The Poisson Channel

Won S. Yoon

6.962: Graduate Seminar in Area I
M.I.T.

March 8, 2001

Outline

<u>I. Introduction</u>

II. Single-User Poisson Channel

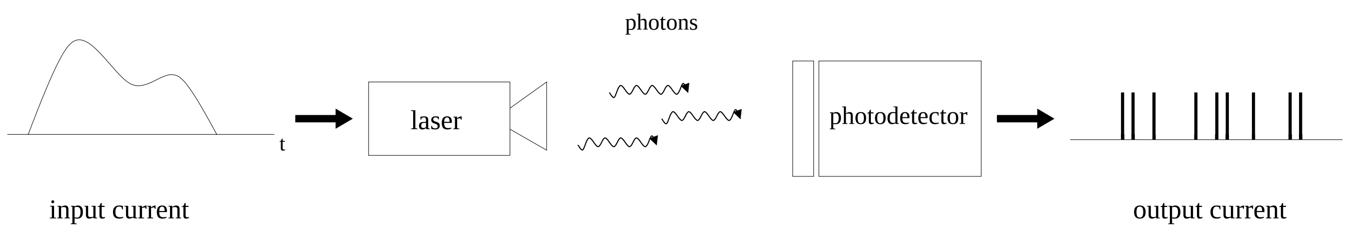
III. Multi-access Poisson Channel

IV. Other Applications

Introduction

- The Poisson channel was introduced around 20 years ago as a model for optical communication.
- Since then, Information-Theorists have studied it extensively.
- However, the theory has yet to have an impact in practice. One possible reason is that the enormous inherent bandwidth of optical fibers has made sophisticated coding techniques unnecessary.

Optical Communication Link



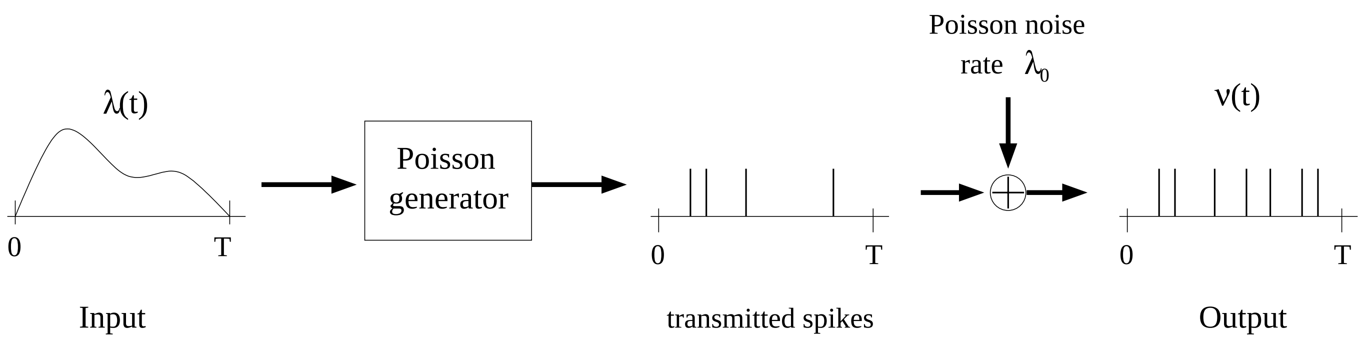
Transmitter (laser): emits photons at a (time-varying) rate which is proportional to the amplitude of the input current.

Receiver (photodetector): detects the arrival times of photons.

Noise: two sources of noise in the laser,

- The laser generates photons according to a random process.
- Background noise: spontaneously emitted photons ("dark current").

The Poisson Channel Model



Input: waveform $\lambda(t)$, (non-negative).

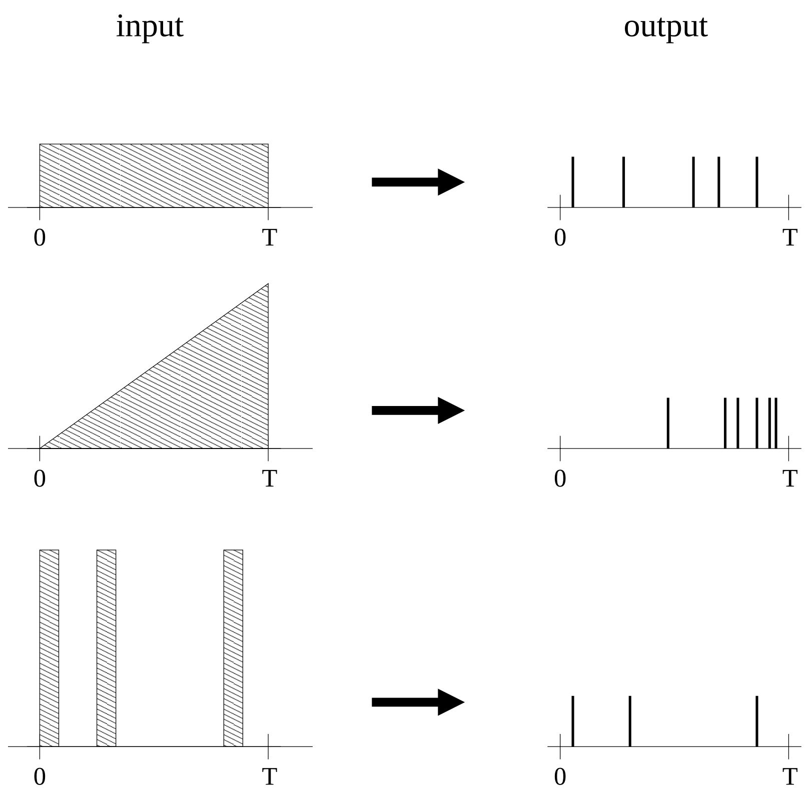
Output: Poisson point process $\nu(t)$ with intensity $\lambda(t) + \lambda_0$.

Noise:

- Randomness in generating spikes from the input $\lambda(t)$.
- Random additive spikes: Poisson with intensity λ_0.

The Poisson Channel (cont.)

Examples of input/output behavior:



The Poisson Channel (cont.)

More precisely:

For input $\lambda(t) = A$, the output in a small time interval $(t, t+\Delta)$ is:

$$\begin{aligned} 1 \text{ spike with probability: } & A\Delta e^{-A\Delta} \\ \text{no spikes with probability: } & e^{-A\Delta} \\ \geq 2 \text{ spikes with probability: } & o(\Delta) \end{aligned}$$

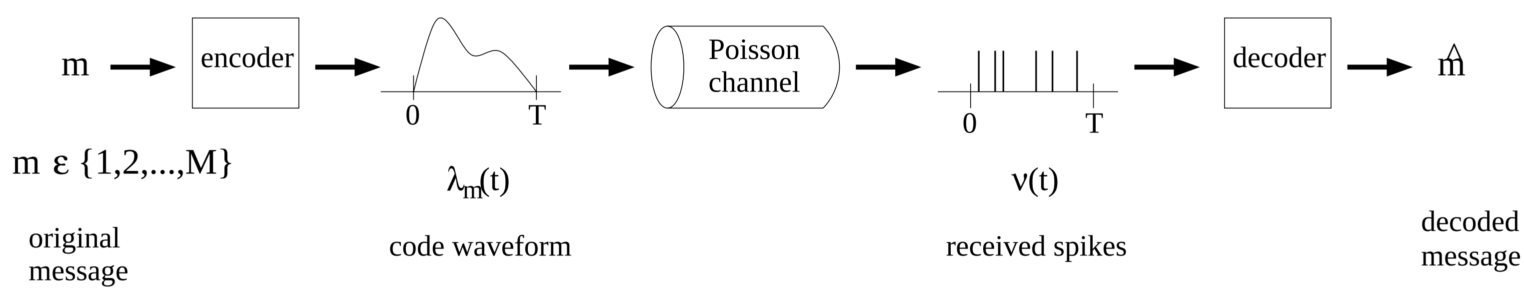
This determines the channel's input-output mutual information:

$$I(\nu(t); \lambda(t)) = H(\nu(t)) - H(\nu(t) \mid \lambda(t))$$

Coding for the Poisson Channel

A code with parameters (M, T) consists of:

1. An index set: {1,2, . . . , M}.
2. Code waveforms: $\{\lambda_m(t)\}_{m=1}^{M}$, where $\lambda_m(t) \geq 0$, $t \in [0, T]$.
3. A decoding function: $D(\nu_0^T) = \hat{m} \in \{1, 2, \ldots, M\}$.



Performance Issues

For an (M, T) code with codewords $\{\lambda_m(t)\}$ and decoder $D(\cdot)$:

- The probability of error is:

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

- The rate of the code is: $\frac{\log_2 M}{T}$ bits/sec.

A rate R is said to be *achievable* if for all $\epsilon > 0$, there exists a code with sufficiently large T and $M \geq 2^{RT}$ such that $P_e \leq \epsilon$.

The *channel capacity* C is defined to be the supremum of all achievable rates, and is equal to:

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

Input Constraints

<u>Peak value:</u> $0 \leq \lambda(t) \leq A.$

<u>Average value:</u> $\frac{1}{T}\int_0^T \lambda(t)dt \leq \sigma A, \quad 0 < \sigma \leq 1.$

Notice that *without* a peak constraint, the capacity is *infinite*.

Example: we can use impulse-like inputs and generate a spike with arbitrarily high time precision.

Preview of Results

Single-user Poisson channel

Wyner '88:

- Found exact error exponent under peak and average constraints.
- Constructed a code which achieves the optimal error exponent.

Multi-user Poisson channel

Lapidoth and Shamai '98:

- Found the capacity region for a 2-user MAC.
- For a general K-user MAC, showed that the maximum total throughput is bounded in the number of users.

Outline

I. Introduction

II. Single-User Poisson Channel

III. Multi-access Poisson Channel

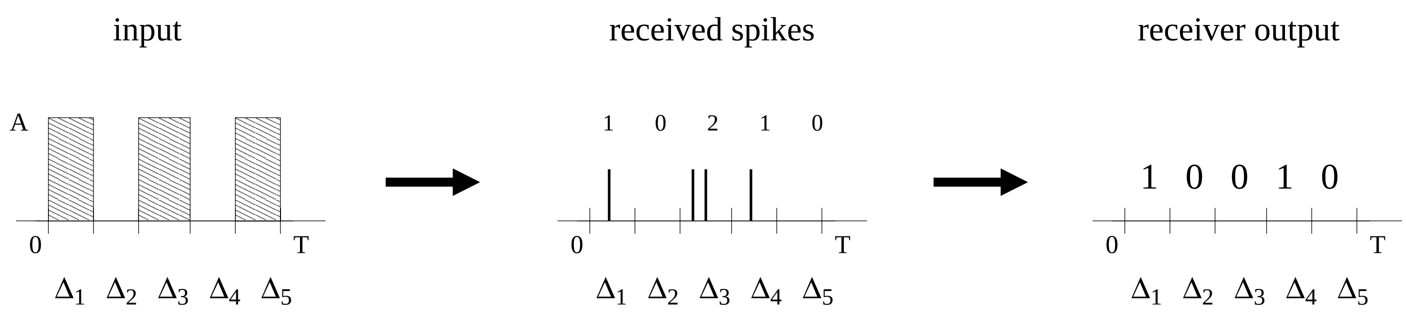
IV. Other Applications

Let's Make a Code!

Intuition: use very large pulses of very short duration.

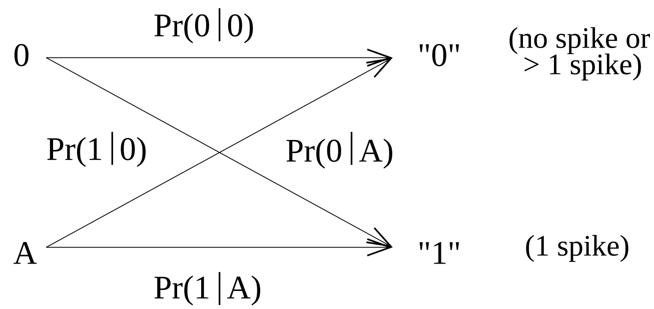
- Discretize the code interval $[0, T]$ into segments of duration Δ.
- In each Δ, let the input take one of two values, 0 or A.
- In each Δ, let the receiver distinguish between only two events:

$$\{\text{receive exactly 1 spike}\} \Longrightarrow \text{output "1"}$$
$$\{\text{receive zero or} \geq 2 \text{ spikes}\} \Longrightarrow \text{output "0"}$$



Lower Bound on Capacity

By the previous assumptions, and by the memoryless property of the Poisson process, each Δ-segment becomes a *binary channel*:



$$\Pr(1|0) = \lambda_o \Delta\, e^{-\lambda_o \Delta}$$

$$\Pr(1|A) = (A+\lambda_o)\Delta\, e^{-(A+\lambda_o)\Delta}$$

This gives us a lower bound to capacity:

$$C_{Poisson} \geq \max \frac{I(X_\Delta; Y_\Delta)}{\Delta}$$

where the max. is over all input distributions $p(X_\Delta)$ s.t. $E[X_\Delta] \leq \sigma A$.

It turns out that as $\Delta \to 0$, this lower bound is *exactly* the capacity.

Code Design

Recall that we have made the following simplifications:

1. Discretize time into Δ-segments, for some Δ.
2. Constrain input waveforms to be binary $\{0, A\}$ in each Δ.

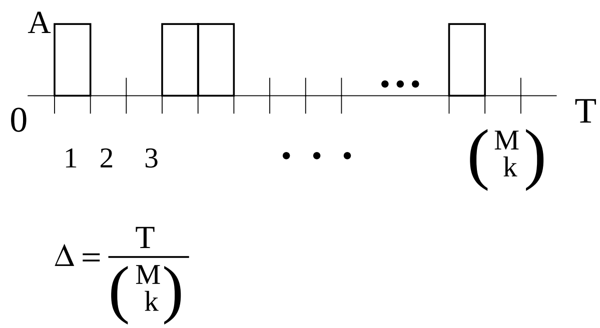
We want to design M waveforms in the interval $[0, T]$ which have maximum Euclidean distance subject to the previous constraints.

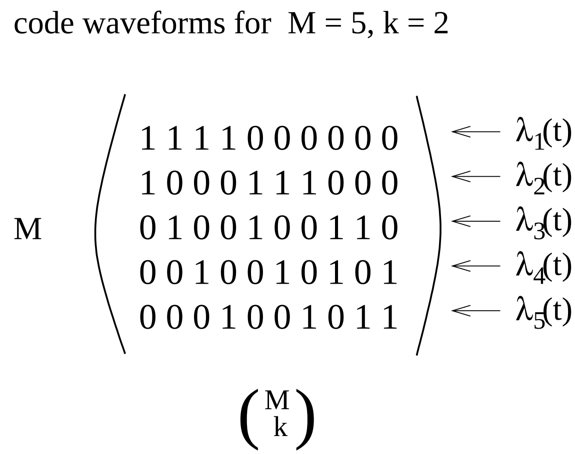
Wyner and Landau '84:

- Obtained an upper bound on the minimum Euclidean distance for any set of waveforms satisfying the previous assumptions.
- Constructed a set of waveforms which achieve that upper bound.

Wyner's Code

Construct the M code waveforms as follows:





- Let $M = 2^{RT}$ and let $k = qM$ (for some $q \leq \sigma$).
  Notice that each $\lambda_m(t)$ satisfies: $\frac{1}{T}\int_0^T \lambda_m(t)dt = \frac{k}{M}A \leq \sigma A$.
- Let $T \to \infty$: notice that $\Delta \to 0$.

Decoder: pick $\hat{m}$ such that $\lambda_{\hat{m}}(t)$ has the maximum number of received spikes during its "on" periods (ML detection).

Performance Analysis

The probability of error for this code is bounded by:

$$P_e \leq \exp\{-\mathrm{T}(\mathrm{Aq} - \mathrm{Aq}^{(1+\rho)} - \rho\mathrm{R})\}$$

Minimizing w.r.t. $\rho \in [0,1]$ and $q \in [0,\sigma]$ yields the tightest bound.

This gives us a lower bound on the optimal error exponent:

$$E^*(R) \geq Aq - Aq^{(1+\rho)} - \rho R$$

(and also a lower bound on capacity).

Upper Bound on the Error Exponent

In part II of his two-part paper, Wyner derives an upper bound on the error exponent which coincides with the lower bound.

Therefore, the optimal probability of error for this channel is:

$$P_e^* = \exp\{-\mathrm{T}(\mathrm{Aq} - \mathrm{Aq}^{(1+\rho)} - \rho\mathrm{R}) + \mathrm{o}(\mathrm{T})\}$$

and this is asymptotically achieved by Wyner's code as $T \to \infty$.

Capacity

For a Poisson channel with a peak input A, avg. input σA, and noise intensity λ_0, the capacity is: [Kabanov/Davis/Wyner]

$$C = A[q^*(1+s)\log(1+s) + (1-q^*)s\log s - (q^*+s)\log(q^*+s)]$$

where

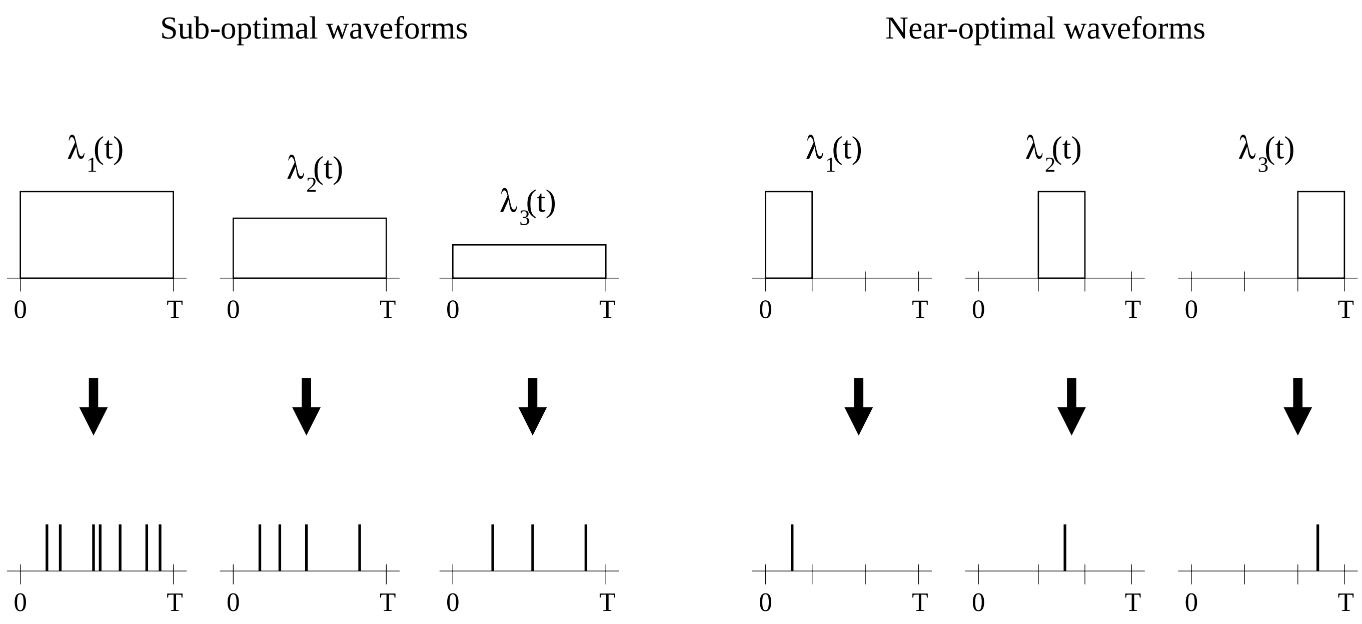
$$\begin{aligned} s &= \frac{\lambda_0}{A} \\ q^* &= \min(\sigma, q_0(s)) \\ q_0(s) &= \frac{(1+s)^{(1+s)}}{s^s e} - s \end{aligned}$$

- When $s = \lambda_0 = 0$: $\underline{C = Aq^* \log \frac{1}{q^*}}$, where $q^* = \min(\sigma, e^{-1})$.
- When $\sigma = 1$ and $s = \lambda_0 = 0$: $\underline{C = Ae^{-1}}$.

Discussion of Wyner's Results

The optimal input waveforms look like a sequence of spikes.

Intuition: this makes the received waveforms as distinct as possible.



Other Results

Single-user channel:

- Kabanov ‘78, Davis ‘80: capacity for peak and avg. constraints.
- Lapidoth and Shamai ‘91: upper and lower bounds on capacity for various input bandwidth constraints. Showed that band-limited inputs *strictly decrease* capacity.
- Lapidoth ‘93: exact error exponent for noiseless feedback.

Bandwidth Constraints

Wyner's optimal code requires the transmitter and receiver to have infinite bandwidth. Can we use finite bandwidth?

Assume the input waveform is peak and avg. constrained and is strictly band-limited to $-B \leq f \leq B$.

- Shamai and Lapidoth ('93) obtained upper and lower bounds on capacity. Showed that non-spike inputs are strictly suboptimal.

Assume the input waveform must be PAM: $\lambda(t) = \sum a_n p(t - nT_s)$, where $p(t)$ is some pulse of duration T_s.

- The capacity of PAM increases as $T_s \rightarrow 0$.
- In the limit $T_s \rightarrow 0$, the optimum distribution of a_s is 2 levels.
- As T_s increases, the optimum distribution of a_s uses more levels.

Feedback

Kabanov and Davis showed that causal, instantaneous, noiseless feedback does not increase capacity of the Poisson channel.

Frey showed that feedback can increase capacity if the dark current is random and time-varying.

Lapidoth has found the exact error exponent under feedback.

Outline

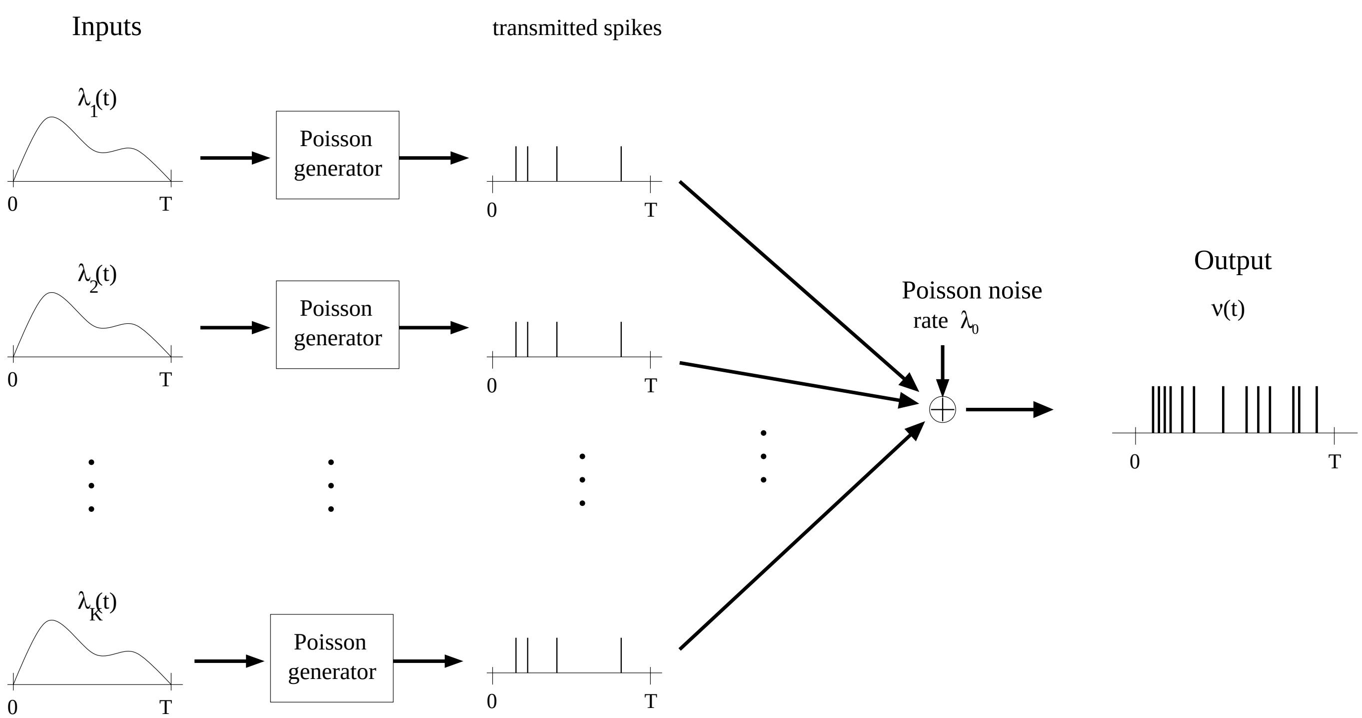
I. Introduction

II. Single-User Poisson Channel

III. Multi-access Poisson Channel

IV. Other Applications

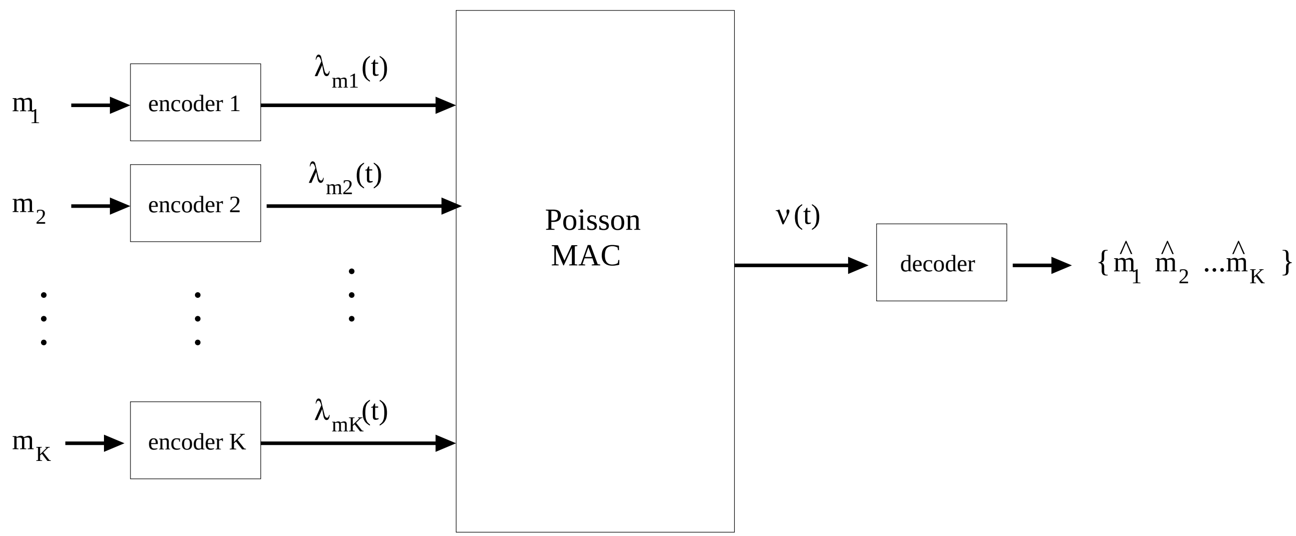
The Multi-access Model



Multi-access Coding and Decoding

The K encoders independently encode their messages.

The decoder tries to decode all messages simultaneously.



Coding and Decoding

Consider the 2-user case:

A (M_1, M_2, T) code consists of:

1. Two index sets: $\{1, \ldots, M_1\}$ and $\{1, \ldots, M_2\}$.
2. Two sets of codewords $\{\lambda_m^{(1)}(t)\}_{m=1}^{M_1}$ and $\{\lambda_m^{(2)}(t)\}_{m=1}^{M_2}$, for $t \in [0, T]$.
3. A decoding function:
   $D(\nu_0^T) = (\hat{m_1}, \hat{m_2}) \in \{1, \ldots, M_1\} \times \{1, \ldots, M_2\}$.

The probability of error is:

$$P_e = \frac{1}{M_1 M_2} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} Pr\{D(\nu_0^T) \neq (m_1, m_2) \mid (\lambda_{m_1}^{(1)}, \lambda_{m_2}^{(2)})\}$$

Capacity Region

A rate pair (R_1, R_2) is said to be *achievable* if $\forall \epsilon > 0$, there exists a (M_1, M_2, T) code with sufficiently large T and $M_i \geq 2^{R_i T}$, $i = 1, 2$ such that $P_e \leq \epsilon$.

The *capacity region* is defined to be the closure of the set of all achievable (R_1, R_2) rate pairs.

Binary Inputs are Optimal

In extending Wyner's results from the single-user channel, Lapidoth and Shamai showed that binary PAM-like inputs do not reduce the capacity region.

Analogous to the single-user case, the Poisson MAC can be simplified to a binary-input, binary-output memoryless MAC.

Capacity Region for 2 Users

Maximum Total Throughput

Maximum total throughput for 2 users:

$$R_{\sum} = \max_{(R_1, R_2) \in \mathcal{C}} (R_1 + R_2)$$

This can be achieved using symmetric rates of the form (R^*, R^*).

For the general case of K users, the maximum total throughput is

- achieved with symmetric rates.
- monotonically increasing in K.
- *bounded above* by the peak amplitude A.

Why does the total throughput *saturate*?

Total Throughput (cont.)

Compare this with the Gaussian MAC: maximum total throughput increases as the *log* of the number of users.

Intuition: look at the *outputs* of the two channels.

- Gaussian MAC: the output is a sum of K indep. Gaussians. As the number of users K increases, the variance of the output increases, hence the output entropy increases.
- Poisson MAC: the output is a sum of K indep. Poisson processes. As K increases, the spacing between output spikes decreases, and this *decreases* the entropy rate. So adding more inputs saturates the output entropy.

Outline

I. Introduction

II. Single-User Poisson Channel

III. Multi-access Poisson Channel

IV. Other Applications

Biological Neuroscience

Our body transmits information through a network of neurons.

The signals look like a train of *spikes* (action potentials).

Neuroscientists believe that information is carried in the *timing* of these spikes, not in their amplitude or shape.

Questions:

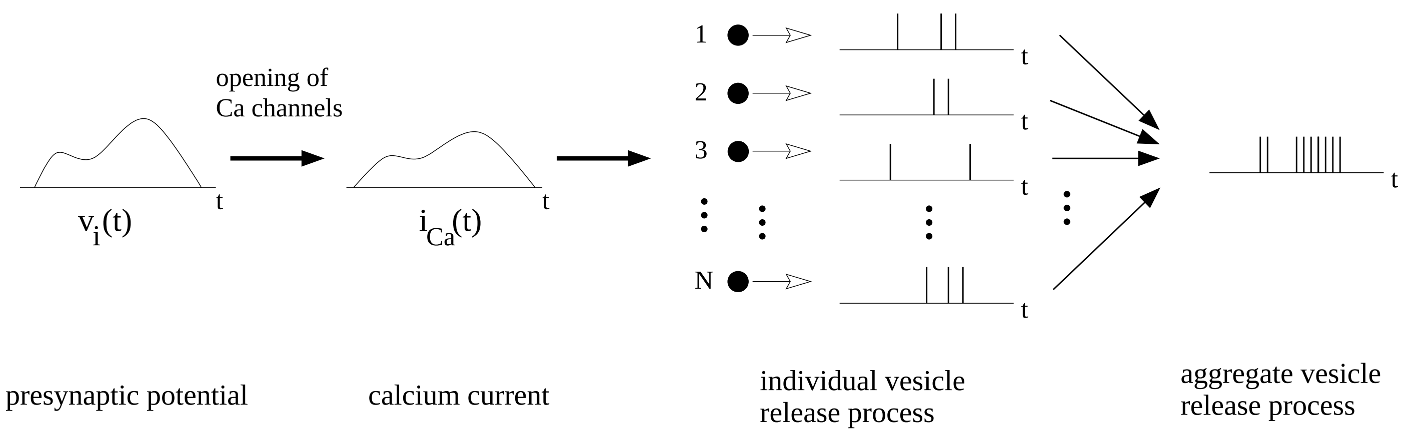
- How does the body “encode” information into these spikes?
- Why does it use spikes?

An Engineering Approach

Try to model the neural system as a noisy communication channel and ask, "How would an engineer build a system for it?"

- Look for a source of noise inherent in the neural system.
- Model the noise and apply Information Theory.
- Compare the theory with reality.

Basic Model of a Synapse



Input: presynaptic potential, $V_i(t)$.

Output: release times of neurotransmitter vesicles, $\nu(t)$.

Assumptions:

- Each vesicle is released according to a Poisson process with time-varying intensity proportional to the input potential $V_i(t)$.
- No refractory period for vesicles (fast replenishment).
- No input bandwidth constraints (fast Ca^{2+} dynamics).

Discussion

- Wyner's result tells us that the best way to send information through such a noisy channel is to use *spike-like* inputs.
- In reality, we know that the body uses spike-like signals.

Is this a coincidence?

Maybe this is how the body has evolved to counteract noise in a synapse?