6.962 Week 5 Summary:

The Poisson Channel

Presenter: Won S. Yoon

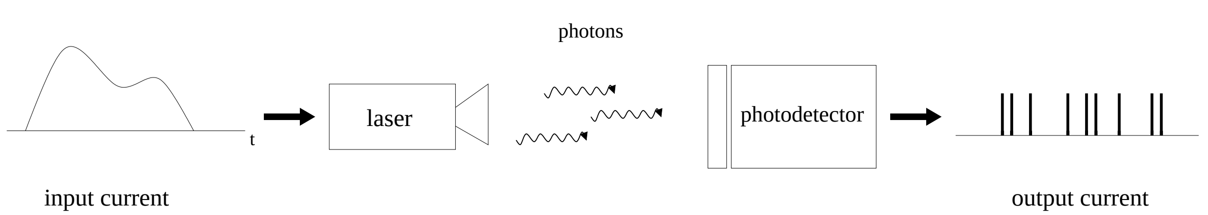
March 8, 2001

1 Introduction

The Poisson channel was originally developed around 20 years ago as a model for an optical communication link. Since then, a rather small subset of Information Theorists have studied this model quite extensively and with remarkable technical success. In practice, however, these theoretical achievements have yet to significantly impact real-world systems. One explanation for this is that the enormous inherent bandwidth and relatively low noise of optical fibers has allowed engineers to get by with simple coded or uncoded techniques. Nonetheless, we might expect that the ever-increasing bandwidth demands of consumers will eventually make intelligent coding methods necessary in future optical networks.

1.1 Optical Communication

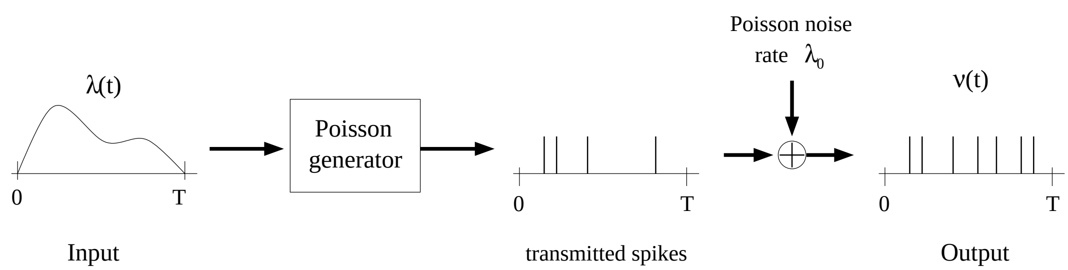
In order to motivate the Poisson channel, we first describe an optical communication link (shown in the figure below). At the transmitter, a laser emits a stream of discrete photons with a (time-varying) rate which is proportional to the amplitude of the input current. The receiver consists of a photodetector which is able to determine the precise arrival times of individual photons.



The stream of photons is typically modeled to be an inhomogeneous Poisson point process with time-varying intensity proportional to the input current. In addition, there may be spontaneous background noise ("dark current") in the form of random photons generated in the laser. This noise is modeled as an additive Poisson process with some fixed rate. Based on these assumptions, people have developed a mathematical model called the Poisson channel.

1.2 The Poisson Channel Model

The input to this channel is a waveform $\lambda(t)$, assumed to be nonnegative. The output of the channel is an inhomogeneous Poisson process with intensity $\lambda(t) + \lambda_0$. The latter term represents additive Poisson noise of intensity λ_0.



So there are two sources of noise in this channel: 1. the probabilistic nature in which spikes are generated at the transmitter, and 2. the extraneous spikes due to background emission (when referring to the generic Poisson channel, we will often use the term "spikes" instead of "photons"). More precisely, for an input $\lambda(t)$, the channel output in a small time increment $(t, t + \Delta)$ has the distribution:

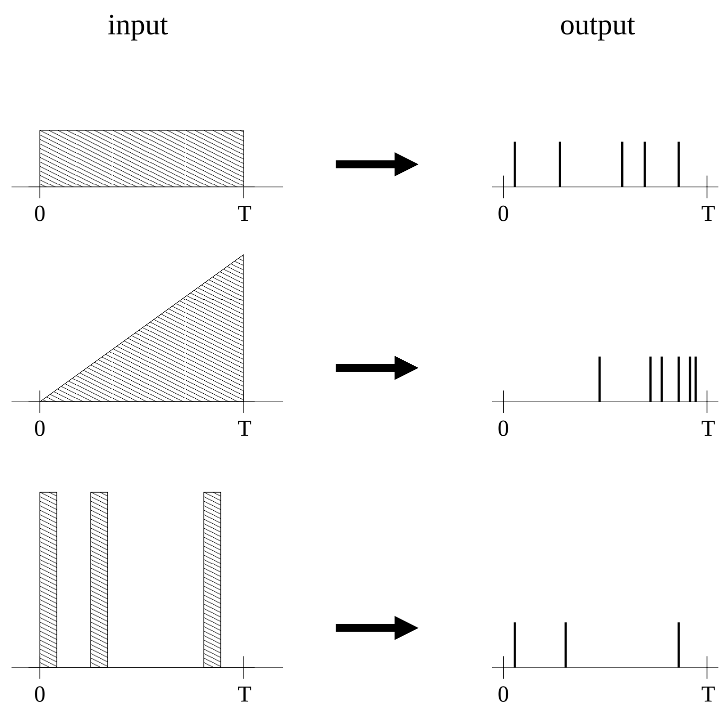
$$\text{1 spike with probability: } \Lambda e^{-\Lambda}$$
$$\text{no spikes with probability: } e^{-\Lambda}$$
$$\geq 2 \text{ spikes with probability: } o(\Delta)$$

where

$$\Lambda = \int_t^{t+\Delta} (\lambda(\tau) + \lambda_0(\tau))dt$$

is the average number of spikes received in the interval $(t, t + \Delta)$.

Some simple examples of the input-output behavior of this channel are shown below (for the case of no background noise). A constant input signal will generate spikes which are statistically uniformly spread over time. As we increase the input amplitude, the frequency of spikes increases. Finally, notice that very short pulses of very large amplitude will be reproduced at the output with high probability.



1.3 Encoding and Decoding

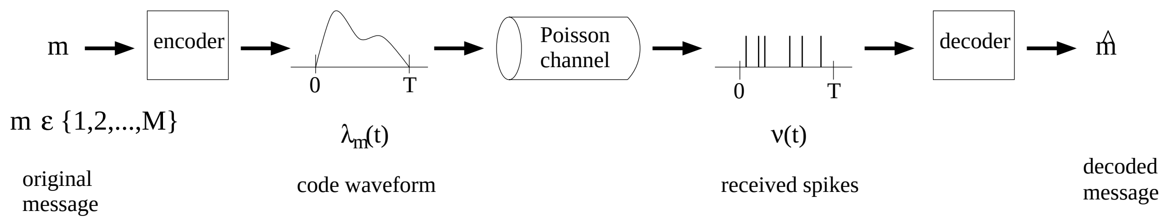
The usual Information-Theoretic notion of channel coding can now be applied to this model. The encoder tries to communicate one of M possible messages via a set of M "code waveforms" which use the channel for some duration T. The decoder observes the arrival times of received photons and tries to determine which message was sent. Specifically, a code with parameters (M, T, P_e) is defined by the following:

- a set of M code waveforms $\{\lambda_m(t)\}_{m=1}^{M}$, which are non-zero only in $[0, T]$.
- a decoding function: $D(\nu_0^T) = \hat{m} \in \{1, 2, \ldots, M\}$ (where ν_0^T denotes the received photon arrival times in the interval $[0, T]$).

- the average probability of error per message is

$$P_e = \frac{1}{M} \sum_{m=1}^{M} Pr\{D(\nu_0^T) \neq m \mid \lambda_m(\cdot)\}$$

The figure below illustrates the encoding and decoding operations.



The rate of the code is $\frac{\log M}{T}$ bits/sec. A rate R is said to be *achievable* if for all $\epsilon > 0$, there exists a code with sufficiently large T and $M \geq e^{RT}$ such that $P_e \leq \epsilon$. The *channel capacity* C is the supremum of all achievable rates. For the Poisson channel, capacity is equal to

$$C = \lim_{T\to\infty} \sup_{p_\lambda(\lambda_0^T)} \frac{1}{T} I(\lambda_0^T; \nu_0^T) \quad \text{bits/sec}$$

The capacity will depend on the types of constraints we impose on the input waveform $\lambda(t)$. We will always constrain the input to be peak-limited, so that $\lambda(t) \leq A$. This makes sense both practically and theoretically, since otherwise we could generate a photon at an arbitrarily precise time and hence convey an arbitrarily large amount of information (assuming the bandwidth of the receiver and channel to be unbounded). We may also want to constrain the average value of the input, so that $\frac{1}{T}\int_0^T \lambda_m(t)dt \leq \sigma A$, where $0 \leq \sigma \leq 1$. We may also want to constrain the bandwidth of the input waveform. All of these scenarios have been addressed to some extent in the literature.

1.4 Summary of Results

There have been numerous studies of the Poisson channel, but we list only a few here. (We will only be discussing the two papers highlighted in bold). A more comprehensive listing of references can be found in [4].

Single-user channel:

- Kabanov '78 [2], Davis '80 [1]: found capacity under peak and average input constraints.
- **Wyner '88 [8]:** obtained the exact error exponent for all rates below capacity and constructed a code which achieves the optimal error exponent.
- Lapidoth and Shamai '91 [6]: considered various input bandwidth constraints and found upper and lower bounds on capacity.
- Lapidoth '93 [3]: considered noiseless feedback and found the exact error exponent.

Multi-user channel:

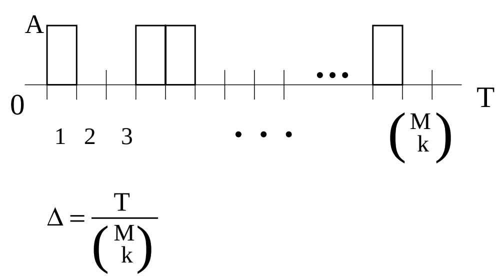
- **Lapidoth and Shamai '98 [4]:** studied the Poisson MAC and found the capacity region for 2 users. Showed that total throughput is bounded in the number of users.
- Lapidoth and Shamai are currently studying a Poisson MAC with feedback.

2 The Single-User Poisson Channel

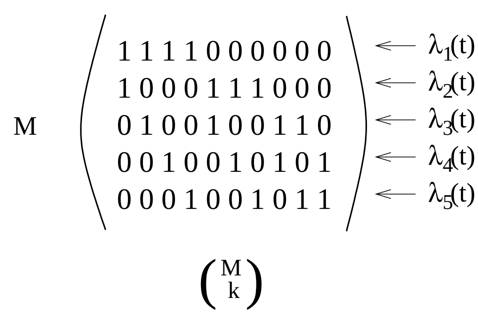
Wyner's results [8] are significant in that very few channels have been characterized so precisely (the only other such channel being the infinite-bandwidth AWGN). An interesting feature of Wyner's capacity-achieving code is that it consists entirely of "on-off" input waveforms which vary infinitely fast. That is, the optimal code consists of PAM-like waveforms which take on only two possible values (zero or the peak value) and with infinitesimally small pulse width. The decoder observes the photon stream and selects the message $\hat{m}$ for which the corresponding code waveform has the maximum number of received photons during its "on" periods (this corresponds to maximum likelihood detection).

Intuitively, we can see why Wyner's code works so well. The amplitude of the input waveform at some time t essentially determines the *probability* of a photon being released at time t; the larger the amplitude, the more likely a photon is released. At every time instant, the receiver can only distinguish between the *presence* or the *absence* of a photon. In order to maximize this distinction, the transmitter should send either at the peak amplitude or zero, and alternate between them as quickly as possible (to maximize data rate).

We now describe Wyner's optimal code in more detail. Let $\mathcal{A}$ be the M-by-$\begin{pmatrix} M \\ k \end{pmatrix}$ binary matrix, the columns of which are all the $\begin{pmatrix} M \\ k \end{pmatrix}$ possible binary M-vectors with exactly k ones and $(M-k)$ zeros. Denote the (m,j)-th entry of $\mathcal{A}$ by a_{mj}. By symmetry, the total number of nonzero entries in each row is $(\frac{k}{M})\begin{pmatrix} M \\ k \end{pmatrix}$. Now partition the coding interval $[0,T]$ into $\begin{pmatrix} M \\ k \end{pmatrix}$ subintervals, each of duration $T/\begin{pmatrix} M \\ k \end{pmatrix}$ seconds. Then for the m-th code waveform, let its amplitude in the j-th subinterval be constant and equal to A times the entry a_{mj} of the matrix $\mathcal{A}$. That is, $\lambda_m(t) = Aa_{mj}$ for t in the j-th subinterval of $[0,T]$. An example of 5 codewords satisfying an average input constraint of $\frac{2}{5}A$ is shown below.



code waveforms for M = 5, k = 2



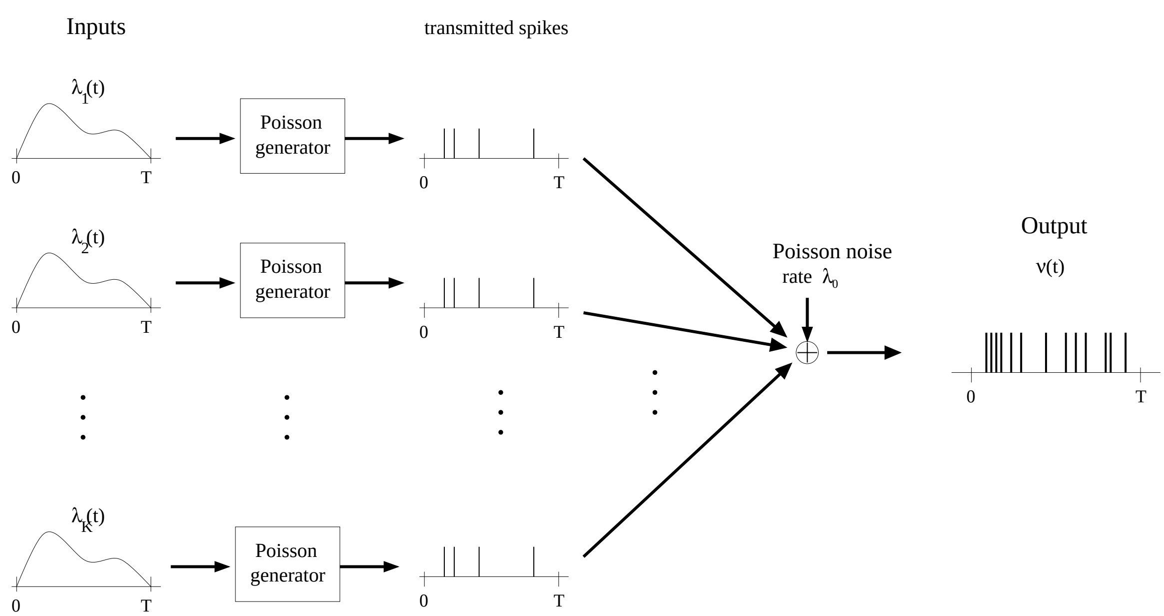
By construction, the average value of each code waveform is $\frac{Ak}{M}$. Now fix $\frac{k}{M} = \sigma$ (so that we satisfy the average input constraint), and $M = 2^{RT}$ (so that we maintain a rate of R bits per second) and let $T \to \infty$. Notice that the length of each of the $\begin{pmatrix} 2^{RT} \\ \sigma 2^{RT} \end{pmatrix}$ subintervals goes to zero with T. Intuitively, the infinitesimally short pulses result from the need to maintain the same average value (number of ones) for each code waveform, which in turn maintains a certain amount of Euclidean distance between the waveforms. Although the duration of each "on-off" period is going to zero, the integral of the waveform remains constant at σA. Wyner then calculates the error probability of this code under the maximum-likelihood detection rule described above.

In part II of his two-part paper [8], Wyner derives an upper bound on capacity that exactly matches the performance of his code. Hence, this code is exponentially optimal, that is, its probability of error decays exponentially in code duration with the optimal error exponent.

The major drawback of this code, however, is that the channel and the receiver must have infinite bandwidth in order to discern the infinitesimally small changes in the code waveforms. Numerous subsequent authors have looked at imposing various bandwidth constraints into the problem, and have obtained lower and upper bounds to capacity [6]. A complete characterization of capacity and optimal codes for the bandwidth-constrained case remains an open problem.

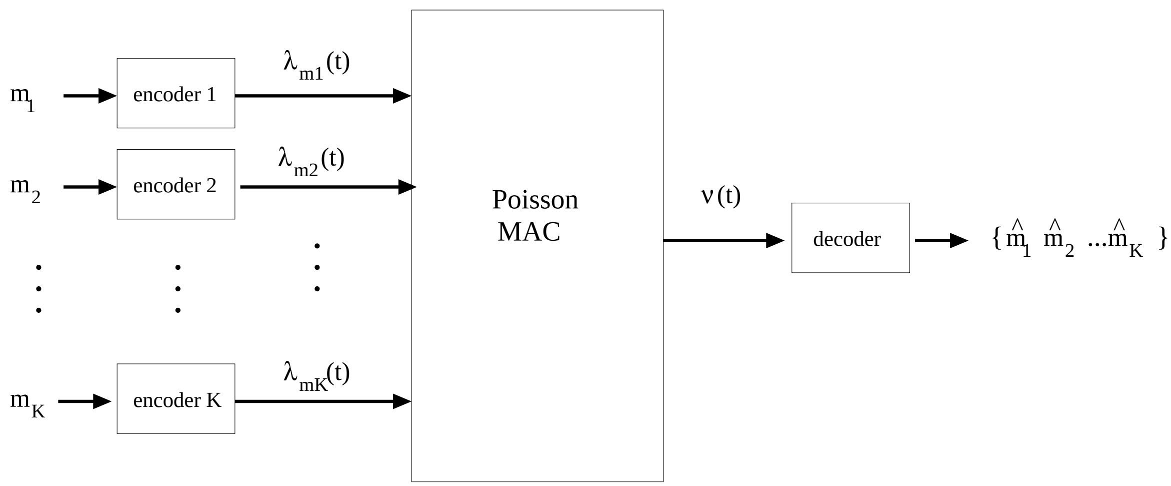
3 The Multi-access Poisson Channel

The multi-access Poisson channel model introduced by Lapidoth and Shamai [4] is shown in the figure below. There are K independent inputs to this channel and one output.



The output can simply be thought of as the superposition of the outputs of K independent single-user Poisson channels. That is, for inputs $\lambda_1(t), \lambda_2(t), \ldots, \lambda_K(t)$ (where the subscripts now index the user) the output of the channel is an inhomogeneous Poisson process, $\nu(t)$, of intensity $\lambda(t) = \sum_{i=1}^{K} \lambda_i(t)$. (A slightly different multi-access model for optical communications has been proposed by others [7], [5], but we will not discuss those here).

As in the usual MAC setup, each encoder wants to independently communicate a message and the decoder tries to decode each of the K messages. An abstract encoding/decoding setup is shown below.



In extending Wyner's result from the single-user channel, Lapidoth and Shamai [4] showed that the capacity region for the Poisson MAC can be achieved by using input waveforms which are binary and which have infinitely-fast time variations. For the case of 2 encoders, Lapidoth and Shamai obtained the exact capacity region. In the general case of K users, they showed that the maximum total throughput is monotonically increasing in the number of users and that it is bounded from above.

This last result may be a little surprising to anyone familiar with the Gaussian MAC, where the maximum total throughput grows unbounded as the log of the number of users. Intuitively, this difference between the Poisson and the Gaussian MACs can be partially understood by comparing the entropy of the respective channel outputs. In a Gaussian MAC, the capacity-achieving output contains the sum of K i.i.d. Gaussians and thus has entropy which grows logarithmically in the number of users. Loosely speaking, the information content of the output increases as the log of the number of users. In contrast, the Poisson

MAC has a capacity-achieving output which is a Poisson process with an intensity λ equal to the sum of the K i.i.d. binary inputs. A simple calculation shows that a Poisson process of intensity λ has entropy rate $\lambda(1 - \log \lambda)$ bits/sec. This does *not* monotonically increase with the input λ, and is in fact concave with a peak at input intensity $\frac{1}{e}$. Therefore, adding more inputs to a Poisson MAC eventually saturates the entropy rate (and hence the information content) of the output.

4 Other Applications

Given the theoretical success of the Poisson channel, it is tempting to look for other applications of this model. One such effort I have recently made is to model a biological neuron as a noisy communication channel. It turns out that the Poisson channel is a good model for the type of noise that exists in a neural synapse.

References

[1] M.A. Davis. Capacity and cutoff rate for poisson-type channels. *IEEE Trans. Info. Theory*, IT-26, 1980.

[2] Y.M. Kabanov. The capacity of a channel of the poisson type. *Theory of Probab. Appl.*, 23, 1978.

[3] A. Lapidoth. On the reliability function of the ideal poisson channel with noiseless feedback. *IEEE Trans. Info. Theory*, IT-39, 1993.

[4] A. Lapidoth and S. Shamai. The poisson multiple-access channel. *IEEE Trans. Info. Theory*, IT-, 1998.

[5] P. Narayan and D.L. Snyder. Signal set design for band-limited memoryless multiple-access channels with soft-decision demodulation. *IEEE Trans. Info. Theory*, IT-33, 1987.

[6] S. Shamai and A. Lapidoth. Bounds on the capacity of a spectrally constrained poisson channel. *IEEE Trans. Info. Theory*, IT-39, 1993.

[7] S. Verdu. Multiple-access channels with point-process observations: optimum demodulation. *IEEE Trans. Info. Theory*, IT-32, 1986.

[8] A.D. Wyner. Capacity and error exponent for the direct detection photon channel - parts i, ii. *IEEE Trans. Info. Theory*, IT-34, 1988.