How to derive the channel capacity of the binary-symmetric channel (BSC) with crossover probability ϵ?

Solution.

Denoting $p(x=0)=u$ and $p(x=1)=1-u$, we obtain

$$\begin{aligned}
p(y) &= \sum_{x'=0}^{1} p(x')p(y|x') \\
&= p(x'=0)p(y|x'=0)+p(x'=1)p(y|x'=1) \\
&= \begin{cases} p(x'=0)p(y=0|x'=0)+p(x'=1)p(y=0|x'=1), & y=0 \\ p(x'=0)p(y=1|x'=0)+p(x'=1)p(y=1|x'=1), & y=1 \end{cases} \\
&= \begin{cases} u(1-\epsilon)+(1-u)\epsilon, & y=0 \\ u\epsilon+(1-u)(1-\epsilon), & y=1 \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
H(Y|X) &= \sum_{x=0}^{1}\sum_{y=0}^{1} p(x)p(y|x)\log_2\frac{1}{p(y|x)} \\
&\left(= \sum_{x=0}^{1}\sum_{y=0}^{1} p(x)\cdot H(Y|X=x)\right) \\
&= p(x=0)p(y=0|x=0)\log_2\frac{1}{p(y=0|x=0)} \\
&\quad +p(x=0)p(y=1|x=0)\log_2\frac{1}{p(y=1|x=0)} \\
&\quad +p(x=1)p(y=0|x=1)\log_2\frac{1}{p(y=0|x=1)} \\
&\quad +p(x=1)p(y=1|x=1)\log_2\frac{1}{p(y=1|x=1)} \\
&= u(1-\epsilon)\log_2\frac{1}{(1-\epsilon)}+u\epsilon\log_2\frac{1}{\epsilon} \\
&\quad +(1-u)\epsilon\log_2\frac{1}{\epsilon}+(1-u)(1-\epsilon)\log_2\frac{1}{(1-\epsilon)} \\
&= \epsilon\log_2\frac{1}{\epsilon}+(1-\epsilon)\log_2\frac{1}{(1-\epsilon)} \\
&= H_b(\epsilon) \qquad \text{(This is called the binary entropy function.)}
\end{aligned}$$

Hence,

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= [u(1-\epsilon) + (1-u)\epsilon] \log_2 \frac{1}{u\epsilon + (1-u)(1-\epsilon)} \\ &\quad + [u\epsilon + (1-u)(1-\epsilon)] \log_2 \frac{1}{u\epsilon + (1-u)(1-\epsilon)} - H_b(\epsilon), \end{aligned}$$

and

$$\begin{aligned} C &= \max_{p(x)=(u,1-u)} I(X;Y) \\ &= \max_{0\le u\le 1} \Bigg([u(1-\epsilon) + (1-u)\epsilon] \log_2 \frac{1}{u\epsilon + (1-u)(1-\epsilon)} \\ &\quad + [u\epsilon + (1-u)(1-\epsilon)] \log_2 \frac{1}{u\epsilon + (1-u)(1-\epsilon)} - H_b(\epsilon) \Bigg) \\ &= -\min_{0\le u\le 1} \Bigg([u(1-\epsilon) + (1-u)\epsilon] \log_2 [u\epsilon + (1-u)(1-\epsilon)] \\ &\quad [u\epsilon + (1-u)(1-\epsilon)] \log_2 [u\epsilon + (1-u)(1-\epsilon)] + H_b(\epsilon) \Bigg). \end{aligned}$$

Taking the derivative of the term inside parentheses with respect to u yields:

$$\begin{aligned} &\Bigg([u(1-\epsilon) + (1-u)\epsilon] \log_2 [u(1-\epsilon) + (1-u)\epsilon] \\ &+ [u\epsilon + (1-u)(1-\epsilon)] \log_2 [u\epsilon + (1-u)(1-\epsilon)] + H_b(\epsilon) \Bigg)' \\ = \;& (1-2\epsilon) \log_2 [u(1-\epsilon) + (1-u)\epsilon] + [u(1-\epsilon) + (1-u)\epsilon] \frac{1-2\epsilon}{\log(2)[u(1-\epsilon) + (1-u)\epsilon]} \\ &+ (2\epsilon - 1) \log_2 [u\epsilon + (1-u)(1-\epsilon)] + [u\epsilon + (1-u)(1-\epsilon)] \frac{2\epsilon - 1}{\log(2)[u\epsilon + (1-u)(1-\epsilon)]} \\ = \;& (1-2\epsilon) \Bigg(\log_2 [u(1-\epsilon) + (1-u)\epsilon] - \log_2 [u\epsilon + (1-u)(1-\epsilon)] \Bigg). \end{aligned}$$

As a result, the above derivative equals zero if, and only if,

$$u(1-\epsilon) + (1-u)\epsilon = u\epsilon + (1-u)(1-\epsilon)$$

which gives $u^* = 1/2$. Finally,

$$\begin{aligned} C &= -\Big([u^*(1-\epsilon) + (1-u^*)\epsilon] \log_2 [u^*\epsilon + (1-u^*)(1-\epsilon)] \\ &\quad [u^*\epsilon + (1-u^*)(1-\epsilon)] \log_2 [u^*\epsilon + (1-u^*)(1-\epsilon)] + H_b(\epsilon) \Big) \\ &= 1 - H_b(\epsilon). \end{aligned}$$