Signal detection over an optical channel

Given an optical channel, a signal on that channel can be described as a coherent state denoted by $|S\rangle$ where $S \in \mathbb{C}$.

Considering a Noise free environment, in order to detect a signal in this type of channel, we use a photon counter receiver. Such a receiver is a direct detection receiver which detects the intensity of the optical and generates a Poisson process, where the rate of the process holds $\lambda = |S|^2$.

Suppose we have two coherent state signals denoted by $|S_0\rangle$, $|S_1\rangle$ and we would like to distinguish between the two binary hypotheses with the corresponding priori probabilities π_0, π_1 respectively under hypotheses $H = 0, 1$, while holding some transmission cost constraint .

One approach was given by **Kennedy** who proposed adding a constant additional coherent state signal $|l\rangle$ before feeding the signal's input to the receiver. Doing so generates a coherent state $|S + l\rangle$ which the receiver in turn outputs a Poisson process with rate $\lambda_i = |S_i + l|^2$.

An additional approach was given by **Dolinar** who suggested as continuation to Kennedy's design to replace the constant signal with a controlled signal $|l(t)\rangle$ which is chosen adaptively based on the photon arrivals up to that moment, in order to achieve more certainty in the hypothesis choice with time .

The core concept of Dolinar's updated design includes a recursive method in which the posterior probabilities of the two possible hypotheses denoted by $\pi_1(t), \pi_2(t)$ are updated after each step of time Δ to yield $\pi_1(t+\Delta), \pi_2(t+\Delta)$.

At this point, by making Δ arbitrarily small we can expect the current Poisson process to return :

$$\begin{cases} 0 & \text{w.p. } (1 - \lambda_i \Delta) \\ 1 & \text{w.p. } (\lambda_i \Delta) \end{cases}$$

Which can be thought of as the following binary channel (Figure 1) :

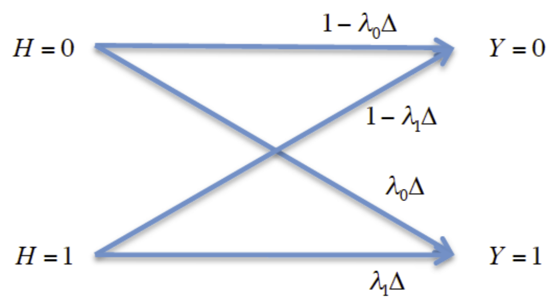

Figure 1: Equivalent binary channel over time Δ

Since we obtained an approximation of a binary channel we may now ask how should $l(t)$ be decided in other to maximize the mutual information of the binary channel .

The Entropy can be calculated as follows :

$$\begin{aligned} H(Y) &= H_b[\Delta(\pi_0\lambda_0 + \pi_1\lambda_1)] \\ H(Y|X) &= \pi_0 H_b(\Delta\lambda_1) + \pi_1 H_b(\Delta\lambda_1) \end{aligned}$$

Which gives us the mutual information of the channel :

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H_b[\Delta(\pi_0\lambda_0 + \pi_1\lambda_1)] \\ &\quad -[\pi_0 H_b(\Delta\lambda_1) + \pi_1 H_b(\Delta\lambda_1)] \end{aligned}$$

Since we want to maximize the mutual information, we shall compare its derivative to zero :

$$\begin{aligned} \frac{dI(X;Y)}{dl} &= \log_2\left[\frac{1-\Delta[\pi_0\lambda_0 + \pi_1\lambda_1]}{\Delta[\pi_0\lambda_0 + \pi_1\lambda_1]}\right] 2\Delta(\pi_0\sqrt{\lambda_0} + \pi_1\sqrt{\lambda_1}) \\ &- \log_2\left[\frac{1-\lambda_0\Delta}{\lambda_0\Delta}\right] 2\Delta\pi_0\sqrt{\lambda_0} \end{aligned}$$

$$
\begin{aligned}
&\quad - \log_2 \left[\frac{1-\lambda_1\Delta}{\lambda_1\Delta}\right] 2\Delta\pi_1\sqrt{\lambda_1} \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{dI(X;Y)}{dl} &= \log_2 \left[\frac{1-\Delta[\pi_0\lambda_0+\pi_1\lambda_1]}{\Delta[\pi_0\lambda_0+\pi_1\lambda_1]}\right] 2\Delta(\pi_0\sqrt{\lambda_0}+\pi_1\sqrt{\lambda_1}) \\
&\quad -2\Delta\left[\log_2\left(\frac{1-\lambda_0\Delta}{\lambda_0\Delta}\right)\pi_0\sqrt{\lambda_0} + \log_2\left(\frac{1-\lambda_1\Delta}{\lambda_1\Delta}\right)\pi_1\sqrt{\lambda_1}\right] \\
&= 0
\end{aligned}
$$

The optimal choice for l is :

$$
\boxed{l^* = \frac{S_0\pi_0 - S_1\pi_1}{\pi_1-\pi_0}}
$$

This optimal solution is numerically proven to be correct. The analytical proof will be included in a future paper .

In the optimal case where $l = l^* = \frac{S_0\pi_0 - S_1\pi_1}{\pi_1-\pi_0}$, we can calculate the following expressions for the λ's :

$$
\lambda_i = (l^*+S_i)^2 = \left(\frac{S_0\pi_0 - S_1\pi_1}{\pi_1-\pi_0} + S_i\right)^2 = \left(\frac{S_0\pi_0 - S_1\pi_1 + S_i(\pi_1-\pi_0)}{\pi_1-\pi_0}\right)
$$

Which yields:

$$
\begin{aligned}
\overline{\lambda} &= \pi_0\lambda_0+\pi_1\lambda_1 \\
&= \pi_0\left(\frac{\pi_1(S_0-S_1)}{\pi_1-\pi_0}\right)^2 + \pi_1\left(\frac{\pi_0(S_0-S_1)}{\pi_1-\pi_0}\right)^2 \\
&= (\pi_0\pi_1^2+\pi_1\pi_0^2)\left(\frac{(S_0-S_1)}{\pi_1-\pi_0}\right)^2 \qquad (1)
\end{aligned}
$$

Since $\pi_0 + \pi_1 = 1$ we have that :

$$\begin{aligned}
\pi_0\pi_1^2 + \pi_1\pi_0^2 &= \pi_0\pi_1(1-\pi_0) + \pi_1\pi_0^2 \\
&= \pi_0\pi_1 - \pi_0\pi_1^2 + \pi_0\pi_1^2 \\
&= \pi_0\pi_1
\end{aligned}$$

Thus we can substitute in (1) and get :

$$(\pi_0\pi_1^2 + \pi_1\pi_0^2)\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2 = \pi_0\pi_1\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2$$

Overall we get the following expression for the λ's:

$$\begin{cases}
\lambda_0 = \left(\dfrac{\pi_1(S_0 - S_1)}{\pi_1 - \pi_0}\right)^2 \\
\lambda_1 = \left(\dfrac{\pi_0(S_0 - S_1)}{\pi_1 - \pi_0}\right)^2 \\
\overline{\lambda} = \pi_0\pi_1\left(\dfrac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2
\end{cases}$$

At this point we can derive the mutual information in the optimal case :

$$\begin{aligned}
I &= H(Y) - H(Y|X) \\
&= H_b[\overline{\lambda}\Delta] - \pi_0 H_b(\lambda_0\Delta) - \pi_1 H_b(\lambda_1\Delta) \\
&= \overline{\lambda}\log\left[\frac{1}{\overline{\lambda}}\right] - \pi_0\lambda_0\log\left[\frac{1}{\lambda_0}\right] - \pi_1\lambda_1\log\left[\frac{1}{\lambda_1}\right]
\end{aligned}$$

$$
\begin{aligned}
= \;& \pi_0\pi_1\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2 \log\left[\frac{1}{\pi_0\pi_1\left(\frac{S_0-S_1}{\pi_1-\pi_0}\right)^2}\right] \\
& -\pi_0\left(\frac{\pi_1(S_0 - S_1)}{\pi_1 - \pi_0}\right)^2 \log\left[\frac{1}{\left(\frac{\pi_1(S_0-S_1)}{\pi_1-\pi_0}\right)^2}\right] \\
& -\pi_1\left(\frac{\pi_0(S_0 - S_1)}{\pi_1 - \pi_0}\right)^2 \log\left[\frac{1}{\left(\frac{\pi_0(S_0-S_1)}{\pi_1-\pi_0}\right)^2}\right] \\
= \;& \pi_0\pi_1\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2 \\
& \Bigg(-\log\left[\pi_0\pi_1\left((\frac{(S_0 - S_1)}{\pi_1 - \pi_0})\right)^2\right] \\
& +\pi_0\log\left[\left(\frac{\pi_1(S_0 - S_1)}{\pi_1 - \pi_0})\right)^2\right] \\
& +\pi_1\log\left[\left(\frac{\pi_0(S_0 - S_1)}{\pi_1 - \pi_0})\right)^2\right]\Bigg) \\
= \;& \pi_0\pi_1\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2 \\
& \Bigg(-\log[\pi_0\pi_1] - \cancel{\log\left[\left((\frac{(S_0 - S_1)}{\pi_1 - \pi_0})\right)^2\right]} \\
& +2\pi_0\log[\pi_0] + \cancel{\pi_0\log\left[\left(\frac{(S_0 - S_1)}{\pi_1 - \pi_0})\right)^2\right]} \\
& +2\pi_1\log[\pi_1] + \cancel{\pi_1\log\left[\left(\frac{(S_0 - S_1)}{\pi_1 - \pi_0})\right)^2\right]}\Bigg) \\
= \;& \pi_0\pi_1\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2 (-log[\pi_0\pi_1] + 2\pi_0\log[\pi_0] + \; 2\pi_1\log[\pi_1]) \\
= \;& \pi_0\pi_1\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2 (-log[\pi_0] - log[\pi_1] + 2\pi_0\log[\pi_0] + \; 2\pi_1\log[\pi_1])
\end{aligned}
$$

$$= \pi_0\pi_1\left(\frac{S_0 - S_1}{\pi_1 - \pi_0}\right)^2((2\pi_0 - 1)log[\pi_0] + (2\pi_1 - 1)log[\pi_1]) \tag{2}$$

Again, since $\pi_0 + \pi_1 = 1$ we have :

$$\begin{cases} 2\pi_0 - 1 = \pi_0 + \pi_0 - 1 = \pi_0 - \pi_1 \\ 2\pi_1 - 1 = \pi_1 + \pi_1 - 1 = \pi_1 - \pi_0 \end{cases}$$

So we can substitute in (2) and get :

$$\begin{aligned} &= \pi_0\pi_1\frac{(S_0 - S_1)^2}{(\pi_1 - \pi_0)^{\cancel{2}})}\cancel{((\pi_0 - \pi_1)}(log[\pi_0] - log[\pi_1])) \\ &= -\pi_0\pi_1\left(\frac{(S_0 - S_1)^2}{\pi_1 - \pi_0}\right)log\left[\frac{\pi_0}{\pi_1}\right] \\ &= \pi_0\pi_1\left(\frac{(S_0 - S_1)^2}{\pi_0 - \pi_1}\right)log\left[\frac{\pi_0}{\pi_1}\right] \end{aligned} \tag{3}$$

At this point we shall parameterize our probabilities with parameter $g = \frac{\pi_0}{\pi_1}$ which will let us describe the recursive procedure in the system. That definition yields the parametrization :

$$g = \frac{\pi_0}{\pi_1} \Longrightarrow \begin{cases} \pi_0 = \dfrac{g}{1+g} \\ \pi_1 = \dfrac{1}{1+g} \end{cases}$$

With that parameterization, we can now continue developing (3) :

$$\begin{aligned} \pi_0\pi_1\left(\frac{(S_0 - S_1)^2}{\pi_0 - \pi_1}\right)log\left[\frac{\pi_0}{\pi_1}\right] &= (\frac{g}{1+g})(\frac{1}{1+g})\frac{(S_0 - S_1)^2}{\frac{g}{1+g} - \frac{1}{1+g}}log\left(\frac{\frac{g}{1+g}}{\frac{1}{1+g}}\right) \\ &= \left(\frac{g}{(1+g)^{\cancel{2}}}\right)\left(\frac{(S_0 - S_1)^2}{\frac{g-1}{\cancel{1+g}}}\right)\log(g) \\ &= \frac{g(S_0 - S_1)^2}{(1-g)(1+g)}\log(g) \end{aligned}$$

We can also recompute the entropy using that parameterization :

$$
\begin{aligned}
H(g) &= H_b(\pi_0, \pi_1) \\
&= \pi_0 \log\left[\frac{1}{\pi_0}\right] + \pi_1 \log\left[\frac{1}{\pi_1}\right] \\
&= \frac{g}{1+g} \log\left[\frac{1}{\frac{g}{1+g}}\right] + \frac{1}{1+g} \log\left[\frac{1}{\frac{1}{1+g}}\right] \\
&= -\frac{1}{1+g}\left(g \log \frac{g}{1+g} + \log \frac{1}{1+g}\right) \\
&= -\frac{1}{1+g}\left(g \log g + (1+g) \log \frac{1}{1+g}\right) \\
&= \frac{1}{1+g}((1+g)\log(1+g) - g \log g) \\
&= \log(1+g) - \frac{g}{1+g}\log g
\end{aligned}
$$

So we can now find the derivative of the entropy as a function of g :

$$
\begin{aligned}
\frac{H(g)}{dg} &= \frac{d(\log(1+g) - \frac{g}{1+g}\log g)}{dg} \\
&= \frac{1}{1+g} - \frac{(1+\cancel{g}) - \cancel{g}}{(1+g)^2}\log g - \frac{1}{\cancel{g}}\frac{\cancel{g}}{(1+g)} \\
&= \cancel{\frac{1}{1+g}} - \frac{\log g}{(1+g)^2} - \cancel{\frac{1}{1+g}} \\
&= -\frac{\log g}{(1+g)^2}
\end{aligned}
$$

Each step of the recursion, the entropy is lowered since we have more certainty of our choice. The quantity of the change in certainty can be describe by the following equation :

$$
H(g(t+\Delta)) = H(g(t)) - \Delta I(g(t))
$$

If we look again at the the derivative of the entropy, we see than indeed it is negative which coincides with the fact that the certainty is increasing .

The definition of the derivative entropy by g, is defined by :

$$\frac{H(g)}{dg} = \frac{H(g(t+\Delta)) - H(g(t))}{g(t+\Delta) - g(t)}$$

which yields :

$$\implies g(t+\Delta) = g(t) + \frac{-\Delta I}{\frac{dH}{dg}} \tag{4}$$

Since we found already the derivative and the mutual information, we can substitute them in (4) :

$$\begin{aligned} g(t+\Delta) &= g(t) + \frac{-\Delta I}{\frac{dH}{dg}} \\ &= g(t) + \frac{-\Delta \frac{g(S_0 - S_1)^2}{(1-g)\cancel{(1+g)}} \cancel{\log(g)}}{-\frac{\cancel{\log(g)}}{(1+g)^{\cancel{2}}}} \\ &= g(t) + \left(\frac{\Delta g(1+g)}{(1-g)} \right) (S_0 - S_1)^2 \end{aligned}$$

We assume that $S_0(t)$ and $S_1(t)$ are constant in time and therefore we shall omit the dependency on t :

$$\begin{aligned} g(t) &= g(0) \cdot \exp\left[\int_0^{\mathrm{t}} \left(\frac{(\mathrm{S}_0 - \mathrm{S}_1)^2 (\mathrm{g}(\tau) + 1)}{\mathrm{g}(\tau) - 1} \right) \mathrm{d}\tau \right] \\ &= g(0) \cdot \prod_{\tau=0}^{t} \left[\frac{(S_0 - S_1)^2 (g(\tau) + 1)}{g(\tau) - 1} \right] \end{aligned}$$

We can now look at the change ratio :

$$
\begin{aligned}
\frac{g(t+\Delta)}{g(t)} &= \frac{g(0)\cdot\prod_{\tau=0}^{t+\Delta}\exp\left[\frac{(\mathrm{S}_0-\mathrm{S}_1)^2(\mathrm{g}(\tau)+1)}{\mathrm{g}(\tau)-1}\right]}{g(0)\cdot\prod_{\tau=0}^{t}\exp\left[\frac{(\mathrm{S}_0-\mathrm{S}_1)^2(\mathrm{g}(\tau)+1)}{\mathrm{g}(\tau)-1}\right]} \\
&= \exp\left[\frac{(\mathrm{S}_0-\mathrm{S}_1)^2(\mathrm{g}(\tau+\Delta)+1)}{\mathrm{g}(\tau+\Delta)-1}\right]
\end{aligned}
$$

Which gives :

$$
\log\left[\frac{g(t+\Delta)}{g(t)}\right] = \frac{(S_0-S_1)^2(g(\tau+\Delta)+1)}{g(\tau+\Delta)-1} \tag{5}
$$

Since by Tailor approximations for $x << 1$ it holds that $\log(1+x) \simeq x - \frac{1}{2}x^2$ we have that :

$$
\begin{aligned}
\log\left[\frac{g(t+\Delta)}{g(t)}\right] &= \log\left[1+\frac{g(t+\Delta)-g(t)}{g(t)}\right] \\
&\simeq \frac{g(t+\Delta)-g(t)}{g(t)}
\end{aligned}
$$

Substituting that in (5) gives :

$$
\frac{g(t+\Delta)-g(t)}{g(t)} \simeq \frac{(S_0-S_1)^2(g(\tau+\Delta)+1)}{g(\tau+\Delta)-1}
$$

Which is what we got to.