Information theory and coding
Take home exam

Solignac Robin 235020

November 28, 2016

Ex 1

(a)

We have $\tilde{S^n}$ independent of S^n and Y^n but with the same marginal $p(s)$. so we have

$$\Pr\left((\tilde{S^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S,Y)\right) = \sum_{(s^n,y^n)\in\mathcal{A}_\epsilon^{(n)}(S,Y)} p(s^n)p(y^n)$$
$$\Pr\left((\tilde{S^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S,Y)\right) \leq \sum_{(s^n,y^n)\in\mathcal{A}_\epsilon^{(n)}(S,Y)} 2^{-n(H(S)-\epsilon)}2^{-n(H(Y)-\epsilon)}$$
$$\Pr\left((\tilde{S^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S,Y)\right) \leq 2^{n(H(S,Y)+\epsilon)}2^{-n(H(S)-\epsilon)}2^{-n(H(Y)-\epsilon)}$$
$$\Pr\left((\tilde{S^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S,Y)\right) \leq 2^{-n(I(S,Y)-3\epsilon)} \tag{1}$$

(b)

We have $\tilde{X^n}$ independent from (S^n, X^n, Y^n) but it has the same marginal as X^n. so in particular $p(s^n, \tilde{x^n}, y^n) = p(\tilde{x^n})p(s^n, y^n)$. And so

$$\Pr\left((S^n, \tilde{X^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S^n,X^n,Y^n)\right) = \sum_{(s^n,x^n,y^n)\in\mathcal{A}_\epsilon^{(n)}(S,X,Y)} p(x^n)p(s^n,y^n)$$
$$\Pr\left((S^n, \tilde{X^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S^n,X^n,Y^n)\right) \leq \sum_{(s^n,x^n,y^n)\in\mathcal{A}_\epsilon^{(n)}(S,X,Y)} 2^{-n(H(X)-\epsilon)}2^{-n(H(Y,S)-\epsilon)}$$
$$\Pr\left((S^n, \tilde{X^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S^n,X^n,Y^n)\right) \leq 2^{n(H(X,Y,S)-\epsilon)}2^{-n(H(X)-\epsilon)}2^{-n(H(Y,S)-\epsilon)}$$
$$\Pr\left((S^n, \tilde{X^n}, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S^n,X^n,Y^n)\right) \leq 2^{-n(I(X;Y,S)-3\epsilon)} \tag{2}$$

Note for later: One thing we can note here is that this inequality does not depend on what the distribution of $p(s^n, y^n)$ and $p(x^n)$ it just rely on the fact that $\tilde{X^n}$ is independent of S^n and Y^n and that it has the same marginal as X^n. But if we switch and take $\tilde{Y^n}$ independent of S^n, X^n (with same marginal as Y^n) then we get

$$\Pr\left((S^n, X^n, \tilde{Y^n}) \in \mathcal{A}_\epsilon^{(n)}(S, X, Y)\right) \leq 2^{-n(I(X;Y,S)-3\epsilon)} \tag{3}$$

(c)

First, if $(s^n, y^n) \in \mathcal{A}_\epsilon^{(n)}(S, Y)$ then we have $p(s^n, y^n) \leq 2^{-n(H(S,Y)-\epsilon)}$ and $p(s^n) \geq 2^{-n(H(S)+\epsilon)}$ so $p(y^n|s^n) = \frac{p(s^n,y^n)}{p(s^n)} \leq \frac{2^{-n(H(S,Y)-\epsilon)}}{2^{-n(H(S)+\epsilon)}} = 2^{-n(H(Y|S)-2\epsilon)}$

By replacing Y by X we get also $p(x^n|s^n) \leq 2^{-n(H(X|S)-2\epsilon)}$

Now we have $\tilde{X^n}$ which is also compute from S^n but independent from the computation of X^n from S^n and so independent of the generation of Y^n form S^n by the intermediate of X^n. This mean $p(s^n, \tilde{x^n}, y^n) = p(\tilde{x^n}, y^n|s)p(s) = p(\tilde{x^n}|s)p(y^n|s)p(s)$

so we get

$$\begin{aligned}
\Pr\left((S^n, \tilde{X}^n, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S^n, X^n, Y^n)\right) &= \sum_{(s^n,x^n,y^n)\in\mathcal{A}_\epsilon^{(n)}(S,X,Y)} p(x^n|s^n)p(y^n|s^n)p(s^n) \\
&\leq 2^{n(H(X,Y,S)-\epsilon)}2^{-n(H(X|S)-2\epsilon)}2^{-n(H(Y|S)-2\epsilon)}2^{-n(H(S)-\epsilon)} \\
&\leq 2^{n(H(X|S)+H(Y|S)-H(X,Y,S)-6\epsilon)}
\end{aligned}$$

$$\Pr\left((S^n, \tilde{X}^n, Y^n) \in \mathcal{A}_\epsilon^{(n)}(S^n, X^n, Y^n)\right) \leq 2^{n(I(X;Y|S)-6\epsilon)} \tag{4}$$

Ex 2

First let's define $t_{i,j}$ the event: "the message pair $(m_1 = i, m_2 = j)$ was transmitted"

(a)

We have

$$\Pr(\varepsilon) = \mathbb{E}\left[\frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\lambda_{m_1,m_2}(C)\right]$$
$$\Pr(\varepsilon) = \sum_C \Pr(C)\frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\lambda_{m_1,m_2}(C)$$
$$\Pr(\varepsilon) = \sum_C \frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\Pr(C)\lambda_{m_1,m_2}(C)$$

But we can use the symmetry of our codebook which implies that λ_{m_1,m_2} does not depend of indexes m_1 and m_2. so $\lambda_{m_1,m_2} = \lambda_{1,1}\,\forall m_1, m_2$

Thus:

$$\Pr(\varepsilon) = \sum_C \frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\Pr(C)\lambda_{1,1}(C)$$
$$\Pr(\varepsilon) = \sum_C \frac{1}{2^{nR_1}}2^{nR_1}\frac{1}{2^{nR_2}}2^{nR_2}\Pr(C)\lambda_{1,1}(C)$$
$$\Pr(\varepsilon) = \sum_C \Pr(C)\lambda_{1,1}(C)$$
$$\Pr(\varepsilon) = \Pr\left(\varepsilon|W_1 = 1, W_2 = 1\right)$$
$$\Pr(\varepsilon) = \mathbb{E}\left[\lambda_{1,1}(C)|W_1 = 1, W_2 = 2\right]$$

(b)

We have

$$\lambda_{1,1} = \Pr\left(\varepsilon_{1,(1,1)} \cup \varepsilon_{2,1}|t_{1,1}\right)$$
$$\lambda_{1,1} \leq \Pr\left(\varepsilon_{1,(m_1,m_2)}|t_{1,1}\right) + \Pr\left(\varepsilon_{2,m_2}|t_{1,1}\right)$$
$$\lambda_{1,1} \leq \lambda_{1,(1,1)} + \lambda_{2,1}$$

So:

$$\Pr(\varepsilon) \leq \sum_C \Pr(C)\left(\lambda_{1,(1,1)} + \lambda_{2,1}\right)$$
$$\Pr(\varepsilon) \leq \left[\lambda_{1,(1,1)}\right] + \left[\lambda_{2,1}\right]$$
$$\Pr(\varepsilon) \leq \left[\lambda_{1,(1,1)}|t_{1,1}\right] + \left[\lambda_{2,1}|t_{1,1}\right]$$

(c)

$$[\lambda_{2,1}|t_{1,1}] = \Pr\left(\varepsilon_{2,1}|t_{1,1}\right)$$

Let's define $E_{2,i}$ the event $s^n(i)$ and Y_2 are jointly typical. where Y_2^n is the message receive by D_1. so $E_{2,i} = (s^n(i), Y_2^n) \in \mathcal{A}_\epsilon^{(n)}(S, Y_2)$

so $\varepsilon_{2,1}|t_{1,1} = E_{2,1}^c \cup \bigcup_{i=2}^{2^{nR_2}} E_{2,i}|t_{1,1}$ and so by using the union bound

$$[\lambda_{2,1}|t_{1,1}] = \Pr\left(E_{2,1}^c \cup \bigcup_{i=2}^{2^{nR_2}} E_{2,i}|t_{1,1}\right)$$

$$[\lambda_{2,1}|t_{1,1}] \leq \Pr(E_{2,1}^c|t_{1,1}) + \sum_{i=2}^{2^{nR_2}} \Pr\left(E_{2,i}|t_{1,1}\right)$$

by join AEP properties: $\Pr(E_{2,1}^c|t_{1,1}) \leq \epsilon$ for n large enough.

from our codebook generation process $s^n(i)$ is independent from $s^n(1)$ for $i \neq 1$ and so we can apply inequality (1): and get $E_i \leq 2^{-n(I(S,Y)-\epsilon)}$

$$[\lambda_{2,1}|t_{1,1}] \leq \epsilon + \sum_{i=2}^{2^{nR_2}} 2^{-n(I(S,Y_2)-3\epsilon)}$$

$$[\lambda_{2,1}|t_{1,1}] \leq \epsilon + 2^{nR_2} * 2^{-n(I(S,Y_2)-3\epsilon)}$$

$$[\lambda_{2,1}|t_{1,1}] \leq \epsilon + 2^{-n(I(S,Y_2)-R_2-3\epsilon)}$$

and this upper bound converge to ϵ with $n \to \infty$ if $R_2 < I(S, Y_2)$.

so with $R_2 < I(S, Y_2)$ we can make $[\lambda_{2,1}|t_{1,1}]$ as small as we want by choosing ϵ and a n large enough. It doesn't depend on R_1.

(d)

$$\left[\lambda_{1,(1,1)}|t_{1,1}\right] = \Pr\left(\varepsilon_{1,(1,1)}|t_{1,1}\right)$$

Let's define $E_{1,(\mathrm{i,j})}$ the event $(s^n(j), x^n(i,j), Y_1^n) \in \mathcal{A}_\epsilon^{(n)}(S, X, Y_1)$ where Y_1^n is the message receive from the channel by D_1, and $E_{1,j}$ the event

$(s^n(j), Y_1^n) \in \mathcal{A}_\epsilon^{(n)}(S, Y_1)$. That's mean:

$$\varepsilon_{1,(1,1)}|t_{1,1} = E^c_{1(1,1)} \cup \bigcup_{i,j\neq 1,1} E_{1,(i,j)}|t_{1,1}$$

$$\varepsilon_{1,(1,1)}|t_{1,1} = E^c_{1(1,1)} \cup \bigcup_{i=2}^{2^{nR_1}} E_{1,(i,1)} \cup \bigcup_{j=2}^{2^{nR_2}} E_{j,(1,j)} \cup \bigcup_{j=2}^{2^{nR_2}} \left(\bigcup_{i=2}^{2^{nR_1}} E_{1,(i,j)} \right) |t_{1,1}$$

$$\Pr\left(\varepsilon_{1,(1,1)}|t_{1,1}\right) \leq \Pr\left(E^c_{1(1,1)}|t_{1,1}\right) + \sum_{i=2}^{2^{nR_1}} \Pr\left(E_{1,(i,1)}|t_{1,1}\right)$$
$$+ \sum_{j=2}^{2^{nR_2}} \Pr\left(E_{j,(1,j)}|t_{1,1}\right) + \sum_{j=2}^{2^{nR_2}} \sum_{i=2}^{2^{nR_1}} \Pr\left(E_{1,(i,j)}|t_{1,1}\right)$$

$$\Pr\left(\varepsilon_{1,(1,1)}|t_{1,1}\right) \leq \Pr\left(E^c_{1(1,1)}|t_{1,1}\right) + \sum_{i=2}^{2^{nR_1}} \Pr\left(E_{1,(i,1)}|t_{1,1}\right) + \sum_{j=2}^{2^{nR_2}} \sum_{i=1}^{2^{nR_1}} \Pr\left(E_{1,(i,j)}|t_{1,1}\right)$$

on this sum we can make the following observation

- By join AEP properties: $\Pr(E^c_{1(1,1)}|t_{1,1}) \leq \epsilon$ for n large enough.
- In $E_{1,(i,1)} = s^n(1), x^n(i,1), Y_1^n$ we have that $x^n(i,1)$ and $Y_1^n$is compute from $s^n(1)$ but independently from each other. this is equivalent to the case covered in Ex 1.(c) and so we can use inequality (4)
- in $E_{1,(i,j)} = s^n(j), x^n(i,j), Y_1^n$ we have $x^n(i,j)$ compute from $s^n(j)$ but Y_1^n is completely independent of both. it's the case covered in Ex 1.(b) and so we can use the inequality (3)

So we have

$$\Pr\left(\varepsilon_{1,(1,1)}|t_{1,1}\right) \leq \epsilon + \sum_{i=2}^{2^{nR_1}} 2^{-n(I(X;Y_1|S)-6\epsilon)} + \sum_{j=2}^{2^{nR_2}} \sum_{i=1}^{2^{nR_1}} 2^{-n(I(X,S;Y_1-3\epsilon)}$$
$$\Pr\left(\varepsilon_{1,(1,1)}|t_{1,1}\right) \leq \epsilon + \sum_{i=1}^{2^{nR_1}} 2^{-n(I(X;Y_1|S)-6\epsilon)} + \sum_{j=1}^{2^{nR_2}} \sum_{i=1}^{2^{nR_1}} 2^{-n(I(X,S;Y_1)-3\epsilon)}$$
$$\Pr\left(\varepsilon_{1,(1,1)}|t_{1,1}\right) \leq \epsilon + 2^{-n(I(X;Y_1|S)-R_1-6\epsilon)} + 2^{-n(I(X,S;Y_1)-R_1-R_2-3\epsilon)}$$

and we can make this upper bound as small as we want the right chose of ϵ and a n large enough if we have both:

- $R_1 < I(X;Y_1|S)$
- $R_1 + R_2 < I(X,S;Y_1)$

Ex 3

Definitions and result for later

We define $h_b(a)$ the binary entropy with probability a. In other word: $H(X)$ with $X \sim \text{Bernoulli}\,(a)$

also we assume that for every symmetric channel with binary flip probability a then $a \leq 0.5$ because otherwise we just have to flip every bit at the output of the flip and get a smaller flip probability $1 - a \leq 0.5$ and so getting a better channel with 0 effort.

we can also note that if we have $a \leq b \leq 0.5$ then $h_b(a) \leq h(b)$ because binary entropy is an increasing function over $[0; 0.5]$.

if we define

$$f_k(a) = a(1 - k) + (1 - a)k$$

for a fixed $0 \leq k \leq 0.5$ and $0 \leq a \leq 0.5$

then $f_k(a) = a(1-2k)+k$ and so $(1-2k) \geq 0$ and so f_k is an increasing function over $0 \leq a \leq 0.5$.

That mean for all $0 \leq a \leq b \leq 0.5$:

$$f_k(a) \leq f_k(b) \leq f_k(0.5) = 0.5$$

$$h_b(f_k(a)) \leq h_b(f_k(b)) \leq 1 \tag{5}$$

Also, here we assume that:

$p(x)$ and $p(y)$ are binary, and $p(s)$ is binary and uniform.

So $p_x(0) = p_s(0)p_{x|s}(0|0) + p_s(1)p_{x|s}(0|1) = \frac{1}{2}(1 - \alpha) + \frac{1}{2}\alpha = \frac{1}{2}$ so marginal $p(x)$ is also uniform.

it's identical for $p_{y1}(0) = p_x(0)p_{y_1|x}(0|0)+p_x(1)p_{y_1|x}(0|1) = \frac{1}{2}(1-q_1)+\frac{1}{2}q_1 = \frac{1}{2}$.

And so on for $p(y_2)$. At the end we can sat that: $p(s)$,$p(x)$,$p(y_1)$,$p(y_2)$ are all binary uniform

Finally as we have $p(x, y, z) = p(s)p(x|s)p(y|x)$.

Then $S \to X \to Y$ is a Markov chain, so $p(y|s, x) = p(y|x)$ and so $H(Y|S, X) = H(Y|X)$

(a)

Recap:

- $R_1 + R_2 < I(X, S; Y_1)$
- $R_1 < I(X; Y_1|S)$
- $R_2 < I(S, Y_2)$

$$\begin{aligned}
&R_1 + R_2 < I(X, S; Y_1)\\
&R_1 + R_2 < H(Y) - H(Y_1|S, X)\\
&R_1 + R_2 < H(Y) - H(Y_1|X)\\
&R_1 + R_2 < 1 - h_b(q_1)
\end{aligned} \tag{6}$$

$$\begin{aligned}
&R_1 < I(X; Y_1|S)\\
&R_1 < H(Y_1|S) - H(Y_1|X, S)\\
&R_1 < H(Y_1|S) - H(Y_1|X, S)\\
&R_1 < H(Y_1|S) - H(Y_1|X)\\
&R_1 < H(Y_1|S) - h_b(q_1)\\
&R_1 < h_b\left(\alpha(1 - q_1) + (1 - \alpha)q_1\right) - h_b(q_1)
\end{aligned} \tag{7}$$

$$\begin{aligned}
&R_2 < I(S, Y_2)\\
&R_2 < H(Y) - H(Y_2|S)\\
&R_2 < 1 - H(Y_2|S)\\
&R_2 < 1 - h_b\left(\alpha(1 - q_2) + (1 - \alpha)q_2\right)
\end{aligned} \tag{8}$$

So in general we have that the region of all possible rate (R_1, R_2) is the region such that

$$\begin{cases}
R_1 < h_b\left(\alpha(1 - q_1) + (1 - \alpha)q_1\right) - h_b(q_1) & (i)\\
R_2 < 1 - h_b\left(\alpha(1 - q_2) + (1 - \alpha)q_2\right) & (ii)\\
R_1 + R_2 < 1 - h_b(q_1) & (iii)
\end{cases} \tag{9}$$

but here has $q_1 < q_2$ from (5) and (ii)

$$R_2 < 1 - h_b\left(\alpha(1 - q_2) + (1 - \alpha)q_2\right) < 1 - h_b\left(\alpha(1 - q_1) + (1 - \alpha)q1\right)$$

and this mean that with addition of (i)

$$R_1 + R_2 < 1 - h_b(q_1)$$

so in this case (i) and (ii) implies (iii) and so we can simplify our region definition:

if $q_1 < q_2$:

$$\begin{cases} R_1 < h_b\left(\alpha(1-q_1) + (1-\alpha)q_1\right) - h_b(q_1) \\ R_2 < 1 - h_b\left(\alpha(1-q_2) + (1-\alpha)q_2\right) \end{cases}$$

(b)

We can start back from (9).

So we have that

$$\begin{cases} R_1 < h_b\left(\alpha(1-q_1) + (1-\alpha)q_1\right) - h_b(q_1) \\ R_2 < 1 - h_b\left(\alpha(1-q_2) + (1-\alpha)q_2\right) \\ R_1 + R_2 < 1 - h_b(q_1) \end{cases}$$

But now as $q_2 < q_1$ then it's not always true that $R_2 < 1 - h_b\left(\alpha(1-q_1) + (1-\alpha)q1\right)$ and so (i) and (ii) doesn't always imply (iii)

Now. If we want to make R_2 higher than $1 - h_b\left(\alpha(1-q_1) + (1-\alpha)q_1\right)$ (so higer than it's the maximum bound in the previous case). this will reduce the bound on R_1.

Indeed, if

$$1 - h_b\left(\alpha(1-q_2) + (1-\alpha)q_2\right) > R_2 = 1 - h_b\left(\alpha(1-q_2) + (1-\alpha)q_3\right) > 1 - h_b\left(\alpha(1-q_1) + (1-\alpha)q_1\right)$$

(so if we take $q_2 < q_3 < q_1$ (5))

then in order to satisfy $R_1 + R_2 < 1 - h_b(q_1)$ we will need

$$\begin{aligned} R_1 + 1 - h_b\left(\alpha(1-q_2) + (1-\alpha)q_3\right) &< 1 - h_b(q_1) \\ R_1 &< h_b\left(\alpha(1-q_2) + (1-\alpha)q_3\right) - h_b(q_1) \end{aligned}$$

and as $q_3 < q_1$ this is a smaller bound than (i).

But we still have $R_1 < h_b\left(\alpha(1-q_1) + (1-\alpha)q_1\right) - h_b(q_1)$ so taking $q_2 < q_1$ instead of $q_1 < q_2$ doesn't lead to any possibility of increasing the bound of R_1 only the one of R_2 at some cost one the one of R_1. So making the channel Y_2 better than Y_1instead of the reverse can let us have a better rate R_2 but it will imply to make the rate R_1 to have a smaller maximum value than before, and will not help us in any way to improve R_1.

Ex 4

(a)

$$\Pr(\varepsilon) = \mathbb{E}\left[\frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\lambda_{m_1,m_2}(C)\right]$$
$$\Pr(\varepsilon) = \sum_C \Pr(C)\frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\lambda_{m_1,m_2}(C)$$
$$\Pr(\varepsilon) = \sum_C \frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\Pr(C)\lambda_{m_1,m_2}(C)$$

But as we draw $u_1^n(m_1, m_t)$ and $u_2^n(m_2, m_s)$ independently from independently from the index m_1, m_2, m_s, m_t. so it's the same for v_i^n and x^n so $\lambda_{m_1,,m_2}$ is independent of m_1 and m_2 so $\lambda_{m_1,,m_2} = \lambda_{1,1}$

$$\Pr(\varepsilon) = \sum_C \frac{1}{2^{nR_1}}\sum_{m_1=1}^{2^{nR_1}}\frac{1}{2^{nR_2}}\sum_{m_2=1}^{2^{nR_2}}\Pr(C)\lambda_{1,1}(C)$$
$$\Pr(\varepsilon) = \sum_C \frac{1}{2^{nR_1}}2^{nR_1}\frac{1}{2^{nR_2}}2^{nR_2}\Pr(C)\lambda_{1,1}(C)$$
$$\Pr(\varepsilon) = \sum_C \Pr(C)\lambda_{1,1}(C)$$
$$\Pr(\varepsilon) = \Pr\left(\varepsilon | W_1 = 1, W_2 = 1\right)$$

(b)

if there's no (m_t, m_s) such that $(u_1^n(1, m_t), u_2^n(1, m_s))$ then the pair $(u_1^n(1, m_t^*), u_2^n(1, m_s^*))$ receive will have one of those property:

- $u_1^n(1, m_t)$ is not typical in U_1^n. so $u_1^n(1, m_t), y_1^n$ will not be typical for any y_1^n even if $m_1 = 1$was send. This will lead to an error a decoding in receiver 1 when $m_1 = 1$ is transmitted
- $u_2^n(1, m_s)$ is not typical in U_2^n. so $u_2^n(1, m_s), y_2^n$ will not be typical for any y_2^n even if $m_2 = 1$was send. This will lead to an error a decoding in receiver 2 when $m_1 = 1$ is transmitted
- $(u_1^n(1, m_t), u_2^n(1, m_s))$ is not typical in (U_1^n, U_2^n) this may not cause error at description.

So $\varepsilon|t_{1,1} \cap \zeta_0 \neq \varnothing$

In the case where $(v_1^n(1), y_1^n)$ is not jointly typical then receiver 1 will ether declare an error (if there's no other typical $(v_1^n(i), y_1^n)$) or output a wrong $m_1 \neq 1$ so $\zeta_{11}|t_{1,1} \subseteq \varepsilon$ when $m_1 = 1$ is transmitted

In the case where there is $(u_1^n(m_1 \neq 1, m_s), y_1^n)$ then there's a chance that receiver 1 decide to output $m_1 \neq 1$ when $m_1 = 1$ is transmitted. so $\varepsilon|t_{1,1} \cap \zeta_{12} \neq \varnothing$

In the case where $(v_2^n(1), y_2^n)$ is not jointly typical then receiver 2 will ether declare an error (if there's no other typical $(v_2^n(i), y_2^n)$) or output a wrong $m_1 \neq 1$ so $\zeta_{11}|t_{1,1} \subseteq \varepsilon$ when $m_2 = 1$ is transmitted

In the case where there is $(u_2^n(m_2 \neq 1, m_s), y_2^n)$ then there's a chance that receiver 2 decide to output $m_1 \neq 1$ when $m_1 = 1$ is transmitted. so $\varepsilon|t_{1,1} \cap \zeta_{12} \neq \varnothing$

tif an error occur it's come from one of this event. so $\varepsilon|t_{1,1} \subseteq \zeta_0 \cup \zeta_{11} \cup \zeta_{12} \cup \zeta_{21} \cup \zeta_{22}$

so

$$\Pr(\varepsilon) = \Pr(\varepsilon|t_{1,1}) \leq \Pr(\zeta_0 \cup \zeta_{11} \cup \zeta_{12} \cup \zeta_{21} \cup \zeta_{22}|t_{1,1})$$

(c)

first we have $(\zeta_0 \cup \zeta_{11} \cup \zeta_{12} \cup \zeta_{21} \cup \zeta_{22}|t_{1,1}) = (\zeta_0 \cup (\zeta_{11} \cap \zeta_0^c) \cup \zeta_{12} \cup (\zeta_{21} \cap \zeta_0^c) \cup \zeta_{22}|t_{1,1})$ (because $\zeta_{1i} = (\zeta_{1i} \cap \zeta_0^c) \cup (\zeta_{1i} \cap \zeta_0)$ and $(\zeta_{1i} \cap \zeta_0) \subseteq \zeta_0$

and so by using union bound:

$$\Pr(\varepsilon) \leq \Pr(\zeta_0 \cup \zeta_{11} \cup \zeta_{12} \cup \zeta_{21} \cup \zeta_{22}|t_{1,1})$$
$$\Pr(\varepsilon) \leq \Pr\left(\zeta_0 \cup (\zeta_{11} \cap \zeta_0^c) \cup \zeta_{12} \cup (\zeta_{21} \cap \zeta_0^c) \cup \zeta_{22}|t_{1,1}\right)$$
$$\Pr(\varepsilon) \leq \Pr\left(\zeta_0|t_{1,1}\right) + \Pr\left(\zeta_{11} \cap \zeta_0^c|t_{1,1}\right) + \Pr\left(\zeta_{12}|t_{1,1}\right) + \Pr\left(\zeta_{21} \cap \zeta_0^c|t_{1,1}\right) + \Pr\left(\zeta_{22}|t_{1,1}\right)$$

(d)

If we fix $m_1 = m_2 = 1$ then we can apply the lemma on the function $(u_1^n(1, m_t), u_2^n(1, m_s))$ because all component of our system matches condition on the lemma. and we have that $(1, m_t)$ has cardinality 2^{nR_t} and $(1, m_s)$ has cardinality 2^{nR_s}.

And so the lemma give us that as long has $R_s + R_t > I(U_1; U_2)$ then

$$\lim_{n\to\infty} \Pr\left(\exists(1, m_s), (1, m_t) \,:\, (u_1^n(1, m_t), u_2^n(1, m_s)) \in \mathcal{A}_\epsilon^{(n)}(U_1, U_2)\right) = 1$$
$$\lim_{n\to\infty} \Pr\left(\exists(m_s, m_t) \,:\, (u_1^n(1, m_t), u_2^n(1, m_s)) \in \mathcal{A}_\epsilon^{(n)}(U_1, U_2)\right) = 1$$
$$\lim_{n\to\infty} \Pr\left(\forall(m_s, m_t) \,:\, (u_1^n(1, m_t), u_2^n(1, m_s)) \notin \mathcal{A}_\epsilon^{(n)}(U_1, U_2)\right) = 1 - 1 = 0$$
$$\lim_{n\to\infty} \Pr\left(\zeta_0|t_{1,1}\right) = 0$$

(e)

We have in our system that $x_i^n(v_1^n(1), v_2^n(1)) = x(v_{1,i}^n(1), v_{2,i}^n(1))$ so in our distribution

$$p(x^n(v_1^n(1), v_2^n(1)), v_1^n(1), v_2^n(1)) = p\left(v_{1,i}^n(1), v_{2,i}^n(1)\right) \prod_{i=1}^{n} p\left(x\left(v_{1,i}^n(1), v_{2,i}^n(1)\right) | v_{1,i}^n(1), v_{2,i}^n(1)\right)$$

And so we can apply the conditional lemma: if $v_1^n(1), v_2^n(1)$ is a typical sequence then $(x^n, v_1^n(1), v_2^n(1))$ is a typical sequence with probability 1

Also we know that (because we assume both channel are memoryless) we have that

$$p\left(y_1^n, y_2^n | x\left(v_1^n(1), v_i^n(1)\right), v_1^n(1), v_2^n(1)\right) = \prod_{i=1}^{n} p\left(y_{1,i}^n, y_{2,i}^n | x\left(v_{1,i}^n(1), v_{2,i}^n(1)\right), v_{1,i}^n(1), v_{2,i}^n(1)\right)$$

And so can also apply the lemma: if $x\left(v_1^n(1), v_i^n(1)\right), v_1^n(1), v_2^n(1)$ is typical then $\left(y_1^n, y_2^n, x\left(v_1^n(1), v_i^n(1)\right), v_1^n(1), v_2^n(1)\right)$ is also typical with probability 1.

But we already have that if $v_1^n(1), v_2^n(1)$ is a typical sequence (so in the event ζ_0^c) then $(x^n, v_1^n(1), v_2^n(1))$ is Typical with probability one (as $x \to \infty$) and so we can extend our result to the combination of these event and have that

$$\lim_{n\to\infty} \Pr\left[y_1^n, y_2^n, x\left(v_1^n(1), v_i^n(1)\right), v_1^n(1), v_2^n(1) \text{is typical} | v_1^n(1), v_2^n(1) \text{is typical}, t_{1,1}\right] = 1$$

$$\lim_{n\to\infty} \Pr\left(\zeta_0^c \cap \zeta_{11} \cap \zeta_{21} | t_{1,1}\right) = 0$$

$$\begin{cases} \lim_{n\to\infty} \Pr\left(\zeta_0^c \cap \zeta_{11} | t_{1,1}\right) &= 0 \\ \lim_{n\to\infty} \Pr\left(\zeta_0^c \cap \zeta_{21} | t_{1,1}\right) &= 0 \end{cases}$$

(f)

From

According to the system specification. We have our distribution $p(u_1^n u_2^n, x^n, y_1^n, y_2^n) = p(u_1^n, u_2^n)p(x^n|u_1^n, u_2^n)p(y_1^n, y_2^n|x^n) = p(u_1^n)p(u_2^n)p(x^n|u_1^n, u_2^n)p(y_1^n|x^n)p(y_2^n|x^n)$

We have that $u_1^n(1, m_t)$ is independent from all other $u_1^n(m_1', m_2')$ with $(1, m_t) \neq (m_1', m_t')$

And also $u_2^n(1, m_{ts})$ is independent from all other $u_2^n(m_2', m_s')$ with $(1, m_s) \neq (m_2', m_s')$

And as

$$p(u_1^n u_2^n, x^n, y_1^n, y_2^n) = p(u_1^n, u_2^n)p(x^n|u_1^n, u_2^n)p(y_1^n, y_2^n|x^n) = p(u_1^n)p(u_2^n)p(x^n|u_1^n, u_2^n)p(y_1^n|x^n)p(y_2^n|x^n)$$

In word: as x^n is draw only depending of the given u_1^n, u_2^n and y_1^n and y_2^n only from x^n then the message receive then as the variable y_1^n and y_2^n which are the message receive for initial messages $((1, m_t), (1, m_s))$ are independent tof the message received receive from $(m_1', m_t') \neq (1, m_t)$ and $(m_2', m_s') \neq (1, m_s)$.

In other word $y_1^n|t_{1,1}$ and $y_2^n|t_{1,1}$ are independent of $u_1^n(m_1', m_t')$ for $(m_1', m_t') \neq (1, m_t)$ and $u_1^n(m_1', m_t')$ for $(m_1', m_t') \neq (1, m_t)$

So we can apply the cuckoo's Egg lemma on 2 different case :

$$\Pr\left(u_1^n(m_1', m_t'), y_1^n|t_{1,1} \in \mathcal{A}_\epsilon^{(n)}(U_1^n, Y^n)\right) < 2^{-n(I(U_1^n;Y_1^n)-3\epsilon)}$$
$$\Pr\left(u_2^n(m_2', m_S'), y_2^n|t_{1,1} \in \mathcal{A}_\epsilon^{(n)}(U_2^n, Y^n)\right) < 2^{-n(I(U_1^n;Y_1^n)-3\epsilon-}$$

and so

$$\Pr(\zeta_{12}|t_{1,1}) = \sum_{\forall(m_1', m_t') \neq (1, m_t)} \Pr\left(u_1^n(m_1', m_t'), y_1^n|t_{1,1} \in \mathcal{A}_\epsilon^{(n)}(U_1, Y)\right)$$
$$\Pr(\zeta_{12}|t_{1,1}) < \sum_{\forall(m_1', m_t') \neq (1, m_t)} 2^{-n(I(U_1^n;Y_1^n)-3\epsilon)}$$
$$\Pr(\zeta_{12}|t_{1,1}) < \sum_{i=1}^{nR_1} \sum_{j=1}^{nR_t} 2^{-n(I(U_1^n;Y_1^n)-3\epsilon)}$$
$$\Pr(\zeta_{12}|t_{1,1}) < 2^{n(R_1+R_t)} 2^{-n(I(U_1^n;Y_1^n)-3\epsilon)}$$
$$\Pr(\zeta_{12}|t_{1,1}) < 2^{-n(I(U_1^n;Y_1^n)-R_1-R_t-3\epsilon)}$$

And by the exact same process on the second inequality we have

$$\Pr(\zeta_{22}|t_{1,1}) = \sum_{\forall(m_2', m_s') \neq (1, m_s)} \Pr\left(u_2^n(m_2', m_s'), y_1^n|t_{1,1} \in \mathcal{A}_\epsilon^{(n)}(U_2^n, Y^n)\right)$$
$$\Pr(\zeta_{22}|t_{1,1}) < \sum_{\forall(m_2', m_s') \neq (1, m_s)} 2^{-n(I(U_2^n;Y_2^n)-3\epsilon)}$$
$$\Pr(\zeta_{22}|t_{1,1}) < \sum_{i=1}^{nR_2} \sum_{j=1}^{nR_s} 2^{-n(I(U_2^n;Y_2^n)-3\epsilon)}$$
$$\Pr(\zeta_{22}|t_{1,1}) < 2^{n(R_2+R_s)} 2^{-n(I(U_2^n;Y_2^n)-3\epsilon)}$$
$$\Pr(\zeta_{22}|t_{1,1}) < 2^{-n(I(U_2^n;Y_2^n)-R_2-R_s-3\epsilon)}$$

So we have that these 2 probability vanish for $n \to \infty$ as long as we have $R_1 + R_t \leq I(U_1^n; Y_1^n)$ and $R_2 + R_s \leq I(U_2^n; Y_2^n)$

(g)

we have

$$\begin{cases} R_1 \leq I(U_1^n; Y_1^n) & (i) \\ R_2 \leq I(U_2^n; Y_2^n) & (ii) \\ R_2 + R_1 \leq I(U_1^n; Y_1^n) + I(U_2^n; Y_2^n) - I(U_1^n; U_2^n) & (iii) \end{cases}$$

If we take $R_t = I(U_1^n; Y_1^n) - R_1$ and $R_s = I(U_2^n; Y_2^n) - R_2$ we get

$$\begin{cases} R_1 + R_t \leq I(U_1^n; Y_1^n) \\ R_2 + R_s \leq I(U_2^n; Y_2^n) \end{cases}$$

These are regular rate because from from (i) and (ii) we have $0 \leq I(U_1^n; Y_1^n) - R_1$, $0 \leq I(U_2^n; Y_2^n) - R_2$

Also from (iii) we have $I(U_1^n; U_2^n) \leq I(U_1^n; Y_1^n) + I(U_2^n; Y_2^n) - R_2 - R_1 \Rightarrow R_s + R_t \geq I(U_1^n; U_2^n)$

so if arbitrary R_1, R_2 satisfy (11) then we can always fix R_s, R_t for which (10) is satisfied

(h)

$$\begin{cases} R_1 + R_t \leq I(U_1^n; Y_1^n) & (i) \\ R_2 + R_s \leq I(U_2^n; Y_2^n) & (ii) \\ R_t + R_s \geq I(U_1^n; U_2^n) & (iii) \end{cases} \tag{10}$$

then as $R_t \geq 0$ and $R_s \geq 0$, by losing precision on (i) and (ii) we get

$$\begin{cases} R_1 \leq I(U_1^n; Y_1^n) \\ R_2 \leq I(U_2^n; Y_2^n) \end{cases}$$

Also by multiplying both side of (iii) by -1 and then add to it (i) we get

$$\begin{aligned} R_2 + R_s + R_1 + R_t - R_t - R_s &\leq I(U_1^n; Y_1^n) + I(U_2^n; Y_2^n) - I(U_1^n; U_2^n) \\ R_2 + R_1 &\leq I(U_1^n; Y_1^n) + I(U_2^n; Y_2^n) - I(U_1^n; U_2^n) \end{aligned}$$

and so we get that (10) implies:

$$\begin{cases} R_1 \leq I(U_1^n; Y_1^n) \\ R_2 \leq I(U_2^n; Y_2^n) \\ R_2 + R_1 \leq I(U_1^n; Y_1^n) + I(U_2^n; Y_2^n) - I(U_1^n; U_2^n) \end{cases} \tag{11}$$

Ex 5

(a)

We can fix $p(u_1, u_2)$ as

$$p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} \sum_{w \in f_1^{-1}(u_1)} \frac{1}{\left|f_1^{-1}(u_1)\right|} \mathbf{1}_{f_2(w)=u_2}$$

because all element are positive (indicator function and size of set) then $p(u_1, u_2) \geq 0$ for all (u_1, u_2) and:

$$\sum_{\forall u_1} \sum_{\forall u_2} p(u_1, u_2) = \sum_{\forall u_1} \sum_{\forall u_2} \frac{1}{|\mathcal{U}_1|} \sum_{w \in f_1^{-1}(u_1)} \frac{1}{\left|f_1^{-1}(u_1)\right|} \mathbf{1}_{f_2(w)=u_2}$$

$$\sum_{\forall u_1} \sum_{\forall u_2} p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} \sum_{\forall u_1} \frac{1}{\left|f_1^{-1}(u_1)\right|} \sum_{w \in f_1^{-1}(u_1)} \sum_{\forall u_2} \mathbf{1}_{f_2(w)=u_2}$$

$$\sum_{\forall u_1} \sum_{\forall u_2} p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} \sum_{\forall u_1} \frac{1}{\left|f_1^{-1}(u_1)\right|} \sum_{w \in f_1^{-1}(u_1)} 1$$

$$\sum_{\forall u_1} \sum_{\forall u_2} p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} \sum_{\forall u_1} \frac{1}{\left|f_1^{-1}(u_1)\right|} \left|f_1^{-1}(u_1)\right|$$

$$\sum_{\forall u_1} \sum_{\forall u_2} p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} \sum_{\forall u_1} 1$$

$$\sum_{\forall u_1} \sum_{\forall u_2} p(u_1, u_2) = 1$$

so this function is indeed a probability function.

Then if $f_1^{-1}(u_1) \cap f_2^{-1}(u_2) \neq \varnothing$ so by definition $\forall w \in f_1^{-1}(u_1) \,:\, f_2(w) \neq u_2$ .in this case

$$p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} \sum_{w \in f_1^{-1}(u_1)} \frac{1}{\left|f_1^{-1}(u_1)\right|} \mathbf{1}_{f_2(w)=u_2}$$

$$p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} \sum_{w \in f_1^{-1}(u_1)} \frac{1}{\left|f_1^{-1}(u_1)\right|} * 0$$

$$p(u_1, u_2) = \frac{1}{|\mathcal{U}_1|} * 0 = 0$$

Let's note that as we know have $\Pr\left(f_1^{-1}(u_1) \cap f_2^{-1}(u_2) \neq \varnothing\right) = 0$ we can rewrite $x(u_1, u_2)$ simply as $x(u_1, u_2) =$ some $a \in f_1^{-1}(u_1) \cap f_2^{-1}(u_2)$ so now $\forall u_1 \in \mathcal{U}_1 \forall u_2 \in \mathcal{U}_2 \,:\, x(u_1, u_2) \in f_1^{-1}(u_1) \wedge x(u_1, u_2) \in f_2^{-1}(u_2)$

Now we want $\Pr(Y_1 = U_1)$. As deterministically we have $y_1 = f_1(x(u_1, y_2))$ then

$$\Pr(Y_1 = U_1) = \Pr(Y_1 = f_1(x(U_1, U_2)) = \Pr\left(x(U_1, U_2) \in f_1^{-1}(U_1)\right) = 1$$

from what we derived in the previous paragraph.

Symmetrically:

$$\Pr(Y_2 = U_2) = \Pr(Y_2 = f_2(x(U_1, U_2)) = \Pr\left(x(U_1, U_2) \in f_2^{-1}(U_2)\right) = 1$$

(b).

case when $f_1^{-1}(u_1) \cap f_2^{-1}(u_2) = \varnothing$ if it exist a and b in such that $f_1^{-1}(a) \cap f_2^{-1}(b) = \varnothing$ then by definition there's no $w \in \mathcal{X}$ such that $f_2(w) = a \wedge f_2(w) = b$.

So $\Pr_{f_1(Z), f_2(Z)}(a, b) = 0$. so if we set in our distribution $p(a, b) = 0$, then $\Pr\left(f_1\left(x(a, b)\right), f_2\left(x(a, b)\right)\right) = 0$ and it's match.

So here we have in distribution $(f_1(Z), f_2(Z)) = (f_1(x(U_1, U_2)), f_2(x(U_1, U_2)))$

case when $f_1^{-1}(u_1) \cap f_2^{-1}(u_2) \neq \varnothing$ Now that we know that

$$\Pr_{U_1, U_2}\left(a, b | f_1^{-1}(a) \cap f_2^{-1}(b) \neq \varnothing\right) = 0$$

then again we have $f_1\left(x(a, b)\right) = a$ and $f_2\left(x(a, b)\right) = b$. so we want to match the distribution of $(f_1(Z), f_2(Z))$ and (U_1, U_2). And we have in general that for any function f $\Pr\left(f(x)\right) = \sum_{w\,:\,f(w) = f(x)} \Pr(w)$. So

$$\Pr\left(f_1(z), f_2(z)\right) = \sum_{w\,:\,f_1(w), f_2(w) = f_1(z), f_2(z)} p_z(w)$$
$$\Pr\left(f_1(z), f_2(z)\right) = \sum_{w\,:\,f_1(w) = f_1(z) \wedge f_2(w) = f_2(z)} p_z(w)$$
$$\Pr\left(f_1(z), f_2(z)\right) = \sum_{w\,:\,f_1(w) = f_1(z) \wedge f_2(w) = f_2(z)} p_z(w)$$
$$\Pr\left(f_1(z), f_2(z)\right) = \sum_{w \in f_1^{-1}(f_1(z)) \cap f_2^{-1}(f_2(z))} p_z(w)$$

So if we set $p(u_1, u_2) = \sum_{z \in f_1^{-1}(u_1) \cap f_2^{-1}(u_2)} p_z(z)$

we get in distribution $((f_1(Z), f_2(Z)) = (U_1, U_2) = (f_1(x(U_1, U_2), f_2(x(U_1, U_2)))$.

Conclusion. Let's note that if $f_1^{-1}(u_1) \cap f_2^{-1}(u_2) = \varnothing$ then $\sum_{z \in f_1^{-1}(u_1) \cap f_2^{-1}(u_2)} p_z(z) = 0$, then we have that the distribution $p(u_1, u_2) = \sum_{z \in f_1^{-1}(u_1) \cap f_2^{-1}(u_2)} p_z(z)$ makes that $(f_1(x(U_1, U_2), f_2(x(U_1, U_2))$ and $(f_1(Z), f_2(Z))$ to have the same distribution

(c)

we have

$$\mathcal{R} = \bigcup_{p(x)} \begin{cases} R_1 \leq I(U_1^n; Y_1^n) \\ R_2 \leq I(U_2^n; Y_2^n) \\ R_2 + R_1 \leq I(U_1^n; Y_1^n) + I(U_2^n; Y_2^n) - I(U_1^n; U_2^n) \end{cases}$$

But for every $p(x)$ we can find $p(u_1, u_2)$ such that the distribution of (U_1^n, U_2^n) match the one of (Y_1^n, Y_2^n) and so we can reach

$$\bigcup_{p(x)} \begin{cases} R_1 \leq I(Y_1^n; Y_1^n) \\ R_2 \leq I(Y_2^n; Y_2^n) \\ R_2 + R_1 \leq I(Y_1^n; Y_1^n) + I(Y_2^n; Y_2^n) - I(Y_1^n; Y_2^n) \end{cases}$$
$$\bigcup_{p(x)} \begin{cases} R_1 \leq H(Y_1^n) \\ R_2 \leq H(Y_2^n) \\ R_2 + R_1 \leq H(Y_1^n) + H(Y_2^n) - (H(Y_1^n) + H(Y_2^n) - H(Y_1^n; Y_2^n)) \end{cases}$$
$$\bigcup_{p(x)} \begin{cases} R_1 \leq H(Y_1^n) \\ R_2 \leq H(Y_2^n) \\ R_2 + R_1 \leq H(Y_1^n; Y_2^n) \end{cases} \tag{12}$$

and so rate region (12) is reachable

Ex 6

(a)

$$X = \begin{cases} \log p_0 + p_1 \\ \log p_2 \\ \log p_3 \end{cases}$$

we have

$$H(p_0 + p_1, p_2, p_3) = -(p_0 + p_1) \log (p_0 + p_1) - p_3 \log(p_3) - p_2 \log(p_2)$$
$$H(p_0 + p_1, p_2, p_3) < -(p_0 + p_1) \log (p_0 + p_1) - (p_3 + p_2) \log \left(\frac{p_3 + p_2}{1 + 1} \right)$$
$$H(p_0 + p_1, p_2, p_3) < H(p_0' + p_1', p_2', p_3')$$

where we use the log rule inequality between first and second line.

Also in the same way

$$
\begin{aligned}
H(p_0, p_1, p_2+p_3) &= -(p_2+p_3)\log(p_2+p_3) - p_1\log(p_1) - p_0\log(p_0) \\
H(p_0+p_1, p_2, p_3) &< -(p_2+p_3)\log(p_2+p_3) - (p_1+p_0)\log\left(\frac{p_1+p_0}{1+1}\right) \\
H(p_0+p_1, p_2, p_3) &< H(p'_0, p'_1, p'_2+p'_3)
\end{aligned}
$$

and finally

$$
\begin{aligned}
H(p_0, p_1, p_2, p_3) &= -p_1\log(p_1) - p_0\log(p_0) - p_2\log(p_2) - p_3\log(p_3) \\
H(p_0, p_1, p_2, p_3) &< -(p_1+p_0)\log\left(\frac{p_1+p_0}{1+1}\right) - (p_3+p_2)\log\left(\frac{p_3+p_2}{1+1}\right) \\
H(p_0, p_1, p_2, p_3) &< -\frac{p_1+p_0}{2}\log\left(\frac{p_1+p_0}{2}\right) - \frac{p_1+p_0}{2}\log\left(\frac{p_1+p_0}{2}\right) \\
&\quad - \frac{p_3+p_2}{2}\log\left(\frac{p_3+p_2}{2}\right) - \frac{p_3+p_2}{2}\log\left(\frac{p_3+p_2}{2}\right) \\
H(p_0, p_1, p_2, p_3) &< H(p'_0, p'_1, p'_2, p'_3)
\end{aligned}
$$

where we used the inequality twice.

So he have that all upper bound in $\mathcal{R}_{p'}$ are bigger than upper bound in $\mathcal{R}_p$ so $\mathcal{R}_p \subseteq \mathcal{R}_{p'}$

(b)

The variable X as 4 possible value, will will denote these probability by the vector $\vec{p_X} = (p_0 p_1, p_3, p_4)$ where $p_i = \Pr[X = i]$ from these probabilities we can compute deterministacly (because $y_i(x)$ are deterministic) the probability vector of Y_1 and Y_2:

$\vec{p}_{Y_1} = (p_0 + p_1, p_2, p_3)$ and $\vec{p}_{Y_2} = (p_0, p_1, p_2 + p_3)$

Also we can see that each value of x gives one unique and distinct pair $y_1(x), y_2(x)$ and so we have

$$
\begin{cases}
p_{Y_1,Y_2}(0,0) = p_X(0) = p_0 \\
p_{Y_1,Y_2}(0,1) = p_X(1) = p_1 \\
p_{Y_1,Y_2}(1,2) = p_X(2) = p_2 \\
p_{Y_1,Y_2}(2,2) = p_X(3) = p_3
\end{cases}
$$

and so by applying it to the rate region from Ex5 we get that that the rate region $\mathcal{R}$ is in fact the rate region $\bigcup_{\forall \vec{p_X}} \mathcal{R}_{\vec{p_X}}$ where $\mathcal{R}_{\vec{p_X}}$ is the region define in (a) for a given vector $\vec{p_X}$

We have that all distribution of the form $\vec{p_X} = (p, p, 0.5 - p, 0.5 - p)$ with $0 \leq p \leq 0,5$ is a subset of all possible $\vec{p_X}$ so

$$\bigcup_{\forall \vec{p_X}} \mathcal{R}_{\vec{p_X}} \supseteq \bigcup_{0 \leq p \leq 0,5} \mathcal{R}_{(p,p,1-p,1-p)} \tag{13}$$

Now for every channel $\vec{p_X} = (p_0 p_1, p_3, p_4)$ if we set that $p' = \frac{p_1+p_0}{2}$ and so $0,5-p' = \frac{p_2+p_3}{2}$ we have from (a) that $\mathcal{R}_{(p_0 p_1, p_3, p_4)} \subseteq \mathcal{R}_{(p',p',0.5-p',0.5-p')}$ and as $p_0 + p_1 \leq 1 \Rightarrow p' = \frac{p_1+p_0}{2} \leq 0.5$ then we have that $\mathcal{R}_{(p_0 p_1, p_3, p_4)} \subseteq \mathcal{R}_{(p,p,1-p,1-p)}$ for a $p = p' \leq 0.5$. so if we take the union on both side we get:

$$\bigcup_{\forall \vec{p_X})} \mathcal{R}_{\vec{p_X}} \subseteq \bigcup_{0 \leq p \leq 0,5} \mathcal{R}_{(p,p,0.5-p,0.5-p)} \tag{14}$$

and so (13) and (14) implies

$$\bigcup_{\forall \vec{p_X}} \mathcal{R}_{\vec{p_X}} = \bigcup_{0 \leq p \leq 0,5} \mathcal{R}_{(p,p,0.5-p,0.5-p)}$$
$$\mathcal{R} = \bigcup_{0 \leq p \leq 0,5} \mathcal{R}_{(p,p,0.5-p,0.5-p)}$$

(c)

We take $p_l = \frac{1}{6}$ and $p_r = \frac{1}{3}$

Preliminary result:

$$H(p,p,0.5-p,0.5-p) = -2p\log(p) - 2(0.5-p)\log(0,5-p)$$
$$H(p,p,0.5-p,0.5-p) = -2p\,(\log(2p) - \log 2) - (1-2p)(\log(1-2p) - \log 2)$$
$$H(p,p,0.5-p,0.5-p) = h_b(2p) + \log 2$$

as we know that binary entropy is increasing for $x \leq 0.5$, if we have $a \leq 0.25$ and $b \leq 0.25$:

$$H(a,a,0.5-a,0.5-a) \leq H(b,b,0.5-b,0.5-b) \iff a \leq b \tag{15}$$

In the same way:

$$H(p,p,1-2p) = -2p\log(p) - (1-2p)\log(1-2p)$$
$$H(p,p,1-2p) = -2p\,(\log(2p) - \log 2) - (1-2p)\log(1-2p)$$
$$H(p,p,1-2p) = h_b(2p) + 2p\log 2$$

Both $h_b(2p)$ and $2p\log 2$ are increasing for $2p \leq 0.5$ so for $a \leq 0.25$ and $b \leq 0.25$:

$$H(a,a,1-2p) \leq H(b,b,1-2b) \iff a \leq b \tag{16}$$

(i)

for every $p \leq \frac{1}{6}$ we have that

- $R_1 \leq H(p+p, 0.5-p, 0.5-p) \leq H(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, because entropy always reach is maximum at equiprobability for the same number of possible value (here 3)
- $R_2 \leq H(p, p, 1-2p) \leq H(\frac{1}{6}, \frac{1}{6}, \frac{2}{3})$ because $p \leq \frac{1}{6} \leq 0.25$ and (16)
- $R_1 + R_2 \leq H(p, p, 0.5-p, 0.5-p) \leq H(\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{3})$ from $p \leq \frac{1}{6} \leq 0.25$ and (15)

So all bound of $\mathcal{R}_p$ are bounded by the bound of $\mathcal{R}_{\frac{1}{6}}$

so $\mathcal{R}_{(p,p,0.5-p,0.5-p)} \subseteq \mathcal{R}_{\frac{1}{6}} \subseteq \bigcup_{\frac{1}{6} \leq p \leq \frac{1}{3}} \mathcal{R}_{(p,p,0.5-p,0.5-p)}$ for all $p < \frac{1}{6}$

Now for every $p \geq \frac{1}{3}$ we have:

- $R_1 \leq H(p+p, 0.5-p, 0.5-p) \leq H(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$
  because
  $$H(p+p, 0.5-p, 0.5-p) \leq H(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}) \iff H(0.5-p, 0.5-p, p+p) \leq H(\frac{1}{6}, \frac{1}{6}, \frac{2}{3})$$
  and $0.5 - p \leq \frac{1}{6} \leq 0.25$ and (16)
- $R_2 \leq H(p, p, 1-2p) \leq H(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ because entropy always reach is maximum at equiprobability for the same number of possible value (here 3)
- $R_1 + R_2 \leq H(p, p, 0.5-p, 0.5-p) \leq H(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6})$
  because
  $$H(p, p, 0.5-p, 0.5-p) \leq H(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}) \iff H(0.5-p, 0.5-p, p, p) \leq H(\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{3})$$
  and $0.5 - p \leq \frac{1}{6} \leq 0.25$ and (15)

So all bound of $\mathcal{R}_p$ are bounded by the bound of $\mathcal{R}_{(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6})}$

$\mathcal{R}_{(p,p,0.5-p,0.5-p)} \subseteq \mathcal{R}_{(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6})} \subseteq \bigcup_{\frac{1}{6} \leq p \leq \frac{1}{3}} \mathcal{R}_{(p,p,0.5-p,0.5-p)}$ for all $p > \frac{1}{3}$

Finally obviously for $\frac{1}{6} \leq p' \leq \frac{1}{3}$: $\mathcal{R}_{(p',p',0.5-p',0.5-p')} \subseteq \bigcup_{\frac{1}{6} \leq p \leq \frac{1}{3}} \mathcal{R}_{(p,p,0.5-p,0.5-p)}$

Conclusion: By combination of all the previous paragraph we get that $\bigcup_{0\leq p\leq 0,5}\mathcal{R}_{(p,p,0.5-p,0.5-p)} \subseteq \bigcup_{\frac{1}{6}\leq p\leq\frac{1}{3}}\mathcal{R}_{(p,p,0.5-p,0.5-p)}$

also as $\frac{1}{6}\leq p\leq\frac{1}{3}$ is a subset of $0\leq p\leq 0,5$ then $\bigcup_{0\leq p\leq 0,5}\mathcal{R}_{(p,p,0.5-p,0.5-p)} \supseteq \bigcup_{\frac{1}{6}\leq p\leq\frac{1}{3}}\mathcal{R}_{(p,p,0.5-p,0.5-p)}$

So

$$\begin{aligned}\bigcup_{0\leq p\leq 0,5}\mathcal{R}_{(p,p,0.5-p,0.5-p)} &= \bigcup_{\frac{1}{6}\leq p\leq\frac{1}{3}}\mathcal{R}_{(p,p,0.5-p,0.5-p)}\\ \mathcal{R} &= \bigcup_{\frac{1}{6}\leq p\leq\frac{1}{3}}\mathcal{R}_{(p,p,0.5-p,0.5-p)}\end{aligned}$$

(ii)

In all the following we assume that $\epsilon < 0$.

We know that the entropy of a random variable with 3 possible value only reach it's maximum $\log 3$ at equiprobability.

So $H(a,b,c)=\log 3 \iff a=b=c=\frac{1}{3}$ otherwise $H(a,b,c)<\log 3$

Lower bound can't be larger than $\frac{1}{6}$: We have:

- $\log 3\leq H(\frac{1}{3},\frac{1}{3},\frac{1}{3})$, because $\log 3 = H(\frac{1}{3},\frac{1}{3},\frac{1}{3})$
- $0\leq H(\frac{2}{3},\frac{1}{6},\frac{1}{6})$, by basic property of entropy and
- $\log 3+0\leq H(\frac{1}{3},\frac{1}{3},\frac{1}{6},\frac{1}{6})$. by explicit computation

so

$$(\log 3,0)\in\mathcal{R}_{(\frac{1}{6},\frac{1}{6},\frac{1}{3},\frac{1}{3})}$$

Now for every $\mathcal{R}_{(p,p,0.5-p,0.5-p)}$ with $\frac{1}{6}+\epsilon\leq p$ we have that $2p=\frac{1}{3}+2\epsilon\neq\frac{1}{3}$ so the first bound we get $R_1\leq H(p+p,0.5-p,0.5-p)<\log 3$ so $(\log 3,0)\notin\mathcal{R}_{(p,p,0.5-p,0.5-p)}$ for any $\frac{1}{6}+\epsilon\leq p$

So

$$(\log 3,0)\notin\bigcup_{\frac{1}{6}+\epsilon\leq p\leq\frac{1}{3}}\mathcal{R}_{(p,p,0.5-p,0.5-p)}$$

Upper bound can't be smaller than $\frac{1}{3}$:

- $0\leq H(\frac{2}{3},\frac{1}{6},\frac{1}{6})$, by basic property of entropy and
- $\log 3\leq H(\frac{1}{3},\frac{1}{3},\frac{1}{3})$, because $\log 3 = H(\frac{1}{3},\frac{1}{3},\frac{1}{3})$

- $0 + \log 3 \leq H(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6})$. by explicit computation

So

$$(0, \log 3) \in \mathcal{R}_{(\frac{1}{3},\frac{1}{3},\frac{1}{6},\frac{1}{6})}$$

Now for every $\mathcal{R}_{(p,p,0.5-p,0.5-p)}$ with $p \leq \frac{1}{3} - \epsilon$ we have that $p = \frac{1}{3} - \epsilon \neq \frac{1}{3}$ so in the second bound we get $R_2 \leq H(p, p, 1-2p) < \log 3$ so $(0, \log 3) \notin \mathcal{R}_{(p,p,0.5-p,0.5-p)}$ for any $p \leq \frac{1}{3} - \epsilon$ so:

$$(0, \log 3) \notin \bigcup_{\frac{1}{6} \leq p \leq \frac{1}{3} - \epsilon} \mathcal{R}_{(p,p,0.5-p,0.5-p)}$$

So our born wan't be any smaller than $\frac{1}{6}$ or larger than $\frac{1}{3}$

(e)

1. $H\left(p_0, 0.5 - \frac{p_0}{2} + 0.5 - \frac{p_0}{2}\right) = H(p_0, 1-p_0) = H(p_0, p_1+p_2)$ so $\mathcal{R}_{(p_0, 0.5-\frac{p_0}{2}, 0.5-\frac{p_0}{2})}$ and $\mathcal{R}_{\vec{p}}$ have the same bounds on R_1
2. $H\left(0.5 - \frac{p_2}{2} + 0.5 - \frac{p_2}{2}, p_2\right) = H(1-p_2, p_2) = H(p_0, +p_1, p_2)$ so $\mathcal{R}_{(0.5-\frac{p_2}{2}, 0.5-\frac{p_2}{2}, p_2)}$ and $\mathcal{R}_{\vec{p}}$ have the same bounds on R_2
3.
$$\begin{aligned} H(p_0, p_1, p_2) &= -p_0 \log p_0 - p_1 \log p_1 - p_2 \log p_2 \\ H(p_0, p_1, p_2) &< -p_0 \log p_0 - (p_1 + p_2) \log \left(\frac{p_1 + p_2}{1 + 1}\right) \\ H(p_0, p_1, p_2) &< -p_0 \log p_0 - 2\left(\frac{1 - p_0}{2}\right) \log \left(\frac{1 - p_0}{2}\right) \\ H(p_0, p_1, p_2) &< H\left(p_0, 0.5 - \frac{p_0}{2}, 0.5 - \frac{p_0}{2}\right) \end{aligned}$$
   Where the inequality come from the log-sum inequality.
   As $H(p_0, p_1, p_2) = H(p_2, p_0, p_1)$ we also have $H(p_0, p_1, p_2) < H\left(p_2, 0.5 - \frac{p_2}{2}, 0.5 - \frac{p_2}{2}\right) = H\left(0.5 - \frac{p_2}{2}, 0.5 - \frac{p_2}{2}, p_2\right)$
   So the bound on R_1+R_2 is $\mathcal{R}_{\vec{p}}$ in smaller than the ones in $\mathcal{R}_{(p_0, 0.5-\frac{p_0}{2}, 0.5-\frac{p_0}{2})}$ and $\mathcal{R}_{(0.5-\frac{p_2}{2}, 0.5-\frac{p_2}{2}, p_2)}$
4. $H\left(p_0 + 0.5 - \frac{p_0}{2}, 0.5 - \frac{p_0}{2}\right) = h_b\left(0.5 - \frac{p_0}{2}\right)$ and $H(p_0 + p_1, p_2) = h_b(p_2)$
5. $H(0.5 - \frac{p_2}{2}, 0.5 + \frac{p_2}{2}) = h_b\left(0.5 - \frac{p_2}{2}\right)$ and $H(p_0, p_1 + p_2) = h_b(p_0)$

If $p_2 \leq 0.5 - \frac{p_0}{2}$ then we have as $p_0 \leq 1$ $p_2 \leq 0.5 - \frac{p_0}{2} \leq 0.5$ and so $h_b(p_2) \leq h_p\left(0.5 - \frac{p_0}{2}\right)$. that's implies, with 1. and 2. that all bound in $\mathcal{R}_{(p_0, 0.5-\frac{p_0}{2}, 0.5-\frac{p_0}{2})}$ are bigger ore equal than the ones of $\mathcal{R}_{\vec{p}}$ and so $\mathcal{R}_{\vec{p}} \subseteq \mathcal{R}_{(p_0, 0.5-\frac{p_0}{2}, 0.5-\frac{p_0}{2})}$

If $p_0 \le 0.5 - \frac{p_2}{2}$ then we have as $p_2 \le 1$ $p_0 \le 0.5 - \frac{p_2}{2} \le 0.5$ and so $h_b(p_0) \le h_p\left(0.5 - \frac{p_2}{2}\right)$. that's implies, with 1. and 2. that all bound in $\mathcal{R}_{(0.5-\frac{p_2}{2},0.5-\frac{p_2}{2},p_2)}$ are bigger ore equal than the ones of $\mathcal{R}_{\vec{p}}$ and so $\mathcal{R}_{\vec{p}} \subseteq \mathcal{R}_{(0.5-\frac{p_2}{2},0.5-\frac{p_2}{2},p_2)}$

If $p_2 > 0.5 - \frac{p_0}{2}$ and $p_0 > 0.5 - \frac{p_2}{2}$ then by addition we get

$$\begin{aligned}
p_0 + p_1 &> 1 - \frac{1}{2}(p_0 + p_1) \\
1 - p_1 &> \frac{2}{3} \\
p_1 &< \frac{1}{3}
\end{aligned}$$

Also we must have for any point (R_1, R_2) :

If $R_1 \le h_p\left(0.5 - \frac{p_2}{2}\right)$ then as $R_2 \le H(p_0+p_1, p_2)$ we have $(R_1, R_2) \in \mathcal{R}_{(0.5-\frac{p_2}{2},0.5-\frac{p_2}{2},p_2)}$

If $R_2 \le h_p\left(0.5 - \frac{p_0}{2}\right)$ then as $R_1 \le H(p_0, p_1+p_2)$ we have $(R_1, R_2) \in \mathcal{R}_{(0.5-\frac{p_0}{2},0.5-\frac{p_0}{2},p_0)}$

If there's a point (R_1, R_2) in $\mathcal{R}_{\vec{p}}$ such that $h_b(p_2) \ge R_1 > h\left(0.5 - \frac{p_0}{2}\right)$ and $h_b(p_0) \ge R_2 > h_p\left(0.5 - \frac{p_2}{2}\right)$

then

$$\begin{aligned}
R_1 + R_2 &> h\left(0.5 - \frac{p_0}{2}\right) + h_p\left(0.5 - \frac{p_2}{2}\right) \\
R_1 + R_2 - H(p_0, p_1, p_2) &> -\left(0.5 - \frac{p_0}{2}\right)\log\left(0.5 - \frac{p_0}{2}\right) - \left(0.5 + \frac{p_0}{2}\right)\log\left(0.5 + \frac{p_0}{2}\right) \\
&\quad - \left(0.5 - \frac{p_2}{2}\right)\log\left(0.5 - \frac{p_2}{2}\right) - \left(0.5 + \frac{p_2}{2}\right)\log\left(0.5 + \frac{p_2}{2}\right) \\
&\quad + p_0 \log(p_0) + (p_2)\log(p_2) + (p_1)\log(p_1) \\
R_1 + R_2 - H(p_0, p_1, p_2) &> 0
\end{aligned}$$

Where we used the fact that $p_2 > 0.5 - \frac{p_0}{2}$ and $p_0 > 0.5 - \frac{p_2}{2}$ then that $-x \log x$ increase for $x > 0.5$

and finally that as $p_1 < \frac{1}{3}$ then $-(p_1)\log(p_1) < \left(\frac{1}{3}\right)\log\left(\frac{1}{3}\right) < 1$ because $-x \log x$ increasing on $x \le \frac{1}{3}$

So a such point can't exist in $\mathcal{R}_{\vec{p}}$ so in all case $\mathcal{R}_{\vec{p}}$ is either in $\mathcal{R}_{(0.5-\frac{p_2}{2},0.5-\frac{p_2}{2},p_2)}$ or $\mathcal{R}_{(p_0,0.5-\frac{p_0}{2},0.5-\frac{p_0}{2})}$ so in every for every $\mathcal{R}_{\vec{p}}$ so

$$\mathcal{R}_{\vec{p_X}} \subseteq \mathcal{R}_{(p_0,0.5-\frac{p_0}{2},0.5-\frac{p_0}{2})} \cup \mathcal{R}_{(0.5-\frac{p_2}{2},0.5-\frac{p_2}{2},p_2)}$$

(f)

form our information in the channel we can deterministically compute the probability vector of Y_1 and Y_2 and (Y_1, Y_2)given $\vec{p}_X = (p_0, p_1, p_2)$

$$\begin{aligned}
\vec{p}_{Y_1} &= (p_0, p_1 + p_2) \\
\vec{p}_{Y_2} &= (p_0 + p_1, p_2) \\
\vec{p}_{Y_1,Y_2} &= (p_0, p_1, p_2)
\end{aligned}$$

So $H(Y_1) = H(p_0, p_1{+}p_2)$, $H(Y_2) = H(p_0{+}p_1, p_2)$ and $H(Y_1, Y_2) = H(p_0, p_1, p_2)$

So by applying this to rate region from Ex 5 we get

$$\mathcal{R} = \bigcup_{\forall \vec{p}_X} \mathcal{R}_{\vec{p}_X}$$

As all $(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2})$ for $0 \le p \le 1$ and all $(0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}, p)$ for $0 \le p \le 1$ are a subset of all possible $\vec{p}_X$ then

$$\mathcal{R} \supseteq \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2})} \right) \cup \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}, p)} \right)$$

but as for all possible $\vec{p}_X$ We have (from (e)) that $\mathcal{R}_{\vec{p}_X} \subseteq \mathcal{R}_{(p_0, 0.5 - \frac{p_0}{2}, 0.5 - \frac{p_0}{2})} \cup \mathcal{R}_{(0.5 - \frac{p_2}{2}, 0.5 - \frac{p_2}{2}, p)}$ then every $\mathcal{R}_{\vec{p}_X}$ is include in $\left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2})} \right) \cup \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}, p)} \right)$ and so the union of all possible $\mathcal{R}_{\vec{p}_X}$ is included too

$$\mathcal{R} \subseteq \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2})} \right) \cup \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}, p)} \right)$$

This mean

$$\mathcal{R} = \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2})} \right) \cup \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}, p)} \right)$$

(g)

(i)

$$\mathcal{R} = \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2})} \right) \cup \left(\bigcup_{0 \le p \le 1} \mathcal{R}_{(0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}, p)} \right)$$

$$\mathcal{R} = \bigcup_{0 \le p \le 1} \left(\mathcal{R}_{(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2})} \cup \mathcal{R}_{(0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}, p)} \right)$$

$$R = \bigcup_{0 \le p \le 1} \left\{ (R_1, R_2) \,:\, R_1 \le h_b(p)\, R_2 \le H\left(0.5 - \frac{p}{2}, 0.5 + \frac{p}{2}\right)\, R_1 + R_2 \le H\left(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}\right) \right\}$$
$$\cup \left\{ (R_1, R_2) \,:\, R_1 \le h_b\left(0.5 - \frac{p}{2}\right)\, R_2 \le h_b(p)\, R_1 + R_2 \le H\left(p, 0.5 - \frac{p}{2}, 0.5 - \frac{p}{2}\right) \right\}$$

We can chose ether $p_l = 0$ or $p_r = \frac{1}{3}$

if $p > \frac{1}{3}$ then we' have that every point of $\mathcal{R}_{(p,0.5-\frac{p}{2},0.5-\frac{p}{2})}$ is in $\mathcal{R}_{(0.5-\frac{0.5-p}{2},0.5-\frac{0.5-p}{2},0.5-p)}$ because as $p > \frac{1}{3}$ then $0.5 - p > p$ and so $h_b(0.5 - p) \geq h_b(p)$ and $h_b(p) \leq h_b(0.5 - \frac{p}{2})$ and also so all 3 bound are over bounded

the reverse work: every point of $\mathcal{R}_{(0.5-\frac{p}{2},0.5-\frac{p}{2},p)}$ is in $\mathcal{R}_{(0.5-p,0.5-\frac{0.5-p}{2},0.5-\frac{0.5-p}{2})}$

(ii)

So $\mathcal{R} = \bigcup_{0\leq p\leq\frac{1}{3}} \left(\mathcal{R}_{(p,0.5-\frac{p}{2},0.5-\frac{p}{2})} \cup \mathcal{R}_{(0.5-\frac{p}{2},0.5-\frac{p}{2},p)}\right)$

if we take $0 + \epsilon$ instead of 0 we have that the point $(0, 1) \in \mathcal{R}_{(0,0.5,0.5)}$ isn't in our rate because because for all $0 < p < \frac{1}{3}$ $h_b(p) < 1$ and $h_b(0.5 - \frac{p}{2}) < 1$ because $0.5 - \frac{p}{2} < 0$ and so isn't in any $\mathcal{R}_{(p,0.5-\frac{p}{2},0.5-\frac{p}{2})} \cup \mathcal{R}_{(0.5-\frac{p}{2},0.5-\frac{p}{2},p)}$ for $0 < p \leq \frac{1}{3}$