for a wider range of attainable data rates without requiring the laser to accommodate a wide range of pulse durations.

However, repeat-PPM requires additional synchronization in order to operate effectively. When slot and symbol synchronization are obtained, the receiver must also determine when sequences of repeated PPM symbols begin so that the noisy super-symbols are properly aligned with the transmitted sequence. In particular, when each PPM symbol is sequentially repeated N times, the receiver must determine which noisy symbols correspond to the beginnings of sequences of N-repeated PPM symbols. We refer to

---

[2] Ryan Rogalin and Meera Srinivasan, "Maximum-likelihood synchronization for pulse position modulation with inter-symbol guard times," submitted to *IEEE Global Communications Conference*, 2016.

this as *super-symbol synchronization.*

B. Overview

The remainder of the article is outlined as follows. In Section II, we further discuss repeat-PPM and show that maximum-likelihood (ML) detection of repeat-PPM can fail when super-symbol synchronization is lost. We also discuss a generalization of repeat-PPM in which signals are spread by a pseudo-noise (PN) sequence. In Section III, we introduce the system model and explicitly define the super-symbol synchronization problem. In Section IV, we derive a technique for obtaining super-symbol synchronization based on an approximation of the ML method. In Section V, we provide some analysis of the estimator, suggesting that in some cases, PN spreading can improve performance. In Section VI, we present simulation results of our super-symbol synchronization method comparing repeating and spreading for the PPM orders and repeat factors specified in the proposed CCSDS standards. In Section VII, we discuss the role of super-symbol synchronization in the proposed CCSDS recommendation for HPE optical communication.

II. Detecting Repeat-PPM

It is known (e.g., [10]) that the ML estimate of a PPM symbol in the presence of noise is to choose the slot with the maximum photon count as the signal slot. In this section, we show that when super-symbol synchronization is obtained, selecting the collapsed super-symbol slot with the largest photon count is the ML estimate of a repeated PPM symbol.

Suppose that the M-ary PPM symbol $x \in \{0, 1, \ldots, M-1\}$ is transmitted N times through an optical channel and that the receiver has obtained slot and symbol synchronization. For each $i \in \{0, 1, \ldots, N-1\}$ and $j \in \{0, 1, \ldots, M-1\}$, let $r_{i,j}$ be the number of photons in the jth slot of the ith noisy symbol

$$\boldsymbol{r}_i = (r_{i,0}, r_{i,1}, \ldots, r_{i,M-1}).$$

When x is transmitted and super-symbol synchronization is obtained, the distribution of $r_{i,j}$ is given by

$$\begin{cases} \mathrm{Poi}\left(K_s + K_b\right), & \text{if } j = x \\ \mathrm{Poi}\left(K_b\right), & \text{if } j \neq x \end{cases}$$

where $\mathrm{Poi}\left(\lambda\right)$ denotes a Poisson distribution with mean λ. The ML estimate of x is given by

$$\hat{x} = \arg\max_{y} \; P(\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_{N-1} \,|\, y \text{ is transmitted}).$$

Since the channel is memoryless, each slot is independently distributed given the M-ary PPM symbol y is transmitted, so

$$
\begin{aligned}
\hat{x} &= \arg\max_y \prod_{i=0}^{N-1} \prod_{j=0}^{M-1} P(r_{i,j} \,|\, y \text{ is transmitted}) \\
&= \arg\max_y \prod_{i=0}^{N-1} \frac{e^{-(K_s+K_b)}(K_s+K_b)^{r_{i,y}}}{r_{i,y}!} \prod_{\substack{j=0\\ j\neq y}}^{M-1} \frac{e^{-K_b}K_b^{r_{i,j}}}{r_{i,j}!} \\
&= \arg\max_y \prod_{i=0}^{N-1} \frac{e^{-(K_s+K_b)}(K_s+K_b)^{r_{i,y}}}{r_{i,y}!} \left(\frac{e^{-K_b}K_b^{r_{i,y}}}{r_{i,y}!}\right)^{-1} \prod_{j=0}^{M-1} \frac{e^{-K_b}K_b^{r_{i,j}}}{r_{i,j}!} \\
&= \arg\max_y \prod_{i=0}^{N-1} \left(\frac{K_s}{K_b}+1\right)^{r_{i,y}} \\
&= \arg\max_y \sum_{i=0}^{N-1} r_{i,y}
\end{aligned}
$$

which is precisely the collapsed super-symbol slot with the largest photon count.

A. Loss of Synchronization

When super-symbol synchronization is not obtained, the noisy super-symbols potentially consist of noisy symbols with distinct signal slots, which can result in incorrect detection. In the following example, a sequence of noisy symbols decodes differently depending on the alignment of the super-symbols.

Suppose the 4-ary PPM sequence $0, 1, 3, 1$ is transmitted with a repeat factor $N = 2$, resulting in the transmitted PPM symbol sequence

$$0, 0, 1, 1, 3, 3, 1, 1$$

and the detector outputs a sequence of photon slot counts, which are grouped into noisy PPM symbols as follows:

$$[0, 2, 1, 0]\,[3, 1, 0, 1]\,[3, 1, 2, 0]\,[0, 3, 2, 0]\,[0, 4, 0, 2]\,[1, 1, 0, 3]\,[2, 0, 0, 4]\,[1, 4, 1, 0]\,[0, 3, 2, 0].$$

The first symbol is, in fact, entirely noise. This could also be viewed as the receiver starting synchronization in the middle of a sequence of symbols, in which case, $[0, 2, 1, 0]$ corresponds to a previously transmitted symbol.

The receiver knows four PPM symbols are sent with a repeat factor of 2 but does not know how the super-symbols align. The receiver could form super-symbols in one of two ways:

$$\underbrace{[0, 2, 1, 0]\,[3, 1, 0, 1]}_{[3,3,1,1]}\,\underbrace{[3, 1, 2, 0]\,[0, 3, 2, 0]}_{[3,4,4,0]}\,\underbrace{[0, 4, 0, 2]\,[1, 1, 0, 3]}_{[1,5,0,5]}\,\underbrace{[2, 0, 0, 4]\,[1, 4, 1, 0]}_{[3,4,1,4]}\,[0, 3, 2, 0]$$

or

$$[0, 2, 1, 0]\,\underbrace{[3, 1, 0, 1]\,[3, 1, 2, 0]}_{[6,2,2,1]}\,\underbrace{[0, 3, 2, 0]\,[0, 4, 0, 2]}_{[0,7,2,2]}\,\underbrace{[1, 1, 0, 3]\,[2, 0, 0, 4]}_{[3,1,0,7]}\,\underbrace{[1, 4, 1, 0]\,[0, 3, 2, 0]}_{[1,7,3,0]}.$$

Using ML detection, the first sequence of super-symbols does not decode uniquely. However, the super-symbols in the second case correctly yield the sequence $0, 1, 3, 1$.

In general, when PPM symbols are repeated N times and super-symbol synchronization is lost, the receiver begins detection on the nth repeat of a symbol, where the *offset* n is an integer between 0 and $N - 1$. If the receiver does not account for the offset, the super-symbols consist of $(N - n)$ noisy symbols with a particular signal slot and n symbols with a potentially disparate signal slot.

We remark that in the previous example, the collapsed super-symbols formed in the second case have more distinctive maximum slot counts. The estimator described in Section IV estimates the offset by determining which alignment yields the collapsed super-symbols with the largest maximum slot counts. In the first case, summing the maximum slot counts of each collapsed super-symbol yields $3 + 4 + 5 + 4 = 16$, while in the second case, the sum is $6 + 7 + 7 + 7 = 27$. Using this technique for determining the offset, the receiver would correctly choose the second case, which corresponds to an offset of $n = 1$.

B. Pseudo-Noise Spreading

We study a generalization of repeat-PPM in which the PPM symbols are first spread by a pseudo-noise sequence, then at the receiver, noisy super-symbols are "unspread" by the same PN sequence. The receiver can decode a PN-spread sequence similarly to a repeated sequence; however, the receiver must undo the PN-spreading when collapsing super-symbols. PN spreading is equivalent to repeat-PPM followed by reordering the slots.

When the M-ary PPM symbol x is *spread* by the M-ary PN sequence $p_0, p_1, \ldots, p_{N-1}$, the sender transmits the PPM sequence

$$(x + p_0 \bmod M),\ (x + p_1 \bmod M),\ \ldots,\ (x + p_{N-1} \bmod M).$$

The special case where $p_0 = p_1 = \cdots = p_{N-1} = 0$ is equivalent to repeat-PPM.

For each $i \in \{0, 1, \ldots, N - 1\}$ and $j \in \{0, 1, \ldots, M - 1\}$, let $s_{i,j}$ be the number of photons in the jth slot of the ith noisy symbol

$$\boldsymbol{s}_i = (s_{i,0}, s_{i,1}, \ldots, s_{i,M-1}).$$

When x is the repeated symbol and super-symbol synchronization is obtained, $s_{i,j}$ is distributed according to

$$\begin{cases} \text{Poi}\left(K_s + K_b\right), & \text{if } j = x + p_i \bmod M \\ \text{Poi}\left(K_b\right), & \text{if } j \neq x + p_i \bmod M. \end{cases}$$

It can be shown (similarly to ML detection of repeat-PPM) that the ML estimate of x is given by

$$\arg\max_{y} \sum_{i=0}^{N-1} s_{i,(y+p_i \bmod M)}$$

which is precisely the collapsed super-symbol slot with the largest photon count, where the noisy PPM symbols are *unspread* by mapping the jth slot count of the ith symbol to the $(j - p_i \bmod M)$th slot of the collapsed super-symbol. An example of spreading and unspreading a PN symbol is given in Figure 1.

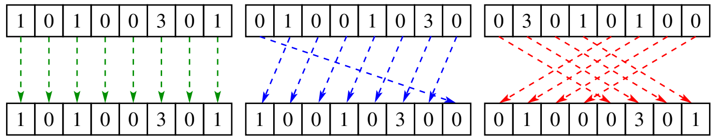

Figure 1. The 8-ary PPM symbol 5 is spread by the 8-ary PN sequence $0, 1, 4$, and the spread sequence $5, 6, 1$ is transmitted through an optical channel which produces the given noisy PPM symbols. The receiver unspreads the noisy symbols as shown and can form the collapsed super-symbol $[2, 1, 1, 1, 0, 9, 0, 2]$. By selecting the collapsed super-symbol slot with the largest photon count, the noisy sequence correctly decodes to 5.

When the spreading is undone correctly, the resulting unspread super-symbol is the same as if spreading were not used. However, when the receiver misaligns the super-symbols, the spreading is not undone correctly, and the signal slots of the resulting "unspread" super-symbol are still spread. This, in theory, makes it easier for the receiver to correctly align super-symbols when spreading is used. This is further discussed in Sections IV-A and V for the particular estimation technique used.

III. Super-Symbol Synchronization Model

An arbitrary M-ary PPM symbol sequence $\boldsymbol{x} = x_0, x_1, \ldots, x_B$ is spread by the M-ary PN sequence $p_0, p_1, \ldots, p_{N-1}$, and the resulting $N(B+1)$ spread symbols are transmitted through an optical channel. The PN sequence is known by both the sender and receiver, and further discussion of selecting a PN sequence is given in Section V-B. Throughout the remainder of this article, for all $i \in \mathbb{Z}$, p_i is assumed to be $p_{(i \bmod N)}$, for simplicity of notation. E.g., $p_{-1} = p_{N-1}$.

The jth repeat of a symbol is permuted by adding p_j to the symbol mod M, so the input to the channel is the sequence

$$\begin{aligned}
&(x_0 + p_0 \bmod M),\ (x_0 + p_1 \bmod M),\ \ldots,\ (x_0 + p_{N-1} \bmod M), \\
&(x_1 + p_0 \bmod M),\ (x_1 + p_1 \bmod M),\ \ldots,\ (x_1 + p_{N-1} \bmod M), \\
&\qquad\vdots \\
&(x_B + p_0 \bmod M),\ (x_B + p_1 \bmod M),\ \ldots,\ (x_B + p_{N-1} \bmod M).
\end{aligned}$$

We assume the detector outputs NB noisy PPM symbols $\boldsymbol{s} = (\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_{NB-1})$, where each noisy symbol is a sequence of photon slot counts $\boldsymbol{s}_i = (s_{i,0}, s_{i,1}, \ldots, s_{i,M-1})$. The sender transmits N more symbols than the receiver detects; this corresponds to the receiver potentially beginning synchronization in the middle of an arbitrary sequence of symbols. Throughout the remainder of this article, for all $j \in \mathbb{Z}$, $s_{i,j}$ is assumed to be $s_{i,(j \bmod M)}$, for simplicity of notation.

For each $n \in \{0, 1, \ldots, N-1\}$, $i \in \{0, 1, \ldots, NB-1\}$, and $j \in \{0, 1, \ldots, M-1\}$, when the offset is equal to n and $\boldsymbol{x}$ is the symbol sequence, the distribution of $s_{i,j}$ is given by

$$\begin{cases} \text{Poi}\left(K_s + K_b\right), & \text{if } j = x_{\lfloor \frac{i+n}{N} \rfloor} + p_{(i+n)} \bmod M \\ \text{Poi}\left(K_b\right), & \text{if } j \neq x_{\lfloor \frac{i+n}{N} \rfloor} + p_{(i+n)} \bmod M. \end{cases}$$

I.e., the noisy symbols correspond to the spread symbols as follows:

$$\underbrace{\boldsymbol{s}_0}_{x_0+p_n}, \ldots, \underbrace{\boldsymbol{s}_{N-n-1}}_{x_0+p_{N-1}}, \underbrace{\boldsymbol{s}_{N-n}}_{x_1+p_0}, \ldots, \underbrace{\boldsymbol{s}_{2N-n-1}}_{x_1+p_{N-1}}, \ldots, \underbrace{\boldsymbol{s}_{BN-N-n}}_{x_{B-1}+p_0}, \ldots, \underbrace{\boldsymbol{s}_{BN-n-1}}_{x_{B-1}+p_{N-1}}, \underbrace{\boldsymbol{s}_{BN-n}}_{x_B+p_0}, \ldots, \underbrace{\boldsymbol{s}_{BN-1}}_{x_B+p_{n-1}}.$$

The first $N-n$ noisy PPM symbols correspond to the last $(N-n)$ transmissions of x_0, the next N noisy PPM symbols correspond to the N transmissions of x_1, the next N noisy PPM symbols correspond to the N transmissions of x_2, and so on, up to the last n noisy PPM symbols, which correspond to the first n transmissions of x_B.

For each $i \in \{0, 1, \ldots, B\}$, $j \in \{0, 1, \ldots, M-1\}$, and $n \in \{0, 1, \ldots, N-1\}$, let $a_{i,j}^{(n)}$ denote the jth slot count of the ith collapsed super-symbol formed from $\boldsymbol{s}$ when the receiver assumes the offset is n:

$$a_{i,j}^{(n)} = \begin{cases} \displaystyle\sum_{l=0}^{N-n-1} s_{l,j+p_{(l+n)}} & i = 0 \\ \displaystyle\sum_{l=Ni-n}^{Ni-n-1} s_{l,j+p_{(l+n)}} & i \in \{1, \ldots, B-1\} \\ \displaystyle\sum_{l=BN-n}^{BN-1} s_{l,j+p_{(l+n)}} & i = B. \end{cases}$$

I.e., the super-symbols are formed as follows:

$$\underbrace{\boldsymbol{s}_0, \ldots, \boldsymbol{s}_{N-n-1}}_{\boldsymbol{a}_0^{(n)}}, \underbrace{\boldsymbol{s}_{N-n}, \ldots, \boldsymbol{s}_{2N-n-1}}_{\boldsymbol{a}_1^{(n)}}, \ldots, \underbrace{\boldsymbol{s}_{BN-N-n}, \ldots, \boldsymbol{s}_{BN-n-1}}_{\boldsymbol{a}_{B-1}^{(n)}}, \underbrace{\boldsymbol{s}_{BN-n}, \ldots, \boldsymbol{s}_{BN-1}}_{\boldsymbol{a}_B^{(n)}}.$$

When the offset is equal to n, the collapsed super-symbols $\boldsymbol{a}_0^{(n)}, \boldsymbol{a}_1^{(n)}, \ldots, \boldsymbol{a}_B^{(n)}$ are correctly aligned with the symbols $x_0, x_1, \ldots, x_B$. The pseudo-noise terms are correctly accounted for, and $\boldsymbol{a}_i^{(n)}$ consists precisely of the symbols corresponding to x_i. However, when the offset is not equal to n, the PN terms are not correctly accounted for, and $\boldsymbol{a}_i^{(n)}$ can consist of symbols corresponding to x_i and x_{i-1}.

A. Example: Repeating vs. Spreading

In the following example, the 4-ary PPM symbols $0, 2, 1$ are each sequentially repeated $N = 3$ times with and without PN spreading, where the PN sequence is $0, 1, 3$. The channel produces noisy symbols, as shown in Figure 2.

Both the PPM symbol sequence and the offset are unknown to the receiver, and the receiver wishes to determine the offset (i.e., obtain super-symbol synchronization). In order to do so, the receiver will form super-symbols, assuming the offset is equal to $0, 1,$ and 2, as shown in Figures 3, 4, and 5, respectively. The receiver then forms an estimate by choosing the offset that produces the "best" collapsed super-symbols. In particular, the maximum slot photon count of each collapsed super-symbol will be summed, and the offset with the largest sum will be selected.

When the receiver undoes the PN spreading and the offset is not correct, the signal slots are still spread, since the incorrect pseudo-noise is subtracted off each term. In the repeat case, the receiver produces the exact same super-symbols when the offset is 1 and 2. However, in the spread case, the receiver correctly chooses the offset to be 2. Using this estimation technique, PN spreading offers an advantage over repeating in this case.

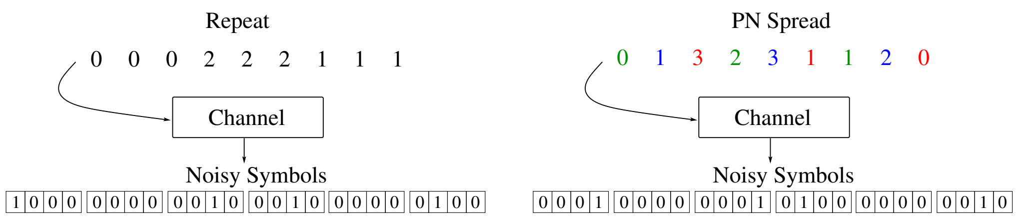

Figure 2. The 4-ary PPM symbols $0, 2, 1$ are each sequentially repeated $N = 3$ times with and without PN spreading, where the PN sequence is $0, 1, 3$. When the repeated/spread sequences are transmitted, the channel produces 6 noisy symbols which are offset by 2, i.e., the noisy symbols correspond to the symbol sequences $0, 2, 2, 2, 1, 1$ and $3, 2, 3, 1, 1, 2$ for the repeat and spread case, respectively. In both cases, no photons are detected in the 1st and 4th noisy symbols.

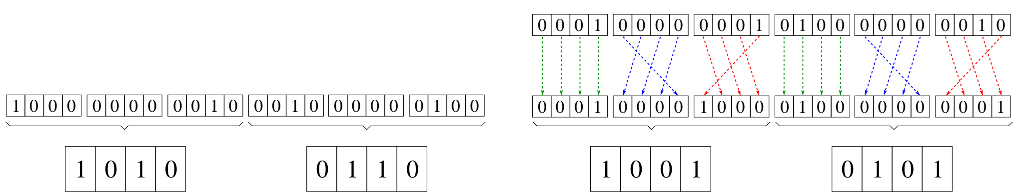

Figure 3. When the assumed offset is 0, the receiver assumes the first noisy symbol is the beginning of a super-symbol. In both the repeat case and the spread case, the super-symbol sum is $1 + 1 = 2$.

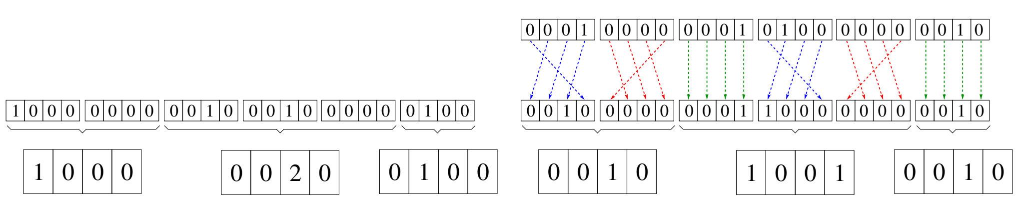

Figure 4. When the assumed offset is 1, the repeat case produces super-symbols whose sum is $1 + 2 + 1 = 4$, and the spread sum is $1 + 1 + 1 = 3$.

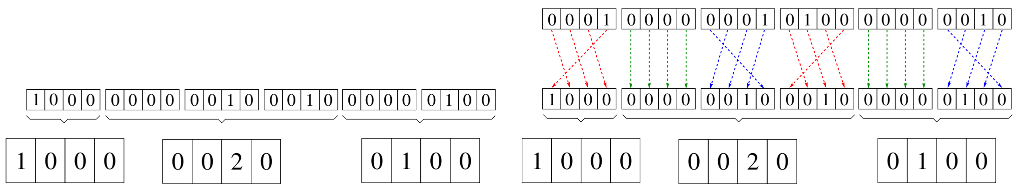

Figure 5. When the assumed offset is 2, the repeat case and the spread case produce the exact same collapsed super-symbols, and the sum is $1+2+1=4$.

IV. ML Estimate of the Offset

Given a sequence of spread PPM symbols, the receiver wishes to determine the value of the offset. Let $\mathcal{X} = \{0, 1, \ldots, M-1\}^{B+1}$ denote the set of possible symbol sequences. We assume each sequence $\boldsymbol{x} \in \mathcal{X}$ is equally likely to be transmitted, and we use the system model described in Section III.

The ML estimate of the offset is given by:

$$\begin{aligned}
&\arg\max_n \; P(\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_{NB-1} \mid \text{offset } = n) \\
&= \arg\max_n \; \frac{1}{M^{B+1}} \sum_{\boldsymbol{x} \in \mathcal{X}} P(\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_{NB-1} \mid \boldsymbol{x}, \text{offset } = n) \\
&= \arg\max_n \; \sum_{\boldsymbol{x} \in \mathcal{X}} \prod_{i=0}^{NB-1} \prod_{j=0}^{M-1} P(s_{i,j} \mid \boldsymbol{x}, \text{Offset } = n) \\
&= \arg\max_n \; \sum_{\boldsymbol{x} \in \mathcal{X}} \prod_{i=0}^{NB-1} \left(e^{-K_s} \prod_{j=0}^{M-1} \frac{e^{-K_b}}{s_{i,j}!} K_b^{s_{i,j}} \right) (1 + K_s/K_b)^{s_{i,\left(x_{\left\lfloor \frac{i+n}{N} \right\rfloor} + p_{(i+n)}\right)}} \\
&= \arg\max_n \; \sum_{\boldsymbol{x} \in \mathcal{X}} \prod_{i=0}^{NB-1} (1 + K_s/K_b)^{s_{i,\left(x_{\left\lfloor \frac{i+n}{N} \right\rfloor} + p_{(i+n)}\right)}} \\
&= \arg\max_n \; \sum_{\boldsymbol{x} \in \mathcal{X}} (1 + K_s/K_b)^{\sum_{i=0}^{NB-1} s_{i,\left(x_{\left\lfloor \frac{i+n}{N} \right\rfloor} + p_{(i+n)}\right)}}.
\end{aligned}$$

For each $\boldsymbol{x} \in \mathcal{X}$, we have

$$\begin{aligned}
&\sum_{i=0}^{NB-1} s_{i,\left(x_{\left\lfloor \frac{i+n}{N} \right\rfloor} + p_{(i+n)}\right)} \\
&= \left(\sum_{j=0}^{N-n-1} s_{j,(x_0 + p_{(j+n)})} \right) + \left(\sum_{i=1}^{B-1} \sum_{j=iN-n}^{(i+1)N-n-1} s_{j,(x_i + p_{(j+n)})} \right) + \left(\sum_{j=BN-n}^{BN-1} s_{j,(x_B + p_{(j+n)})} \right) \\
&= \sum_{i=0}^{B} a_{i,x_i}^{(n)}
\end{aligned}$$

which is the sum of the collapsed super-symbol photon slot counts corresponding to the symbol sequence $\boldsymbol{x}$ when the offset is assumed to be n. Hence the ML estimate of the offset is given by

$$\arg\max_{n} \sum_{\boldsymbol{x}\in\mathcal{X}} (1+K_s/K_b)^{\sum_{i=0}^{B} a_{i,x_i}^{(n)}}.$$

In the proposed CCSDS standard for HPE optical communication, an M-ary PPM encoder produces codeword blocks of $B = 15120/\log_2 M$ PPM symbols. For each block of B symbols, there are $M^B = M^{15120/\log_2 M} = 2^{15120}$ possible sequence of symbols. In these cases, summing over every possible sequence of symbols is not feasible. However, we can approximate this by only considering the symbol sequence that yields the maximum term. In particular, we use the approximation

$$\sum_{\boldsymbol{x}\in\mathcal{X}} (1+K_s/K_b)^{\sum_{i=0}^{B} a_{i,x_i}^{(n)}} \approx \max_{\boldsymbol{x}\in\mathcal{X}} (1+K_s/K_b)^{\sum_{i=0}^{B} a_{i,x_i}^{(n)}}. \tag{1}$$

We have

$$\max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i=0}^{B} a_{i,x_i}^{(n)} = \sum_{i=0}^{B} \max_{j} a_{i,j}^{(n)}$$

and since $(1+K_s/K_b) > 1$, this approximated ML estimate in (1) can be written as

$$\widehat{\text{offset}} = \arg\max_{n} \sum_{i=0}^{B} \max_{j} a_{i,j}^{(n)}.$$

We refer to this as the *super-symbol-sum estimate*. This estimator forms super-symbols assuming the offset is n, then sums the maximum photon slot count in each collapsed super-symbol. The estimator then selects the offset that produces the largest such sum. We note that this is precisely the estimator used in the example in Section III-A. Additionally, the receiver requires no knowledge of the average signal or background photon count to form this estimate.

Our results in Section VI suggest that this super-symbol synchronization method outperforms other aspects of the communication system in which it will potentially be used. However, some other synchronization technique may be able to achieve similar performance with lower runtime or with a smaller block size B. Since the ML estimate is not feasible to compute in many cases, considering other approximations or estimation techniques may be of interest for future work.

A. Pseudo-Noise and Correlation

The study of pseudo-noise sequences has led to the development of sequences with good autocorrelation properties [11]. That is, the sequence is highly correlated with itself but not with shifted versions of itself. If we use a good PN sequence to spread the repeated PPM messages, then the noisy PPM symbols will be uncorrelated with one another. Additionally, when a misaligned PN sequence is used to unspread a super-symbol, the resulting symbols will still be uncorrelated with one another.

When n is not equal to the offset, the collapsed super-symbol $\boldsymbol{a}_i^{(n)}$ is a sum of relatively uncorrelated noisy PPM symbols, since the spreading is not undone correctly. When n is equal to the offset, $\boldsymbol{a}_i^{(n)}$ is a sum of highly correlated noisy PPM symbols, since the spreading is undone correctly. Generally speaking, the maximum slot count of a collapsed super-symbol which is a sum of uncorrelated symbols should be smaller than the maximum slot count of a sum of highly correlated symbols. This is good from a super-symbol synchronization perspective, since ideally the term $\sum_{i=0}^{B} \max_j \; a_{i,j}^{(n)}$ is large for n equal to the offset but and is small for n not equal to the offset.

On the other hand, if spreading is not used, then when n is not equal to the offset, $\boldsymbol{a}_i^{(n)}$ is a sum of noisy PPM symbols that are still somewhat correlated with one another, which, in general, will cause $\sum_{i=0}^{B} \max_j \; a_{i,j}^{(n)}$ to be larger than when spreading is used.

V. Justifying Results

In this section, we attempt to provide some analytic justification for the simulation results in Section VI.

If we assume the offset is equally likely to take on any value in $\{0, 1, \ldots, N-1\}$, then by a symmetry argument, the probability the super-symbol-sum estimate is incorrect is given by

$$\begin{aligned} P_e^{PN} &= P(\widehat{\text{offset}} \neq 0 \mid \text{ offset } = 0) \\ &= P\left(\bigcup_{n=1}^{N-1} \left\{ \sum_{i=0}^{B} \max_j \; a_{i,j}^{(n)} > \sum_{i=0}^{B} \max_j \; a_{i,j}^{(0)} \right\} \middle| \text{ offset } = 0 \right). \end{aligned}$$

Generally speaking, errors occur when $\max\limits_j \; a_{i,j}^{(n)}$ is large when n is not equal to the offset and $\max\limits_j \; a_{i,j}^{(n)}$ is small when n is equal to the offset. Ideally, PN spreading helps to prevent the former, since as discussed in Section IV-A, incorrectly unspread symbols are less correlated with one another.

When the offset is 0, the received noisy symbols correspond to the spread symbol sequence in the following way:

$$\underbrace{\boldsymbol{s}_0, \ldots, \boldsymbol{s}_{N-1}}_{x_0}, \underbrace{\boldsymbol{s}_N, \ldots, \boldsymbol{s}_{2N-1}}_{x_1}, \ldots, \underbrace{\boldsymbol{s}_{BN-N}, \ldots, \boldsymbol{s}_{BN-1}}_{x_{B-1}}.$$

For each $n \in \{0, 1, \ldots, N-1\}$, the super-symbols align with the noisy symbols accordingly:

$$\underbrace{\boldsymbol{s}_0, \ldots, \boldsymbol{s}_{N-n-1}}_{\boldsymbol{a}_0^{(n)}}, \underbrace{\boldsymbol{s}_{N-n}, \ldots, \boldsymbol{s}_{2N-n-1}}_{\boldsymbol{a}_1^{(n)}}, \ldots, \underbrace{\boldsymbol{s}_{BN-N-n}, \ldots, \boldsymbol{s}_{BN-n-1}}_{\boldsymbol{a}_{B-1}^{(n)}}, \underbrace{\boldsymbol{s}_{BN-n}, \ldots, \boldsymbol{s}_{BN-1}}_{\boldsymbol{a}_B^{(n)}}.$$

When $B \gg 0$, the contributions from $\boldsymbol{a}_0^{(n)}$ and $\boldsymbol{a}_B^{(n)}$ to the estimator sum become negligible, so we focus on the contributions from $\boldsymbol{a}_1^{(n)}, \ldots, \boldsymbol{a}_{B-1}^{(n)}$.

For each $i \in \{1, 2, \ldots, B-1\}$ and $n \in \{0, 1, \ldots, N-1\}$, the super-symbol $\boldsymbol{a}_i^{(n)}$ consists of the last n noisy symbols corresponding to x_{i-1} followed by the first $(N-n)$ noisy symbols corresponding to x_i. However, the super-symbols do not correctly account for the pseudo-noise. Hence, for each $j \in \{0, 1, \ldots, M-1\}$, given $\boldsymbol{x}$ is the symbol sequence, the distribution of $a_{i,j}^{(n)}$ is given by

$$\sum_{l=0}^{n-1} \begin{cases} \mathrm{Poi}\left(K_s + K_b\right), & \text{if } j + p_l = x_{i-1} + p_{(l-n)} \mod M \\ \mathrm{Poi}\left(K_b\right), & \text{if } j + p_l \neq x_{i-1} + p_{(l-n)} \mod M \end{cases}$$
$$+ \sum_{l=n}^{N-1} \begin{cases} \mathrm{Poi}\left(K_s + K_b\right), & \text{if } j + p_l = x_i + p_{(l-n)} \mod M \\ \mathrm{Poi}\left(K_b\right), & \text{if } j + p_l \neq x_i + p_{(l-n)} \mod M \end{cases}$$

where $\mathrm{Poi}\left(\lambda\right) + \mathrm{Poi}\left(\sigma\right)$ denotes a sum of independent Poisson random variables with mean λ and σ, respectively. The particular distribution of $a_{i,j}^{(n)}$ depends on the particular pseudo-noise sequence chosen. Each collapsed slot count $a_{i,j}^{(n)}$ is a sum of some signal slot counts and some noise slot counts.

For a "good" PN sequence, the signal slots corresponding to x_i and x_{i-1} are evenly distributed over all of the collapsed slots when $n \neq 0$. We discuss this further in Section V-B.

We also note that $a_{i,j}^{(0)}$ is distributed according to

$$\begin{cases} \mathrm{Poi}\left(N(K_s + K_b)\right), & \text{if } j = x_i \\ \mathrm{Poi}\left(NK_b\right), & \text{if } j \neq x_i \end{cases}$$

which is independent of the PN sequence. Because of this, we expect

$$\sum_{i=0}^{B-1} \max_j \, a_{i,j}^{(0)}$$

to have the same distribution regardless of the PN sequence. This agrees with the earlier example, where the repeated and spread sequence yielded the same super-symbols when the offset was correct. Additionally, since the spreading is correctly undone in this case, the PN sequence itself should not matter.

A. Special Case: N=2

Now suppose the repeat factor is $N = 2$. Then $a_{i,j}^{(1)}$ is distributed according to

$$\begin{cases} \mathrm{Poi}\left(2(K_s + K_b)\right), & \text{if } j = x_i + p_0 - p_1 = x_{i-1} + p_1 - p_0 \\ \mathrm{Poi}\left(K_s + 2K_b\right), & \begin{array}{l}\text{if } j = x_i + p_0 - p_1 \neq x_{i-1} + p_1 - p_0, \\ \text{or } j = x_{i-1} + p_1 - p_0 \neq x_i + p_0 - p_1\end{array} \\ \mathrm{Poi}\left(2K_b\right), & \text{if } x_{i-1} + p_1 - p_0 \neq j \neq x_i + p_0 - p_1. \end{cases}$$

For all integers $n \geq 0$, let $F(\lambda, n) = \sum_{k=0}^{n} \frac{e^{-\lambda}\lambda^k}{k!}$, i.e., $F(\lambda, n)$ is the cumulative distribution function of a Poisson random variable with mean λ. Then

$$\begin{aligned}
& P\left(\max_j a_{i,j}^{(1)} \leq n\right) \\
& = P\left(a_{i,0}^{(1)} \leq n,\, a_{i,1}^{(1)} \leq n,\, \ldots,\, a_{i,M-1}^{(1)} \leq n\right) \\
& = \frac{1}{M} P\left(a_{i,0}^{(1)} \leq n,\, a_{i,1}^{(1)} \leq n,\, \ldots,\, a_{i,M-1}^{(1)} \leq n \mid x_i + p_0 - p_1 = x_{i-1} + p_1 - p_0\right) \\
& \quad + \frac{M-1}{M} P\left(a_{i,0}^{(1)} \leq n,\, a_{i,1}^{(1)} \leq n,\, \ldots,\, a_{i,M-1}^{(1)} \leq n \mid x_i + p_0 - p_1 \neq x_{i-1} + p_1 - p_0\right) \\
& = \frac{1}{M} F(2(K_s + K_b), n)\, F(2K_b, n)^{M-1} \\
& \quad + \frac{M-1}{M} F(K_s + 2K_b, n)^2\, F(2K_b, n)^{M-2}
\end{aligned}$$

which is independent of the PN sequence p_0, p_1. Hence the distributions of both $\max_j a_{i,j}^{(1)}$ and $\max_j a_{i,j}^{(0)}$ are independent of the PN sequence, so we expect both

$$\sum_{i=0}^{B-1} \max_j a_{i,j}^{(0)} \quad \text{and} \quad \sum_{i=0}^{B-1} \max_j a_{i,j}^{(1)}$$

to also be independent of the PN sequence. Therefore when the repeat factor is 2, we do not expect PN spreading to offer improvements over repeating.

B. Good PN Sequences

When the offset is 0, the distribution of $a_{i,j}^{(n)}$ is given by

$$\begin{aligned}
& \sum_{l=n}^{N-1} \begin{cases} \mathrm{Poi}\,(K_s + K_b) & j = x_i + p_{(l-n)} - p_l \mod M \\ \mathrm{Poi}\,(K_b) & j \neq x_i + p_{(l-n)} - p_l \mod M \end{cases} \\
& \quad + \sum_{l=0}^{n-1} \begin{cases} \mathrm{Poi}\,(K_s + K_b) & j = x_{i-1} + p_{(l-n)} - p_l \mod M \\ \mathrm{Poi}\,(K_b) & j \neq x_{i-1} + p_{(l-n)} - p_l \mod M, \end{cases}
\end{aligned}$$

and the distribution of $a_{i,j}^{(0)}$ is independent of the PN sequence. When $n > 0$, we want to select a PN sequence which distributes the signal slots of x_i and x_{i-1} evenly over $\{0, 1, \ldots, M-1\}$, since this will, on average, result in smaller values of $\max_j a_{i,j}^{(n)}$. When signal slots are less spread, $\max_j a_{i,j}^{(n)}$ will tend to be larger.

If

$$(p_{(l-n)} - p_l) = (p_{(m-n)} - p_m) \mod M$$

for some distinct $l, m \in \{0, 1, \ldots, n-1\}$, then some collapsed slot count consists of a sum of multiple signal slot counts. Hence when $p_0 = p_1 = \cdots = p_{N-1}$, the $(N-n)$ signal slots from x_i are all in the same super-symbol slot and the n signal slots from x_{i-1} are all in the same super-symbol slot. In fact, in this case, the distribution of $a_{i,j}^{(n)}$

is given by

$$\begin{cases} \mathrm{Poi}\left(N(K_s+K_b)\right), & \text{if } j = x_i = x_{i-1} \\ \mathrm{Poi}\left((N-n)K_s+NK_b\right), & \text{if } j = x_i \neq x_{i-1} \\ \mathrm{Poi}\left(nK_s+NK_b\right), & \text{if } j = x_{i-1} \neq x_i \\ \mathrm{Poi}\left(NK_b\right), & \text{if } x_{i-1} \neq j \neq x_i. \end{cases}$$

Ideally, we have

$$(p_{(l-n)} - p_l) \neq (p_{(m-n)} - p_m) \bmod M$$

for all $n \in \{1, 2, \ldots, N-1\}$ and all distinct $l, m \in \{0, 1, \ldots, n-1\}$, since then the all of the signal slots corresponding to x_i are in distinct collapsed slots and all of the signal slots corresponding to x_{i-1} are in distinct collapsed slots, which on average causes $\max_j a_{i,j}^{(n)}$ to be smaller than when signal slots are not spread. When $M < N-1$, it is not possible to spread the signal slots into distinct collapsed slots, but we would like the sequence to evenly distribute the signal slots over the collapsed slots.

Optimizing the PN sequence in this way for general M and N appears to be a non-trivial task. However as an example, when $N = 4$ and $M \geq 4$, we can select

$$(p_0, p_1, p_2, p_3) = (0, 1, 3, 2).$$

Then

$$(p_1 - p_0, p_2 - p_1, p_3 - p_2) = (1, 2, M-1)$$
$$(p_2 - p_0, p_3 - p_1) = (3, 1)$$

so the signal slots are completely spread over the collapsed slots.

However, for the proposed CCSDS scheme, selecting $p_0, p_1, \ldots, p_{N-1}$ to be a "good" PN sequence appears to be sufficient.

VI. Simulation Results

In the CCSDS proposed recommendation, a serially-concatenated convolutionally-coded M-ary PPM encoder produces codewords consisting of $B = 15120/\log_2 M$ PPM symbols. The PPM symbols are either repeated or spread by a factor of N and transmitted through an optical channel. The receiver uses the estimator described in Section IV with either repeating or spreading to estimate the offset, which is assumed to be uniform over $\{0, 1, \ldots, N-1\}$. We use the M-ary PN sequence described in [1].

In Figures 6 to 20, the performance of the repeat and spread estimators versus K_s/M are plotted for PPM orders $M \in \{2, 4, 8, 16, 32, 64, 128, 256\}$, repeat factors $N \in \{2, 4, 8, 16, 32\}$, and $K_b \in \{0, 0.01, 1\}$. When repeating (or spreading) by a factor of N and super-symbol synchronization is attained, the effective M-ary PPM optical channel is a memoryless Poisson process with NK_s average received signal photons and NK_b average received background photons. The soft-decision capacity of such a channel is derived in terms of bits per PPM symbol in [4].

In the proposed scheme, prior to mapping binary symbols to PPM symbols, the binary symbols are convolutionally encoded at a rate of either $R = 1/3, 1/2$, or $2/3$ [1] as shown in Table 1. When more than one rate is allowed for a particular PPM order, we compare the super-symbol synchronization performance to the lower rate code, since lower signal levels are needed for reliable communication at lower rates. For each PPM order, repeat factor, and K_b level, the capacity-bounded bit error rate is plotted alongside the super-symbol synchronization performance plots.

PPM Order, M	Allowed Code Rate(s), R
4	2/3
8	2/3
16	2/3, 1/2
32	1/2
64	1/2
128	1/2, 1/3
256	1/2, 1/3

Table 1. Allowed rates of convolutional encoder for particular PPM orders, given in [1]. In Figures 6 to 20, the bit error rates correspond to the capacity bounds when the code rates are the values in this table.

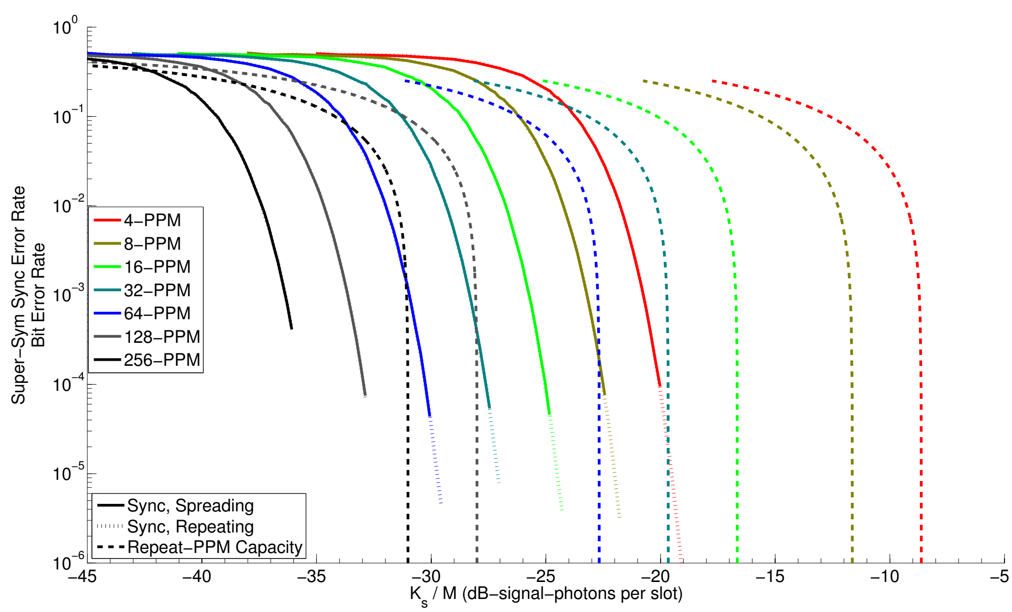

Figure 6. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 2$, $K_b = 0$, and convolutional code rates specified in Table 1.

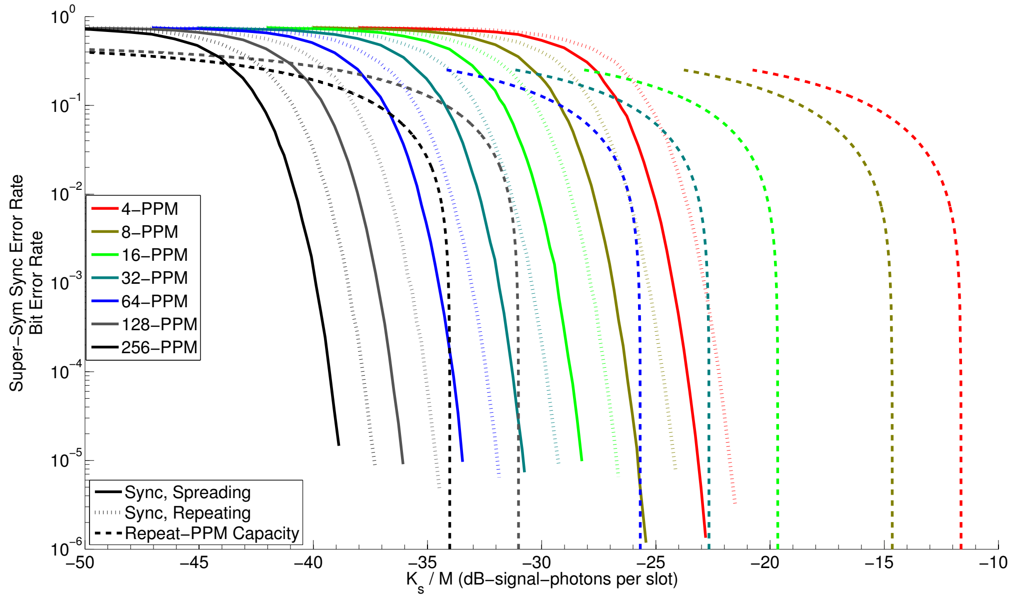

Figure 7. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 4$, $K_b = 0$, and convolutional code rates specified in Table 1.

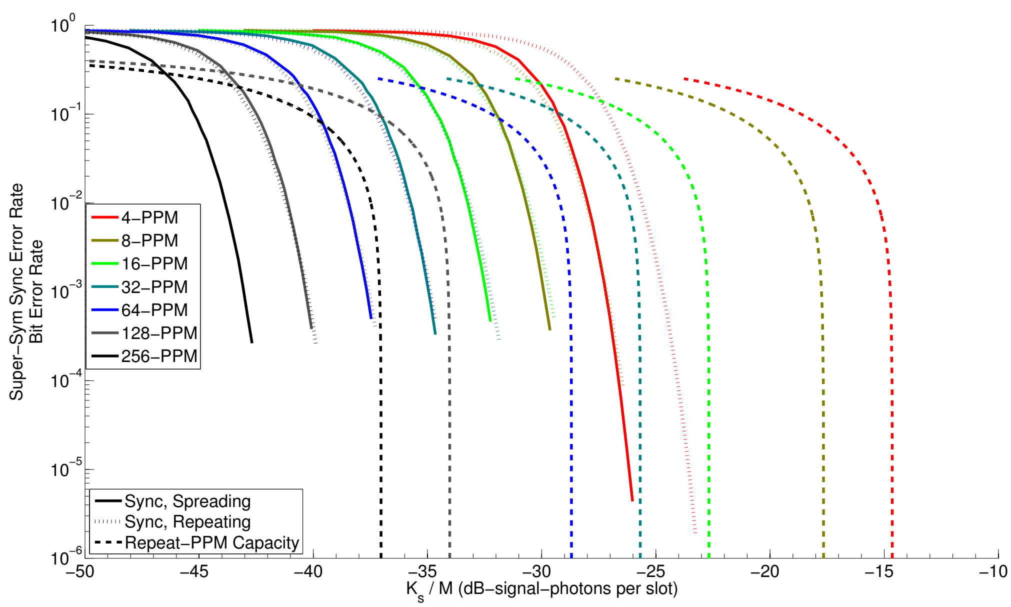

Figure 8. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 8$, $K_b = 0$, and convolutional code rates specified in Table 1.

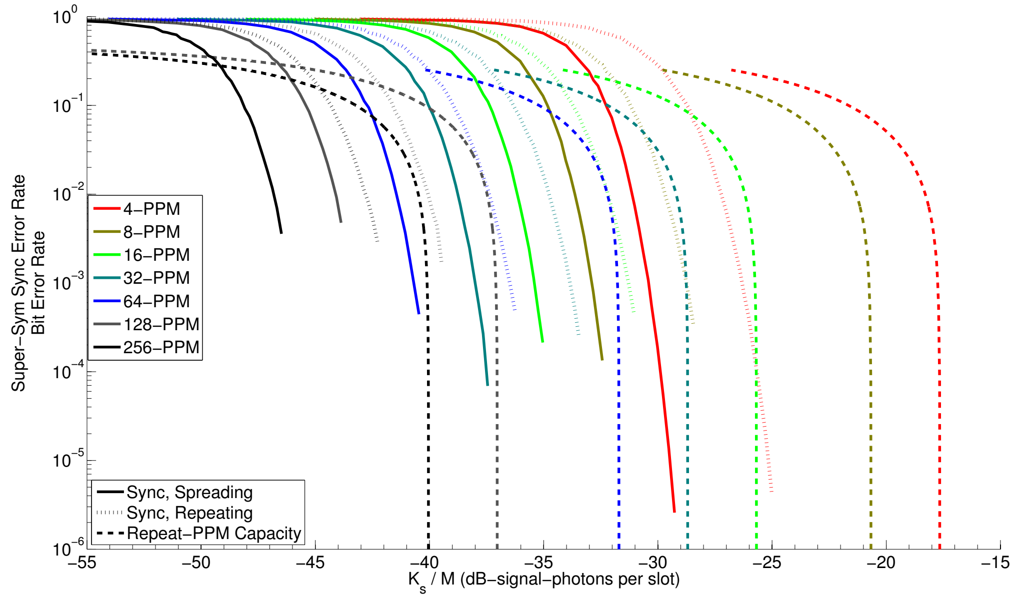

Figure 9. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 16$, $K_b = 0$, and convolutional code rates specified in Table 1.

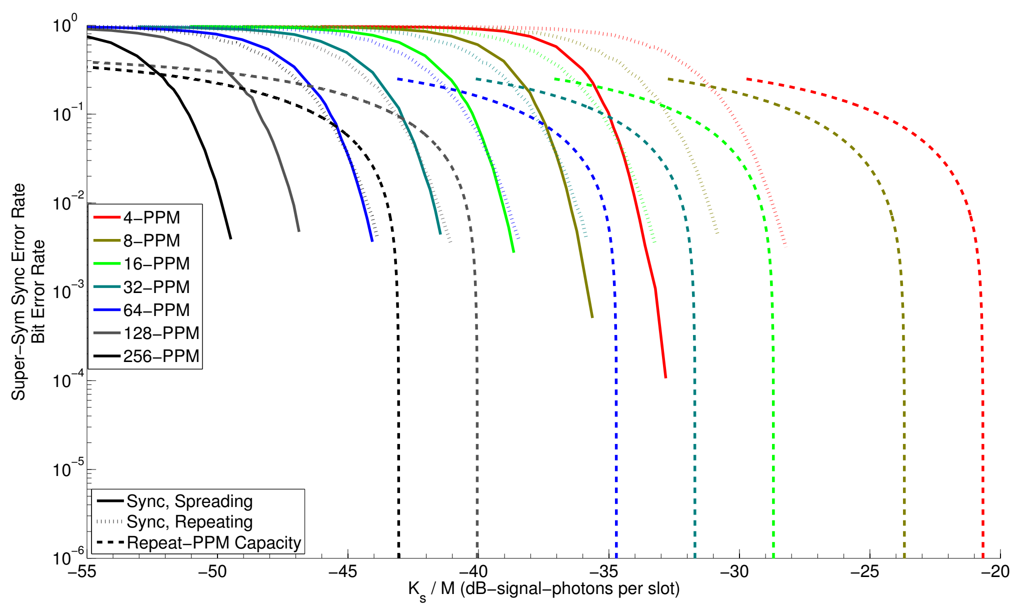

Figure 10. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 32$, $K_b = 0$, and convolutional code rates specified in Table 1.

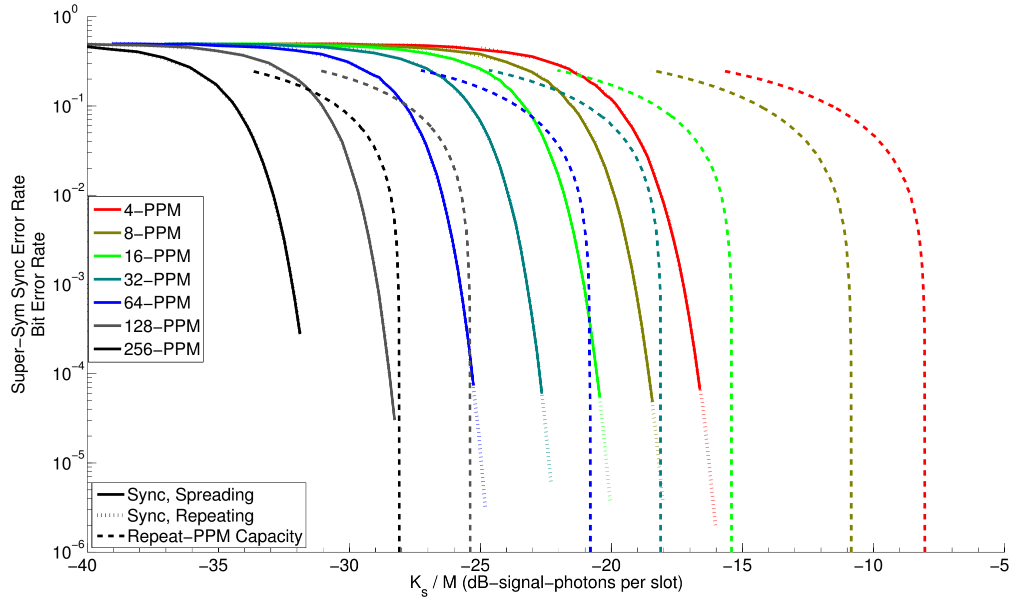

Figure 11. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 2$, $K_b = 0.01$, and convolutional code rates specified in Table 1.

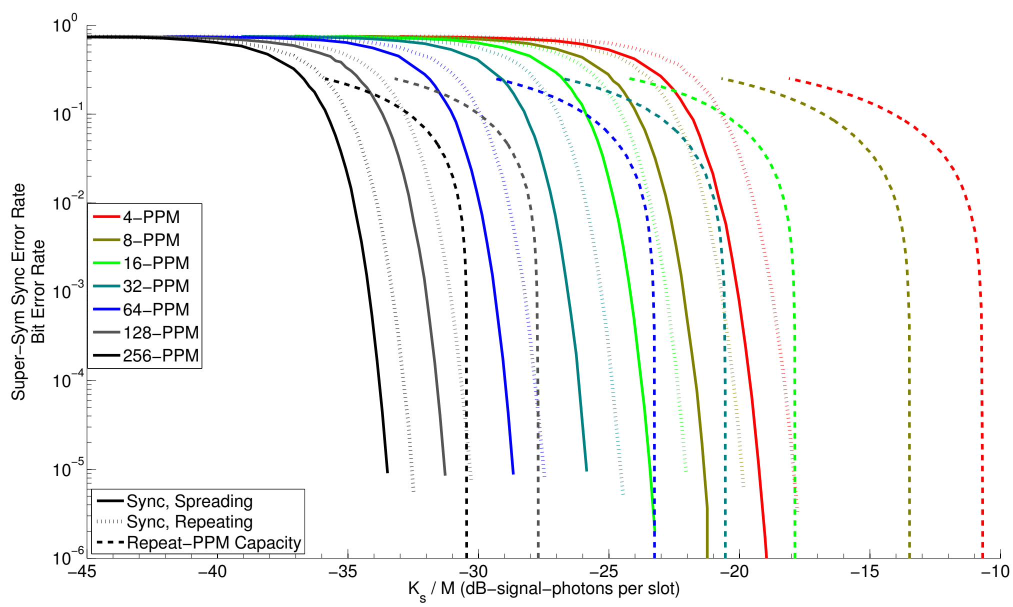

Figure 12. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 4$, $K_b = 0.01$, and convolutional code rates specified in Table 1.

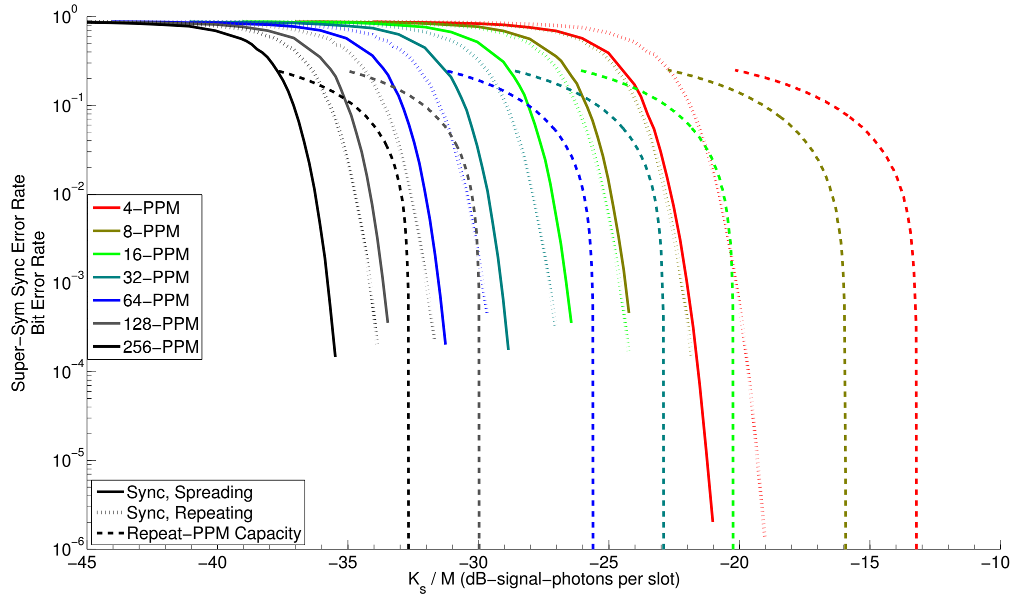

Figure 13. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 8$, $K_b = 0.01$, and convolutional code rates specified in Table 1.

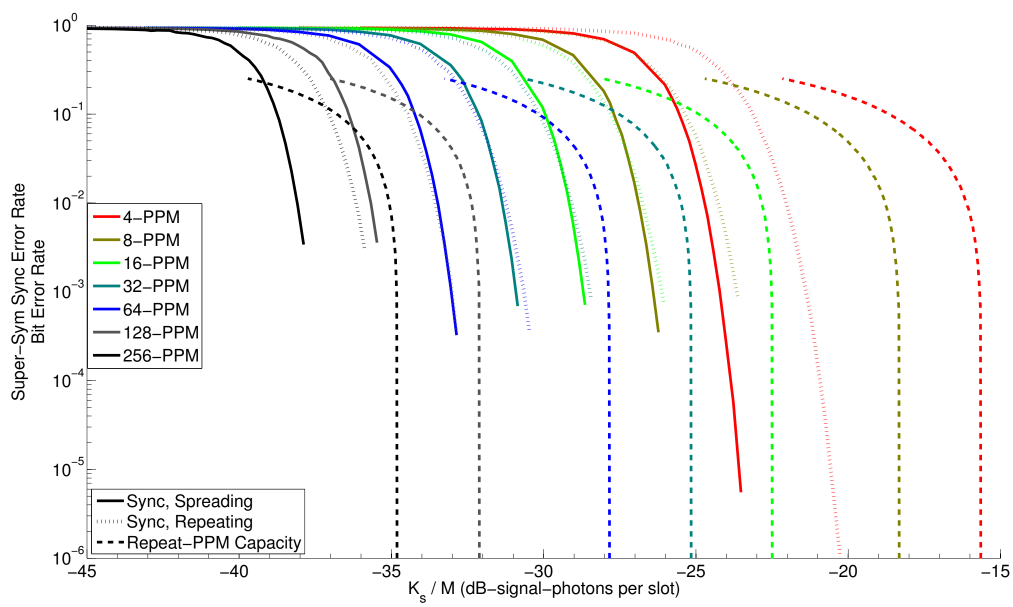

Figure 14. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 16$, $K_b = 0.01$, and convolutional code rates specified in Table 1.

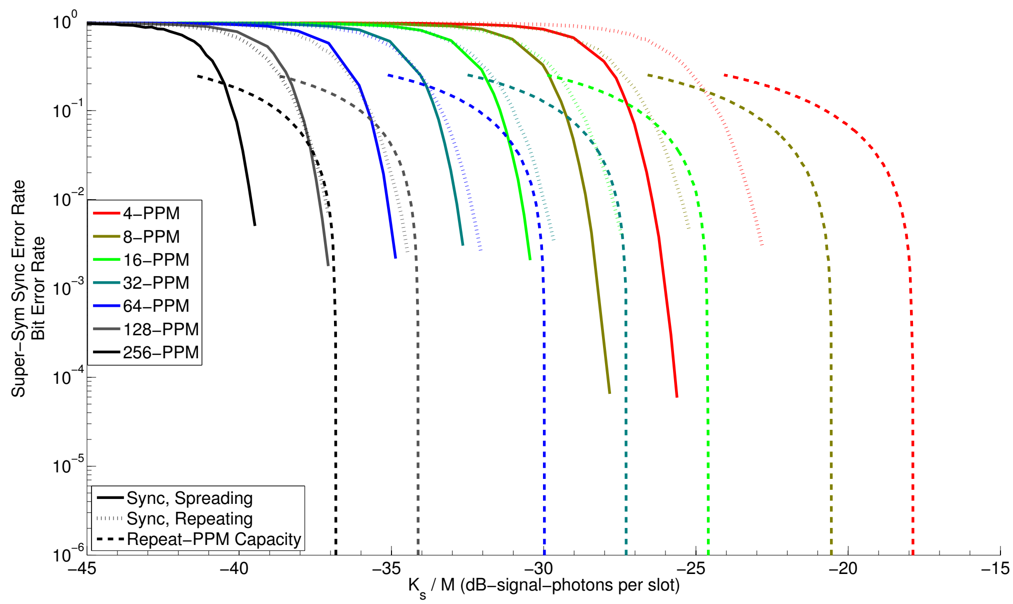

Figure 15. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 32$, $K_b = 0.01$, and convolutional code rates specified in Table 1.

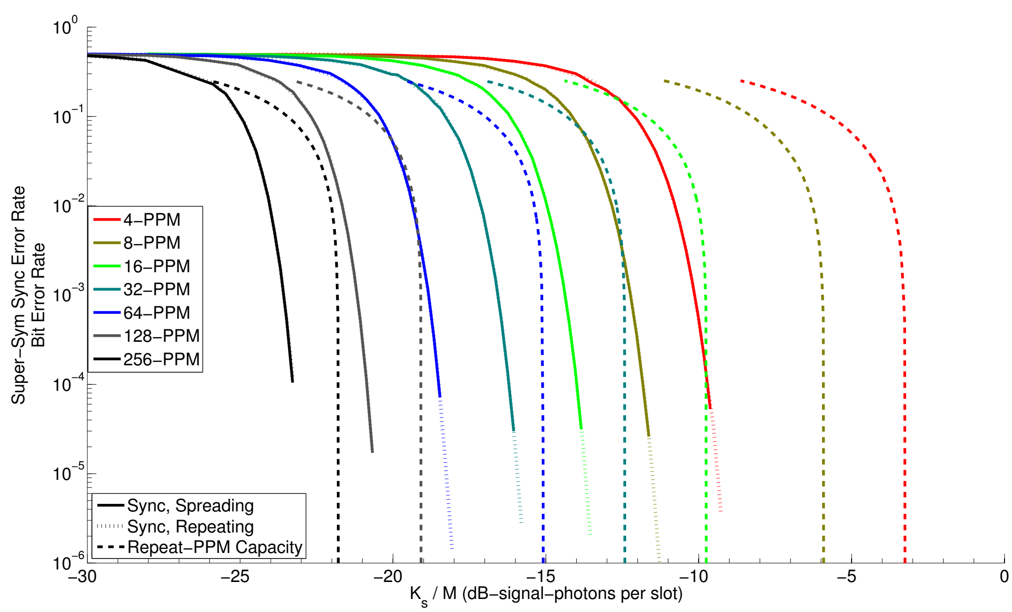

Figure 16. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 2$, $K_b = 1$, and convolutional code rates specified in Table 1.

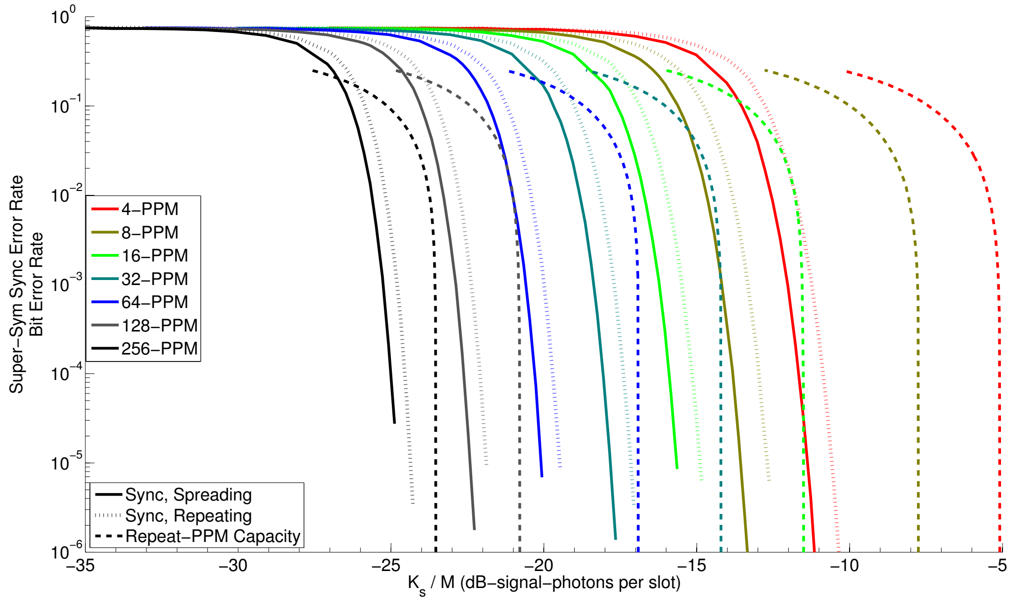

Figure 17. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 4$, $K_b = 1$, and convolutional code rates specified in Table 1.

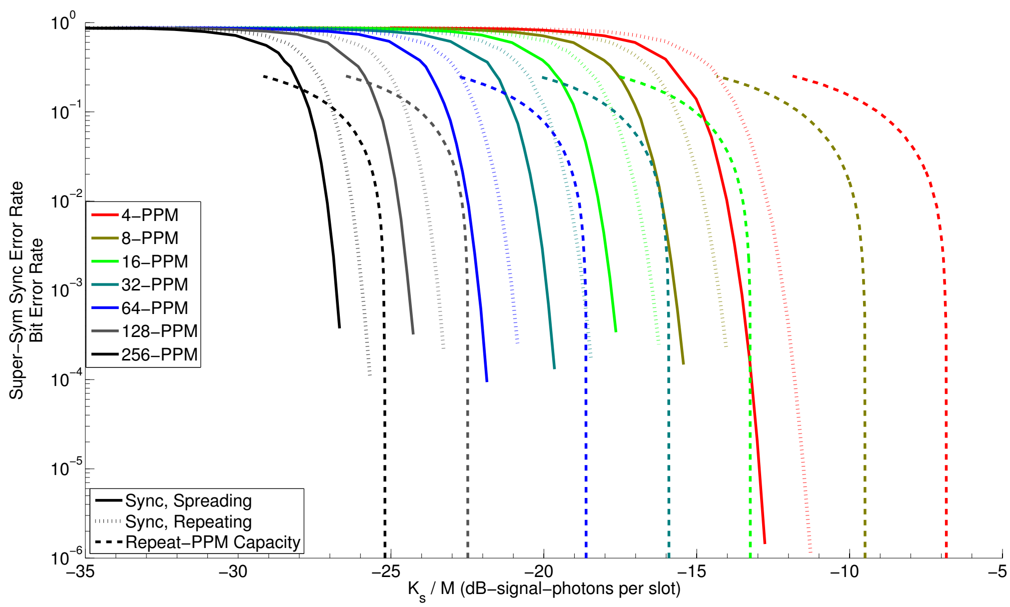

Figure 18. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 8$, $K_b = 1$, and convolutional code rates specified in Table 1.

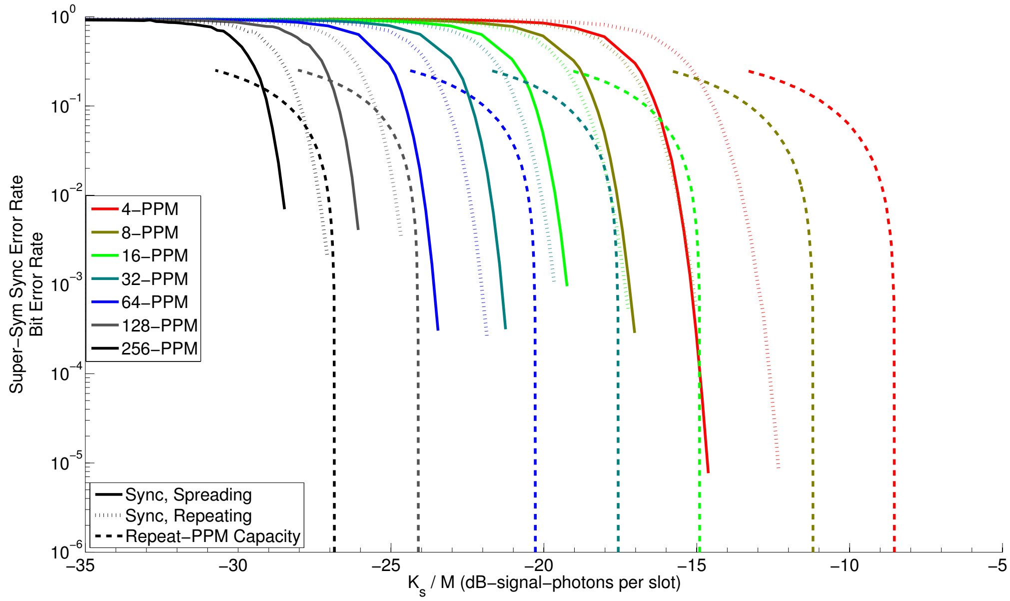

Figure 19. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 16$, $K_b = 1$, and convolutional code rates specified in Table 1.

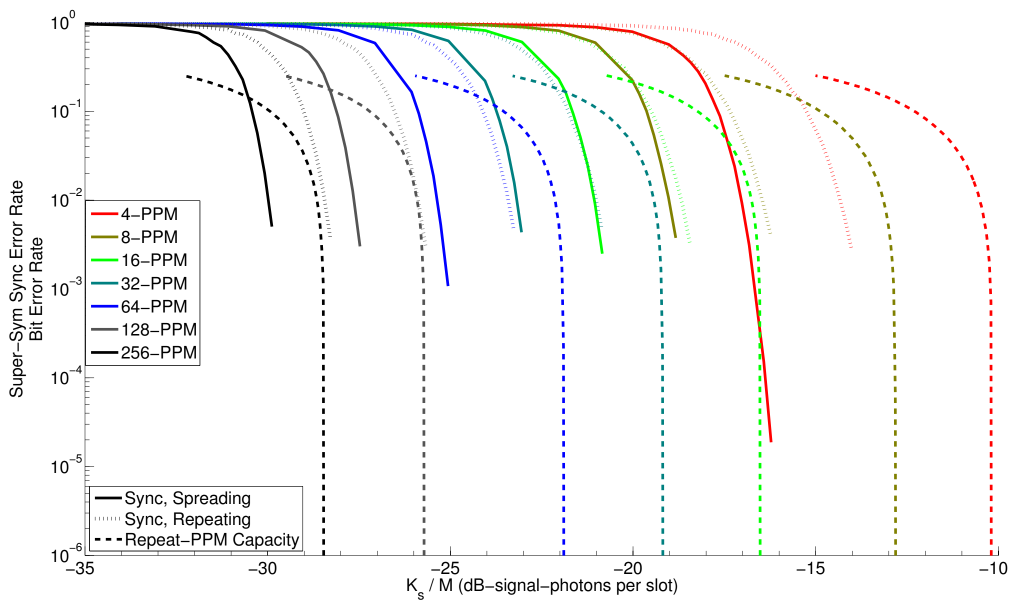

Figure 20. Super-symbol synchronization performance and repeat-PPM capacity-bounded bit error rate, for repeat/spread factor $N = 32$, $K_b = 1$, and convolutional code rates specified in Table 1.

VII. Conclusions

The simulation results demonstrate that when the receiver has obtained slot and symbol synchronization and PPM symbols are spread, the receiver can reliably obtain super-symbol synchronization at the signal levels needed for reliable communication at the code rates specified in the proposed CCSDS standard. The super-symbol synchronization technique presented in this report requires neither a priori knowledge of the signal and noise levels nor explicit synchronization sequences. Additionally, the runtime of the given technique is $O(BMN^2)$, where the M-ary PPM symbols are spread by a factor of N and the receiver uses NB noisy symbols to obtain super-symbol synchronization.

For fixed PPM order and spread factor, super-symbol synchronization performance can be improved by increasing the number of noisy symbols considered (i.e., increasing B). The runtime of our technique increases linearly with B, and the tradeoff between runtime and performance of this (or other super-symbol synchronization) techniques may be of interest for future studies. Particularly, for a fixed runtime, what techniques maximize super-symbol synchronization performance?

Acknowledgment

The author would like to thank Jon Hamkins for his guidance and technical insight throughout the development of this work.

References

[1] CCSDS Proposed Recommended Draft Standard, "High Photon Efficiency Optical Communications, Coding & Synchronization," White Book, September 2016.

[2] CCSDS Proposed Recommended Draft Standard, "High Photon Efficiency Optical Communications, Physical Layer," White Book, September 2016.

[3] A. Wyner, "Capacity and Error Exponent for the Direct Detection Photon Channel Part I," *IEEE Transactions on Information Theory*, vol. 34, pp. 1449–1461, November 1988.

[4] H. Hemmati, *Deep Space Optical Communications*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.

[5] G. Ling and R. Gagliardi, "Slot synchronization in optical PPM communications," *IEEE Transactions on Communications*, vol. 34, pp. 1202–1208, Dec. 1986.

[6] V. Vilnrotter, E. Rodemich, and H. Tan "A synchronization technique for optical PPM signals," *The Telecommunications and Data Acquisition Progress Report*, vol. 42-87, Jet Propulsion Laboratory, Pasadena, California, pp. 24–31, Nov. 15, 1986.

[7] C. Georghiades, "On the synchronizability and detectability of random PPM sequences," *IEEE Transactions on Information Theory*, vol. 35, pp. 146–156, Jan 1989.

[8] K. Quirk, J. Gin, and M. Srinivasan, "Optical PPM synchronization for photon counting receivers," *IEEE Military Communications Conference*, pp. 1–7, November 2008.

[9] Y. Fujiwara, "Self-synchronizing pulse position modulation with error tolerance," *IEEE Transactions on Information Theory*, vol. 59, pp. 5352–6362, Sept. 2013.

[10] V. Vilnrotter, M. Simon, and M. Srinivasan, "Maximum likelihood detection of PPM signals governed by arbitrary point-process plus additive Gaussian noise," *IEEE Electronic Letters*, vol. 35, pp. 1132–1133, July 1999.

[11] C. Georghiades, "On PPM sequences with good autocorrelation properties," *IEEE Transactions on Information Theory*, vol. 34, pp. 571–576, May 1988.