

# On the Limits of Communication over Optical On-Off Keying Channels with Crosstalk

Hongchao Zhou and Gregory Wornell  
 Research Laboratory of Electronics  
 Massachusetts Institute of Technology  
 Cambridge, MA 02139  
 {hongchao, gww}@mit.edu

**Abstract**—In this paper, we investigate the limits of communication over optical on-off-keying channels with 1-D or 2-D crosstalk, where photons are transferable between adjacent time slots or spatial pixels, and the receiver is equipped with single-photon detectors. We observe that high transmission power (measured by the expected number of photons emitted in each signal slot or pixel) may not lead to high information rate; the maximum capacity is typically achieved in a low-photon regime – with about expected 3 to 8 photons received in each signal slot or pixel. Furthermore, we study the selection of slot length for maximizing the channel bandwidth, as the slot length affects the crosstalk probability and hence the channel capacity. It reveals that optimum optical-communication systems do not minimize the level of crosstalk between slots or pixels.

## I. INTRODUCTION

Optical communication, supporting high-speed long-distance point-to-point information delivery, serves as the backbone of the Internet and has important applications in space-based systems such as satellite-to-satellite relays. Today's optical communication explores multiplexing techniques in different dimensions including time, wavelength, polarization, and space, where crosstalk is a key issue that limits the channel capacities. For the time-division multiplexing (TDM), where the time domain is divided into small slots, photons from one slot might be observed in adjacent slots, resulting in temporal crosstalk, e.g., caused by electronics jitters [1]. For the wavelength-division multiplexing (WDM), photons from one channel corresponding to a wavelength can leak into another when the channels are not spaced far apart on wavelength [2]. Two-dimensional spatial crosstalk is often observed in multi-core fibers and multi-spatial-mode communication. In multi-core fibers, photons are transferable between neighboring cores, determined by some factors such as coupling between cores, fiber length, fiber layout (e.g., bends and twists), and the operation wavelength [3]. In multi-spatial-mode communication, crosstalk happens between adjacent closely-packed channels at the photon detector array, caused by, e.g., the presence of turbulence [4].

In this paper, we consider optical communication with both photon losses and crosstalk, which can be one-dimensional or two-dimensional depending on the deployed multiplexing technique. Specifically, we focus on the on-off keying (OOK) modulation, where the optical light is pulsed on or off for

This work was supported in part by AFOSR under Grant No. FA9550-11-1-0183, and by the DARPA InPho program under Contract No. HR0011-10-C-0159.

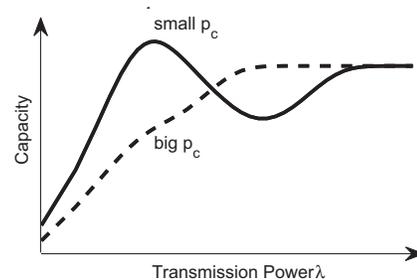


Fig. 1. Capacity behaviors of optical on-off-keying channels.

each slot (or each pixel for the 2-D case) to transmit a bit 1 or 0. Given a signal slot, which transmits 1, the number of photons emitted in the slot is a Poisson random variable  $k$  with mean  $\lambda$ . We call  $\lambda$  the transmission power. The receiver is equipped with single-photon detectors such as avalanche photodiodes and photomultiplier tubes that can detect whether there are incoming photons in a slot or a pixel. In our model, we assume that each photon emitted by the transmitter has a probability  $\eta$  to be detected by the receiver (if there are no other photons), and it has a probability  $p_c$  to appear in adjacent slots (or pixels). We call  $\eta$  the detection probability and  $p_c$  the crosstalk probability.

To maximize the communication bandwidth, it is essential to understand the relation between the channel capacity and the transmission power  $\lambda$ . Here, we refer the channel capacity as the maximal information rate with a specific transmission power  $\lambda$ , rather than the maximal information rate with a transmission power upper bounded by  $\lambda$ . In fact, the optical on-off-keying channels with crosstalk demonstrate completely different behaviors from most well-studied channels such as AWGN channels, where the channel capacity is a monotonically increasing function of the signal-noise-ratio (SNR). In this paper, we derive the lower and upper bounds on the capacities of optical on-off-keying channels with 1-D or 2-D crosstalk. Our results show that the channel capacity is a monotonically increasing function of the transmission power  $\lambda$  only when the crosstalk probability  $p_c$  is sufficiently large. For small values of  $p_c$ , the channel capacity has two peaks with different transmission powers: Numerical results show that the first peak locates in a low-photon regime with  $3 \leq \lambda\eta \leq 8$  specified by (8) for both the 1-D and 2-D cases, and the second

peak appears at  $\lambda \rightarrow \infty$ , see Fig. 1 as a demonstration. When  $p_c$  is smaller than a threshold (e.g., 0.068 for the 1-D case), the first peak is higher than the second one, implying that only a small amount of photons should be transmitted within a light pulse. In addition, we study the maximal information rates with simple input distributions for the purpose of constructing error-correcting codes, and we investigate the effect of slot length on the communication bandwidth.

The rest of this paper is organized as follows. Section II describes the channel model and provides some basic definitions. Section III and Section IV derive bounds on the capacities of optical on-off-keying channels with 1-D or 2-D crosstalk, respectively. Section V investigates how to select slot length (or pixel size) appropriately for maximizing the communication bandwidth. Due to the space limit, we introduce the technical ideas briefly, and mainly focus on useful observations and conclusions for optical-communication system design.

## II. CHANNEL CAPACITY AND INFORMATION RATE

### A. Channel Model

For optical channels with 1-D crosstalk, each slot has two adjacent slots, and a photon occurs in a particular adjacent slot with a probability  $\frac{p_c}{2}$ , where  $p_c$  is the crosstalk probability. Let  $X = x_1 x_2 \dots x_m \in \{0, 1\}^m$  be the input sequence and  $Y = y_1 y_2 \dots y_m \in \{0, 1\}^m$  be the output sequence, where  $x_i$  and  $y_i$  are the bits sent and received through slot  $i$ . The number of photons received by the receiver at slot  $i$  is a Poisson random variable, denoted by  $k_i$ , and its mean is

$$\lambda_i = \left[ \frac{p_c}{2}(x_{i-1} + x_{i+1}) + (1 - p_c)x_i \right] \lambda \eta, \quad (1)$$

where  $\lambda$  is the transmission power and  $\eta$  is the detection probability. For each slot, the receiver can only determine whether there are photons or not, without resolving the exact number. Hence, the output of slot  $i$  is  $y_i = 1$  if and only if  $k_i > 0$ ; otherwise,  $y_i = 0$ . These notations can be naturally extended to the 2-D case: we consider an  $m \times m$  rectangular array, where each pixel has four adjacent pixels, and a photon from a pixel falls into a particular adjacent pixel with a probability  $\frac{p_c}{4}$ .

The capacity of the Poisson channels, without considering crosstalk, has been well studied [5], [6]. With considering the crosstalk effect, in [7], Kachelmer and Boroson studied the soft decision capacity of an  $M$ -ary pulse-position modulation (PPM) receiver for the 1-D case. But such a capacity (more accurately we should call it information rate) with the PPM constraint is not the maximum information rate that the channel can achieve.

### B. Channel Capacity

Given transmission power  $\lambda$ , crosstalk probability  $p_c$ , and detection probability  $\eta$ , the *capacity* of an optical channel with 1-D crosstalk is the maximal expected number of bits that can be transmitted per slot, namely,

$$C(\lambda, p_c, \eta) = \lim_{m \rightarrow \infty} \max_{P_X} \frac{I(X; Y)}{m}, \quad (2)$$

where  $X, Y \in \{0, 1\}^m$  are the input/output sequences,  $P_X$  is the distribution of  $X$ , and  $I(X; Y)$  is the mutual information

between  $X$  and  $Y$ . Note that the properties of the optical channel keep unchanged if we scale  $\lambda$  and  $\eta$  together while keeping  $\lambda \eta$  fixed. It implies that  $C(\lambda, p_c, \eta) = C(\lambda \eta, p_c, 1)$ . In order to study the behaviors of channel capacity, we only need to consider the case that  $\eta = 1$ , and in this case we write the channel capacity as  $C(\lambda, p_c)$ .

### C. Information Rates

In practice, it is usually difficult to construct error-correcting codes to achieve the capacity of an optical channel with crosstalk, since the input distribution needs to be biased and correlated. Hence, besides channel capacities, we are also interested in the information rates that can be achieved with simple input distributions. In this paper, we study two types of input distributions: i.i.d. distributions and uniform distributions. Given  $\lambda$  and  $p_c$ , we call the maximal information rate with an i.i.d. input distribution as the *i.i.d. information rate*, and we call the maximal information rate with the uniform input distribution as the *symmetric information rate*. A question is how close these information rates are to the channel capacity so that we can construct error-correcting codes with near-capacity performance for optical channels with crosstalk.

## III. CHANNELS WITH 1-D CROSSTALK

An optical on-off-keying channel with 1-D crosstalk satisfies

$$P_{Y|X}(y_1^m | x_1^m) = \prod_{i=1}^m P(y_i | x_{i-1}^{i+1}) \quad (3)$$

where  $x_a^b = x_a x_{a+1} \dots x_b$  and  $x_0 = x_{m+1} = 0$ . It belongs to inter-symbol interference (ISI) channels, or more generally, channels with memory. The computation of the capacity of channels with memory has attracted much attention in the information theory literatures. However, for most types of channels with memory, computing the exact capacity is still an unsolved problem. Instead, there are extensive studies focusing on deriving bounds. A general approach to obtaining a lower bound on capacity is to calculate the information rate of the channel when its input is an order- $k$  Markov chain. By optimizing the distribution of the Markov chain and gradually increasing  $k$ , we get a lower bound approaching the channel capacity [8], [9]. Given a known Markov chain as the input, an efficient simulation-based method to estimate the information rate of a channel with memory was given in [10], which requires the generation of a very long output sequence based on the channel character. On the other hand, the upper bound on the capacity for channels with memory was studied by Vontobel and Arnold [11] and Yang et al. [12]. In what follows, we derive the lower and upper bounds on the capacity of an optical on-off-keying channel with 1-D crosstalk.

To obtain a lower bound, we assume that the input of the channel is an order- $k$  stationary Markov chain with transition probability  $\Psi$ . The information rate of the channel, denoted by  $R(\Psi)$ , is a lower bound of the channel capacity. It can be written as

$$R(\Psi) = \lim_{m \rightarrow \infty} \frac{\sum_{i=1}^m [H(x_i | x_1^{i-1}) - H(x_i | y_1^m, x_1^{i-1})]}{m},$$

and it is larger than

$$\lim_{m \rightarrow \infty} \frac{\sum_{i=k+1}^{m-d+1} [H(x_i|x_{i-k}^{i-1}) - H(x_i|x_{i-k}^{i-1}, y_{i-1}^{i-1+d})]}{m}.$$

This leads to

$$R(\Psi) \geq H(x_i|x_{i-k}^{i-1}) - H(x_i|x_{i-k}^{i-1}, y_{i-1}^{i-1+d}). \quad (4)$$

Here,  $d$  and  $k$  are small integers, and  $i$  in (4) satisfies  $1 \leq i - k, i - 1 + d \leq m$ .

Hence, a **lower bound** on the capacity for optical channels with 1-D crosstalk is

$$c_l(\lambda, p_c) = \max_{\Psi} H(x_i|x_{i-k}^{i-1}) - H(x_i|x_{i-k}^{i-1}, y_{i-1}^{i-1+d}) \quad (5)$$

$$\text{s.t. } \Psi \in \mathbb{R}^k \text{ with } \mathbf{0} \leq \Psi \leq \mathbf{1}.$$

Given the transition probability  $\Psi$ , we can compute the stationary distribution of the input Markov chain, denoted by  $u(\Psi)$ , and further compute the objective function. Although the objective function may not be concave with some parameters  $\lambda$  and  $p_c$ , we still can apply convex-programming techniques to obtain a local maximum as a lower bound on capacity. Numerical results imply that the local maximum is most likely the global maximum when the starting search point is  $\Psi_0 = [\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}]$ .

Now, let's consider the upper bound. According to [8], there exists a stationary distribution on the input that achieves the capacity of a channel with memory. Also, due to the symmetric property of optical channels with 1-D crosstalk, it can be proved that there exists an optimal input distribution  $P_X$ , which is stationary and symmetric. Based on this conclusion, we get

$$\begin{aligned} C(\lambda, p_c) &= \lim_{m \rightarrow \infty} \max_{P_X} \frac{\sum_{i=1}^m [H(y_i|y_1^{i-1}) - H(y_i|y_1^{i-1}, x_1^m)]}{m} \\ &\leq \max_{P_X} H(y_i|y_{i-t}^{i-1}) - H(y_i|x_{i-1}^{i+1}), \end{aligned} \quad (6)$$

for a small integer  $t$ , s.t.,  $P_X$  is stationary and symmetric.

Let  $\theta$  be the marginal distribution of  $P_X$  on  $x_{i-t-1}^{i+1}$ . Since  $P_X$  is stationary and symmetric,  $\theta$  is also stationary and symmetric, which can be written as linear constraints on  $\theta$ :  $\theta(x_{i-t-1}x_{i-t} \dots x_{i+1}) = \theta(x_{i+1}x_{i-t} \dots x_{i-t-1})$  and  $\theta(x_{i-t-1}^i) = \theta(x_{i-t}^{i+1})$ . Hence, we get the following **upper bound** on the capacity for optical channels with 1-D crosstalk.

$$c_u(\lambda, p_c) = \max_{\theta} H(y_i|y_{i-t}^{i-1}) - H(y_i|x_{i-1}^{i+1}) \quad (7)$$

s.t.  $\theta$  is a stationary and symmetric distribution.

Note that in order to get an upper bound, we need to get a global optimal solution for the optimization problem. Fortunately, the objective function in this problem is a concave function of  $\theta$ , and the constraints on  $\theta$  are linear. Hence, we can solve this optimization problem by using convex-programming techniques, and hence we get an upper bound on capacity.

Based on (5) and (7), we can get very tight bounds on the capacity of optical channels with 1-D crosstalk, e.g., by setting  $k = 6, d = 4, t = 5$ . Fig. 2 plots the channel capacities for a wide range of  $\lambda$  (logarithmic scale) with

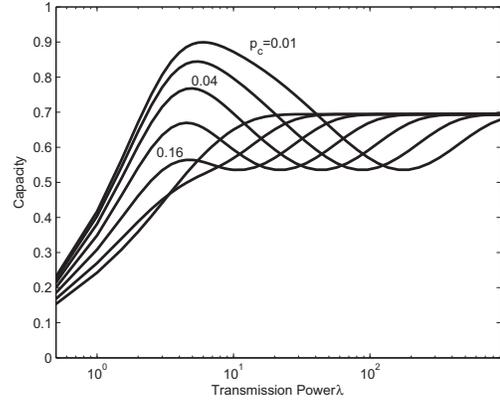


Fig. 2. Capacities of optical channels with 1-D crosstalk.

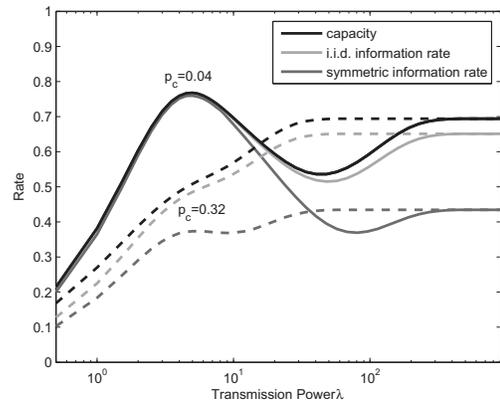


Fig. 3. Information rates of optical channels with 1-D crosstalk.

different  $p_c$  chosen from  $[0.01, 0.02, 0.04, \dots, 0.64]$ . It shows that the channel capacity converges to a constant that is independent of  $p_c$  as  $\lambda$  goes to infinity. With techniques from constrained coding, it can be proved that this capacity with an infinite transmission power is exactly 0.6942. An interesting observation is that as  $\lambda$  changes, the channel capacities with different values of  $p_c$  demonstrate different behaviors:

- When  $p_c \geq 0.22$ , the capacity is a non-decreasing function of  $\lambda$ . But in most practical applications, it barely meets such a high crosstalk probability.
- If  $p_c < 0.22$ , then the capacity has two local maximums, one with a low transmission power  $\lambda^*$  that can be excellently approximated by

$$\lambda^* \eta \simeq -2 \log_{10}(p_c) + 2, \quad (8)$$

and the other with an infinite transmission power. The first local maximum (with a low transmission power) is larger than the second one when  $p_c \leq 0.068$ .

Fig. 3 compares the channel capacity, the i.i.d. information rate, and the symmetric information rate when  $p_c = 0.04$  or  $0.32$ . As  $\lambda$  increases, the i.i.d. information rate converges to 0.6508 and the symmetric information rate converges to 0.4343, independent of the crosstalk probability  $p_c$ . It shows that the i.i.d. information rate is much closer to the capacity than the symmetric information rate, implying that in practice

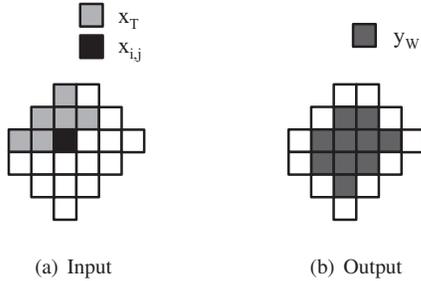


Fig. 4. The region to compute a lower bound on capacity with 2-D crosstalk.

we can use an i.i.d. input distribution to achieve near-capacity performance. As  $\lambda$  decreases, the gaps between both the information rates and the capacity decrease. In particular, when  $p_c$  is small, both the maximal i.i.d. information rate and the maximal symmetric information rate (where  $\lambda$  is not large) are very close to the maximal capacity.

#### IV. CHANNELS WITH 2-D CROSSTALK

In this section, we extend the discussions to optical channels with 2-D crosstalk, where the input and output are two arrays  $\mathcal{X}, \mathcal{Y} \in \{0, 1\}^{m \times m}$  with  $m$  large enough. In [13], Chen and Siegel introduced a method to compute upper bounds on the symmetric information rate of 2-D finite-state inter-symbol interference (ISI) channels, which appear, e.g., in magnetic and optical recording devices. Another approach that applies simulation-based methodologies, the “generalized belief propagation” algorithm, to estimate the symmetric information rate of 2-D ISI channels was proposed by Shental et al. [14]. In general, due to the high computational complexity, few studies have been conducted to obtain tight bounds on the capacity of 2-D channels with memory, except for the capacity of some special 2-D constraints. Specifically, the lower bounds on the capacity of some 2-D constraints were presented and analyzed in [15]–[17], based on either bit-stuffing encoders or tiling encoders; and the upper bounds were provided by Forchhammer and Justesen [18], Tal and Roth [19].

To compute a lower bound on the capacity for channels with 2-D crosstalk, it is computationally difficult to obtain an optimal input distribution. In the previous section, we showed that the i.i.d. information rate is very close to the capacity for the 1-D case. If this property also holds for the 2-D case, then we can use the i.i.d. information rate as a lower bound, where we assume that the probability of each input bit is  $q$ . As a result, we get the following **lower bound** on the capacity for optical channels with 2-D crosstalk:

$$r_l(\lambda, p_c) = \max_{0 \leq q \leq 1} H(x_{i,j}) - H(x_{i,j}|x_T, y_W). \quad (9)$$

Here,  $x_{i,j}$  indicates the input bit at pixel  $(i, j)$ , and  $x_T$  indicates the input bits at a set of pixels  $T$ . As shown in Fig. 4, we use  $T$  to denote a set of pixels on the left of or above  $(i, j)$ , and we use  $W$  to denote a set of pixels including  $(i, j)$ .

We then extend the idea for the 1-D case to compute an upper bound on the capacity for the 2-D case. In [19], it was demonstrated that there exists a stationary and symmetric input

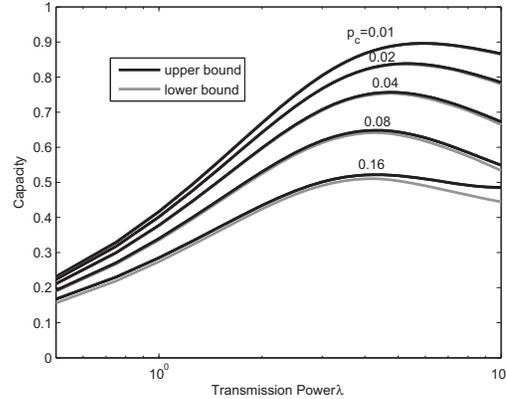


Fig. 5. Capacities of optical channels with 2-D crosstalk.

distribution that achieves the capacity of some special 2-D constraints. In fact, we can prove that this conclusion also holds for optical on-off-keying channels with 2-D crosstalk. Hence, we can get an **upper bound** on the capacity for optical channels with 2-D crosstalk:

$$c_u(\lambda, p_c) = \max_{\theta} H(y_{i,j}|y_T) - H(y_{i,j}|y_T, x_W) \quad (10)$$

s.t.  $\theta$  is a stationary and symmetric distribution.

The definition of  $T$  and  $W$  is the same as above. According to [19], the stationary and symmetric requirement on  $\theta$  can be represented by a group of linear constraints. Since the objective function is a concave function of  $\theta$ , we can solve the optimization problem with convex programming. However, the number of variables in the convex-programming problem has an exponential dependence on the size of  $T$  and  $W$ . In order to make the problem computationally practical, we need to consider relatively small regions for  $T$  and  $W$ , hence the resulting upper bound might not be very tight.

Fig. 5 shows the upper and lower bounds on the capacity of optical channels with 2-D crosstalk when  $0 \leq \lambda \leq 10$ . We see that when  $p_c$  and  $\lambda$  are small, the upper and lower bounds are tight. But their gap increases as  $p_c$  or  $\lambda$  increases. In fact, the capacity of optical channels with 2-D crosstalk has very similar behaviors as that of optical channels with 1-D crosstalk. If  $p_c$  is small, the communication limit (the maximal capacity) can be achieved with a low transmission power  $\lambda^*$ , which can be excellently approximated by (8), consistent with the 1-D case. With techniques of computing the bounds on the capacity of some 2-D constraints, we can prove that as the transmission power  $\lambda$  increases, the capacity of optical channels with 2-D crosstalk converges to a constant  $C_\infty \in [0.56394, 0.6126]$ , which is independent of  $p_c$ .

Fig. 6 shows the lower and upper bounds on the i.i.d. information rate and the symmetric information rate when  $p_c = 0.04$  or  $0.32$ . From this figure, we observe that the bounds on the i.i.d. information rate are tight when  $\lambda$  is not large. In general, given  $\lambda$  and  $p_c$ , the optimal i.i.d. input distribution is more efficient than the uniform input distribution for achieving high information rate, especially when  $p_c$  or  $\lambda$  is large.

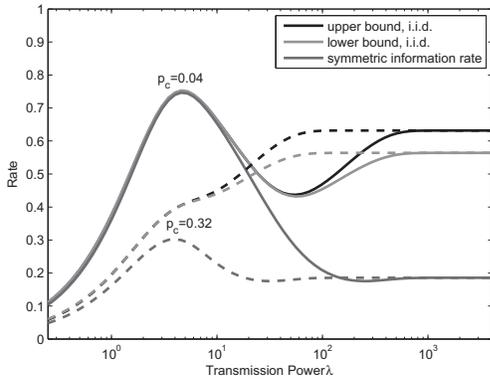


Fig. 6. Information rates of optical channels with 2-D crosstalk.

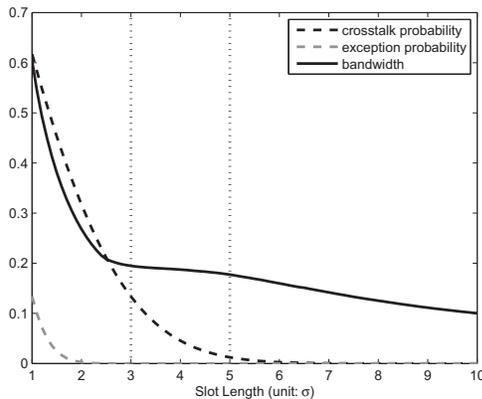


Fig. 7. Bandwidth versus slot length when photons are emitted at the centers of slots with Gaussian noises  $\mathcal{N}(0, \sigma)$ .

## V. THE EFFECT OF SLOT LENGTH

In many optical communication systems, the slot length (or pixel size) is an important parameter that significantly affects the performance of the systems. In this section, we investigate how to select the slot length appropriately for maximizing the bandwidth of optical channels with 1-D crosstalk (we only show a simple scenario due to the space limit). A similar analysis can be extended to optical channels with 2-D crosstalk, e.g., the Gaussian-beam channels studied in [20].

Let  $\ell$  be the length of each slot. Typically, the crosstalk probability  $p_c$  is a non-increasing function of  $\ell$ , denoted by  $p_c(\ell)$ . We want to find the slot length  $\ell$  that maximizes

$$B(\ell) = \frac{\max_{\lambda} C(\lambda, p_c(\ell), \eta)}{\ell}. \quad (11)$$

For instance, if the slot is in the time dimension, then  $B(\ell)$  is the bandwidth of the channel, i.e., the maximal number of bits that can be transmitted per second.

We consider a simple scenario that all the photons are emitted at the centers of slots, but each photon is shifted by a Gaussian noise  $z \sim \mathcal{N}(0, \sigma)$  when it arrives at the receiver. In this scenario, the crosstalk probability is given by  $p_c(\ell) = P_r(|t| \geq \frac{\ell}{2})$ . The probability that the crosstalk happens between non-adjacent slots is given  $p_e(\ell) = P_r(|t| \geq \frac{3\ell}{2})$ , and

we call it the exception probability, which is required to be ignorable. Fig. 7 shows the relation between the bandwidth  $B(\ell)$  and the slot length  $\ell$ . In order to make the exception probability ignorable, we should let  $\ell > 3\sigma$ . On the other hand, when  $\ell > 5\sigma$ , the bandwidth of the channel quickly decreases as the slot length  $\ell$  increases. A reasonable choice for the slot length is between  $3\sigma$  and  $5\sigma$ ; and in this region, the bandwidth is not very sensitive to the slot length. For example, if we choose  $\ell = 4\sigma$ , then the crosstalk probability is about 0.06 and the optimal transmission power  $\lambda$  satisfies  $\lambda\eta = 4.4$  photons per slot, with  $\eta$  the detection probability.

## REFERENCES

- [1] B. Moision and W. Farr, "Communication limits due to photon detector jitter," *IEEE Photonics Technology Letters*, vol. 20, pp. 715–717, 2008.
- [2] G.P. Agrawal, *Fiber-Optic Communication Systems*, vol. 222, John Wiley & Sons, 2010.
- [3] J. M. Fini, B. Zhu, T. F. Taunay, M. F. Yan, and K. S. Abedin, "Crosstalk in multicore fibers with randomness: gradual drift vs. short-length variations," *Opt. Express*, vol. 20, no. 2, pp. 949–959, 2012.
- [4] N. Chandrasekaran, and J. H. Shapiro, "Turbulence-induced crosstalk in multiple-spatial-mode optical communication," *CLEO: Science and Innovations*, Optical Society of America, 2012.
- [5] I. Bar-David, "Communication under the Poisson regime," *IEEE Trans. Info. Theory*, vol. 15, pp. 31–37, Jan. 1969.
- [6] A. Lapidot and S. M. Moser, "On the capacity of the discretetime Poisson channel," *IEEE Trans. Info. Theory*, vol. 55, no. 1, pp. 303–322, Jan. 2009.
- [7] A.L. Kachelmeyer, D.M. Boroson, "Efficiency penalty of photon-counting with timing jitter," *Optical Engineering and Applications*, International Society for Optics and Photonics, 2007.
- [8] J. Chen and P. H. Siegel, "Markov processes asymptotically achieve the capacity of finite state intersymbol interference channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1295–1303, Mar. 2008.
- [9] P. O. Vontobel, A. Kavcic, D. M. Arnold, and H. A. Loeliger, "A generalization of the Blahut Arimoto algorithm to finite-state channels," *IEEE Trans. Inf. Theory*, vol. 54, pp. 1887–1918, 2008.
- [10] D. M. Arnold, H. A. Loeliger, et al. "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.
- [11] P. O. Vontobel and D. M. Arnold, "An upper bound on the capacity of channels with memory and constraint input," in *Proc. IEEE Information Theory Workshop*, Cairns, Australia, pp. 147–149, Sep. 2001.
- [12] S. Yang, A. Kavcic, and S. Tatikonda, "Feedback capacity of finite-state machine channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 799–810, Mar. 2005.
- [13] J. Chen and P. H. Siegel, "On the symmetric information rate of two dimensional finite state ISI channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 227–236, Jan. 2006.
- [14] O. Shental, N. Shental, et al. "Discrete-input two-dimensional Gaussian channels with memory: Estimation and information via graphical models and statistical mechanics," *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1500–1513, Apr. 2008.
- [15] S. Halevy, J. Chen, R. M. Roth, P. H. Siegel, and J. K. Wolf, "Improved bit-stuffing bounds on two-dimensional constraints," *IEEE Trans. Inf. Theory*, vol. 50, pp. 824–838, 2004.
- [16] I. Tal and R. M. Roth, "Bounds on the rate of 2-D bit-stuffing encoders," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2561–2567, 2010.
- [17] A. Sharov and R. M. Roth, "Two-dimensional constrained coding based on tiling," *IEEE Trans. Inf. Theory*, vol. 56, pp. 1800–1807, 2010.
- [18] S. Forchhammer and J. Justesen, "Bounds on the capacity of constrained two-dimensional codes," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2659–2666, 2000.
- [19] I. Tal, and R. M. Roth, "Convex programming upper bounds on the capacity of 2-D constraints," *IEEE Trans. Inf. Theory*, vol. 57, pp. 381–391, 2011.
- [20] Y. Kochman and G. W. Wornell, "On high-efficiency optical communication and key distribution," in *Proc. Inform. Theory Appl. (ITA)*, Feb. 2012.