# ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE 

School of Computer and Communication Sciences

Handout 25
Solutions to homework 10

Information Theory and Coding
December 09, 2014

Problem 1.
(a) We have

$$
\begin{aligned}
\mathbb{P}\left[U(1) \neq U^{n} \mid U^{n}=u^{n}\right] & =\mathbb{P}\left[U(1) \neq u^{n} \mid U^{n}=u^{n}\right] \stackrel{(*)}{=} \mathbb{P}\left[U(1) \neq u^{n}\right] \\
& =1-\mathbb{P}\left[U(1)=u^{n}\right]=1-\prod_{i=1}^{n} \mathbb{P}\left[U(1)_{i}=u_{i}\right]=1-\prod_{i=1}^{n} p_{U}\left(u_{i}\right),
\end{aligned}
$$

where $(*)$ follows from the independence of $U(1)$ and $U^{n}$.
(b) An encoding failure happens if and only if $U(m) \neq U^{n}$ for every $1 \leq m \leq M$. Therefore,

$$
\begin{aligned}
\mathbb{P}\left[\text { "failure" } \mid U^{n}=u^{n}\right] & =\mathbb{P}\left[U(m) \neq U^{n}, \forall 1 \leq m \leq M \mid U^{n}=u^{n}\right] \\
& =\mathbb{P}\left[U(m) \neq u^{n}, \forall 1 \leq m \leq M \mid U^{n}=u^{n}\right] \\
& =\mathbb{P}\left[U(m) \neq u^{n}, \forall 1 \leq m \leq M\right] \\
& =\prod_{m=1}^{M}\left(1-\prod_{i=1}^{n} p_{U}\left(u_{i}\right)\right)=\left(1-\prod_{i=1}^{n} p_{U}\left(u_{i}\right)\right)^{M} .
\end{aligned}
$$

(c) Note that if $u^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)$, then $\prod_{i=1}^{n} p_{U}\left(u_{i}\right) \geq 2^{-n H(U)(1+\epsilon)}$, which implies that

$$
\begin{aligned}
\mathbb{P}\left[\text { "failure" } \mid U^{n}=u^{n}\right] & =\left(1-\prod_{i=1}^{n} p_{U}\left(u_{i}\right)\right)^{M} \leq\left(1-2^{-n H(U)(1+\epsilon)}\right)^{M} \\
& \stackrel{(*)}{\leq}\left(e^{-2^{-n H(U)(1+\epsilon)}}\right)^{M}=e^{-M 2^{-n H(U)(1+\epsilon)}}=e^{-2^{n R-n H(U)(1+\epsilon)}},
\end{aligned}
$$

where ( $*$ ) follows from the inequality $1-x \leq e^{-x}$. Therefore, we have

$$
\begin{aligned}
\mathbb{P}\left[\text { "failure" } \mid U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right] & =\frac{\mathbb{P}\left[\text { "failure" }, U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]}{\mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]} \\
& =\frac{\sum_{u^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)} \mathbb{P}\left[\text { "failure", } U^{n}=u^{n}\right]}{\mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]} \\
& =\frac{\sum_{u^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)} \mathbb{P}\left[\text { "failure" } \mid U^{n}=u^{n}\right] \mathbb{P}\left[U^{n}=u^{n}\right]}{\mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}\left(p_{U}\right)\right]} \\
& \leq \frac{\left.\sum_{u^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)} e^{-2^{n R-n H(U)(1+\epsilon)} \mathbb{P}\left[U^{n}\right.}=u^{n}\right]}{\mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]} \\
& =e^{-2^{n R-n H(U)(1+\epsilon)} \frac{\sum_{u^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)} \mathbb{P}\left[U^{n}=u^{n}\right]}{\mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]}} \\
& =e^{-2^{n R-n H(U)(1+\epsilon)} \mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]} \mathbb{\mathbb { P } [ U ^ { n } \in \mathcal { T } _ { \epsilon } ^ { n } ( p _ { U } ) ]}=e^{-2^{n R-n H(U)(1+\epsilon)}} .
\end{aligned}
$$

(d) Assume that $R>H(U)$, then there exists $\epsilon>0$ such that $R>H(U)+\epsilon$. We have

$$
\begin{aligned}
\mathbb{P}[\text { "failure"] } & =\mathbb{P}\left[\text { "failure", } U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]+\mathbb{P}\left[\text { "failure" }, U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)^{c}\right] \\
& =\mathbb{P}\left[\text { "failure" } \mid U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right] \mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]+\mathbb{P}\left[\text { "failure", } U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)^{c}\right] \\
& \leq \mathbb{P}\left[\text { "failure" } \mid U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)\right]+\mathbb{P}\left[U^{n} \in \mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)^{c}\right] \\
& \leq e^{-2^{n R-n H(U)(1+\epsilon)}+\mathbb{P}_{U^{n}}\left(\mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)^{c}\right) .}
\end{aligned}
$$

On the other hand, $\mathbb{P}_{U^{n}}\left(\mathcal{T}_{\epsilon}^{n}\left(p_{U}\right)^{c}\right) \rightarrow 0$ as $n \rightarrow \infty$, and $e^{-2^{n R-n H(U)(1+\epsilon)}} \rightarrow 0$ as $n \rightarrow \infty$ since $R>H(U)+\epsilon$. Therefore, if $R>H(U)$ then $\mathbb{P}$ ["failure"] $\rightarrow 0$ as $n \rightarrow \infty$.

## Problem 2.

(a) For every $0 \leq p \leq 1$, define $\bar{p}:=1-p$. We have:

$$
\begin{equation*}
h_{2}(\bar{p})=-\bar{p} \log \bar{p}-p \log p=-p \log p-\bar{p} \log \bar{p}=h_{2}(p) . \tag{1}
\end{equation*}
$$

On the other hand, it is easy to check that for every $0 \leq p^{\prime}, p^{\prime \prime} \leq 1$, we have:

$$
\overline{p^{\prime}} * p^{\prime \prime}=p^{\prime} * \overline{p^{\prime \prime}}=\overline{p^{\prime} * p^{\prime \prime}} \quad \text { and } \overline{p^{\prime}} * \overline{p^{\prime \prime}}=p^{\prime} * p^{\prime \prime} .
$$

Now (1) implies that

$$
\begin{equation*}
h_{2}\left(\overline{p^{\prime}} * p^{\prime \prime}\right)=h_{2}\left(p^{\prime} * \overline{p^{\prime \prime}}\right)=h_{2}\left(\overline{p^{\prime}} * \overline{p^{\prime \prime}}\right)=h_{2}\left(p^{\prime} * p^{\prime \prime}\right) . \tag{2}
\end{equation*}
$$

Let $p^{\prime}=\mathbb{P}\left[X_{1}=1\right]$ and $p^{\prime \prime}=\mathbb{P}\left[X_{2}=1\right]$. We have the following:

- $\mathbb{P}\left[X_{1} \oplus X_{2}=1\right]=\mathbb{P}\left[X_{1}=1\right] \mathbb{P}\left[X_{2}=0\right]+\mathbb{P}\left[X_{1}=0\right] \mathbb{P}\left[X_{2}=1\right]=p^{\prime} \overline{p^{\prime \prime}}+\overline{p^{\prime}} p^{\prime \prime}=$ $p^{\prime} * p^{\prime \prime}$. Therefore, $H\left(X_{1} * X_{2}\right)=h_{2}\left(p^{\prime} * p^{\prime \prime}\right)$.
- Since $H\left(X_{1}\right)=h_{2}\left(p_{1}\right)$, then we have either $p^{\prime}=p_{1}$ or $p^{\prime}=1-p_{1}$. I.e., we have $p_{1}=p^{\prime}$ or $p_{1}=1-p^{\prime}=\overline{p^{\prime}}$.
- Since $H\left(X_{2}\right)=h_{2}\left(p_{2}\right)$, then we have either $p^{\prime \prime}=p_{2}$ or $p^{\prime \prime}=1-p_{2}$. I.e., we have $p_{2}=p^{\prime \prime}$ or $p_{2}=1-p^{\prime \prime}=\overline{p^{\prime \prime}}$.

Now (2) implies that $H\left(X_{1} \oplus X_{2}\right)=h_{2}\left(p^{\prime} * p^{\prime \prime}\right)=h_{2}\left(p_{1} * p_{2}\right)$.
(b) We have $H\left(X_{1} \mid Y\right)=\sum_{y \in \mathcal{Y}} H\left(X_{1} \mid Y=y\right) \mathbb{P}_{Y}(y)=\sum_{y \in \mathcal{Y}} h_{2}\left(p_{1}(y)\right) q(y)$.

Now for every $y \in \mathcal{Y}, X_{1}$ and $X_{2}$ are independent conditioned on $Y=y$. Moreover, $H\left(X_{1} \mid Y=y\right)=h_{2}\left(p_{1}(y)\right)$ and $H\left(X_{2} \mid Y=y\right)=H\left(X_{2}\right)=h_{2}\left(p_{2}\right)$ since $X_{2}$ and $Y$ are independent. Therefore, Part (a) implies that $H\left(X_{1} \oplus X_{2} \mid Y=y\right)=h_{2}\left(p_{1}(y) * p_{2}\right)$.

We conclude that

$$
\begin{aligned}
H\left(X_{1} \oplus X_{2} \mid Y\right) & =\sum_{y \in \mathcal{Y}} H\left(X_{1} \oplus X_{2} \mid Y=y\right) \mathbb{P}_{Y}(y) \\
& =\sum_{y \in \mathcal{Y}} h_{2}\left(p_{1}(y) * p_{2}\right) q(y)=\sum_{y \in \mathcal{Y}} h_{2}\left(p_{2} * p_{1}(y)\right) q(y) .
\end{aligned}
$$

(c) Note that $p_{2} * p=p\left(1-p_{2}\right)+p_{2}(1-p)=\beta p+p_{2}$, where $\beta=1-2 p_{2} \geq 0$. Let $g(p)=\frac{\frac{\partial}{\partial p} h_{2}\left(p_{2} * p\right)}{\frac{\partial}{\partial p} h_{2}(p)}=\frac{\frac{\partial}{\partial p} h_{2}\left(\beta p+p_{2}\right)}{\frac{\partial}{\partial p} h_{2}(p)}=\frac{\beta h_{2}^{\prime}\left(\beta p+p_{2}\right)}{h_{2}^{\prime}(p)}$. We have

$$
\begin{aligned}
g^{\prime}(p) & =\frac{\beta^{2} h_{2}^{\prime \prime}\left(\beta p+p_{2}\right) h_{2}^{\prime}(p)-\beta h_{2}^{\prime \prime}(p) h_{2}^{\prime}\left(\beta p+p_{2}\right)}{h_{2}^{\prime}(p)^{2}} \\
& =\frac{\beta h_{2}^{\prime \prime}\left(\beta p+p_{2}\right) h_{2}^{\prime \prime}(p)}{h_{2}^{\prime}(p)^{2}}\left[\beta \frac{h_{2}^{\prime}(p)}{h_{2}^{\prime \prime}(p)}-\frac{h_{2}^{\prime}\left(\beta p+p_{2}\right)}{h_{2}^{\prime \prime}\left(\beta p+p_{2}\right)}\right] .
\end{aligned}
$$

Note that $h_{2}^{\prime}(p)=\log \frac{1-p}{p}$ and $h_{2}^{\prime \prime}(p)=\frac{-1}{p(1-p) \ln 2}$, which implies that $h_{2}^{\prime \prime}\left(\beta p+p_{2}\right) \leq 0$ and $h_{2}^{\prime \prime}(p) \leq 0$. Therefore, $\frac{\beta h_{2}^{\prime \prime}\left(\beta p+p_{2}\right) h_{2}^{\prime \prime}(p)}{h_{2}^{\prime}(p)^{2}} \geq 0$ and so it is sufficient to show that we have $\beta \frac{h_{2}^{\prime}(p)}{h_{2}^{\prime \prime}(p)}-\frac{h_{2}^{\prime}\left(\beta p+p_{2}\right)}{h_{2}^{\prime \prime}\left(\beta p+p_{2}\right)} \geq 0$. Now define $\alpha=1-2 p$. It is easy to check the following:

- $p=\frac{1}{2}(1-\alpha)$.
- $1-p=\frac{1}{2}(1+\alpha)$.
- $\beta p+p_{2}=\frac{1}{2}(1-\alpha \beta)$.
- $1-\left(\beta p+p_{2}\right)=\frac{1}{2}(1+\alpha \beta)$.

Therefore, we have

$$
\beta \frac{h_{2}^{\prime}(p)}{h_{2}^{\prime \prime}(p)}=-\beta(\ln 2) p(1-p) \log \frac{1-p}{p}=-\frac{\beta \ln 2}{4}\left(1-\alpha^{2}\right) \log \frac{1+\alpha}{1-\alpha},
$$

and

$$
\frac{h_{2}^{\prime}\left(\beta p+p_{2}\right)}{h_{2}^{\prime \prime}\left(\beta p+p_{2}\right)}=-(\ln 2)\left(\beta p+p_{2}\right)\left(1-\beta p-p_{2}\right) \log \frac{1-\beta p-p_{2}}{\beta p+p_{2}}=-\frac{\ln 2}{4}\left(1-(\alpha \beta)^{2}\right) \log \frac{1+\alpha \beta}{1-\alpha \beta} .
$$

Using the formula $\log (1+x)=\sum_{k \geq 1}(-1)^{k-1} \frac{x^{k}}{k}$, we get

$$
\begin{aligned}
\log \frac{1+x}{1-x}=\log (1+x)-\log (1-x) & =\left(\sum_{k \geq 1}(-1)^{k-1} \frac{x^{k}}{k}\right)-\left(\sum_{k \geq 1}(-1)^{k-1} \frac{(-x)^{k}}{k}\right) \\
& =\sum_{k \geq 1}\left((-1)^{k-1}+1\right) \frac{x^{k}}{k}=2 \sum_{\substack{k \geq 1 \\
k \text { is odd }}} \frac{x^{k}}{k} .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
-\left(1-x^{2}\right) \log \frac{1+x}{1-x} & =-2 \sum_{\substack{k \geq 1 \\
k \text { is odd }}} \frac{x^{k}}{k}+2 \sum_{\substack{k \geq 1 \\
k \text { is odd }}} \frac{x^{k+2}}{k}=-2 x-2 \sum_{\substack{k \geq 3 \\
k \text { is odd }}} \frac{x^{k}}{k}+2 \sum_{\substack{k \geq 3 \\
k \text { is odd }}} \frac{x^{k}}{k-2} \\
& =-2 x+2 \sum_{\substack{k \geq 3 \\
k \text { is odd }}}\left(\frac{1}{k-2}-\frac{1}{k}\right) x^{k} .
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\beta \frac{h_{2}^{\prime}(p)}{h_{2}^{\prime \prime}(p)} & =-\frac{\beta \ln 2}{4}\left(1-\alpha^{2}\right) \log \frac{1+\alpha}{1-\alpha}=\frac{\beta \ln 2}{4}\left[-2 \alpha+2 \sum_{\substack{k \geq 3 \\
k \text { is odd }}}\left(\frac{1}{k-2}-\frac{1}{k}\right) \alpha^{k}\right] \\
& =-\frac{\alpha \beta \ln 2}{2}+\frac{\ln 2}{2} \sum_{\substack{k \geq 3 \\
k \text { is odd }}}\left(\frac{1}{k-2}-\frac{1}{k}\right) \beta \alpha^{k},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{h_{2}^{\prime}\left(\beta p+p_{2}\right)}{h_{2}^{\prime \prime}\left(\beta p+p_{2}\right)} & =-\frac{\ln 2}{4}\left(1-(\alpha \beta)^{2}\right) \log \frac{1+\alpha \beta}{1-\alpha \beta}=\frac{\ln 2}{4}\left[-2 \alpha \beta+2 \sum_{\substack{k \geq 3 \\
k \text { is odd }}}\left(\frac{1}{k-2}-\frac{1}{k}\right)(\alpha \beta)^{k}\right] \\
& =-\frac{\alpha \beta \ln 2}{2}+\frac{\ln 2}{2} \sum_{\substack{k \geq 3 \\
k \text { is odd }}}\left(\frac{1}{k-2}-\frac{1}{k}\right) \beta^{k} \alpha^{k} .
\end{aligned}
$$

We conclude that

$$
\beta \frac{h_{2}^{\prime}(p)}{h_{2}^{\prime \prime}(p)}-\frac{h_{2}^{\prime}\left(\beta p+p_{2}\right)}{h_{2}^{\prime \prime}\left(\beta p+p_{2}\right)}=\frac{\ln 2}{2} \sum_{\substack{k \geq 3 \\ k \text { is odd }}}\left(\frac{1}{k-2}-\frac{1}{k}\right)\left(\beta-\beta^{k}\right) \alpha^{k} \stackrel{(*)}{\geq} 0,
$$

where $(*)$ follows from the fact that $\beta=1-2 p_{2} \leq 1$ which implies that $\beta^{k} \leq \beta$. Therefore, $g^{\prime}(p) \geq 0$ and so $g(p)$ is increasing. We conclude that the function $f$ is convex.
(d) We have

$$
\begin{aligned}
H\left(X_{1} \oplus X_{2} \mid Y\right) & =\sum_{y \in \mathcal{Y}} h_{2}\left(p_{2} * p_{1}(y)\right) q(y)=\sum_{y \in \mathcal{Y}} h_{2}\left(p_{2} * h_{2}^{-1}\left(H\left(X_{1} \mid Y=y\right)\right)\right) q(y) \\
& =\sum_{y \in \mathcal{Y}} f\left(H\left(X_{1} \mid Y=y\right)\right) q(y) \stackrel{(*)}{\geq} f\left(\sum_{y \in \mathcal{Y}} H\left(X_{1} \mid Y=y\right) q(y)\right) \\
& =f\left(H\left(X_{1} \mid Y\right)\right)=h_{2}\left(p_{2} * h_{2}^{-1}\left(H\left(X_{1} \mid Y\right)\right)\right)=h_{2}\left(p_{2} * p_{1}\right)=h_{2}\left(p_{1} * p_{2}\right),
\end{aligned}
$$

where $(*)$ follows from the convexity of the function $f$.
(e) For every $y_{1} \in \mathcal{Y}_{1}$, let $0 \leq p_{1}\left(y_{1}\right) \leq \frac{1}{2}$ be such that $H\left(X_{1} \mid Y_{1}=y_{1}\right)=h_{2}\left(p_{1}\left(y_{1}\right)\right)$ and let $q_{1}\left(y_{1}\right)=\mathbb{P}_{Y_{1}}\left(y_{1}\right)$. Similarly, for every $y_{2} \in \mathcal{Y}_{2}$, let $0 \leq p_{2}\left(y_{2}\right) \leq \frac{1}{2}$ be such that $H\left(X_{2} \mid Y_{2}=y_{2}\right)=h_{2}\left(p_{2}\left(y_{2}\right)\right)$ and let $q_{2}\left(y_{2}\right)=\mathbb{P}_{Y_{2}}\left(y_{2}\right)$. For every $y_{1} \in \mathcal{Y}_{1}$, define the mapping $f_{y_{1}}:[0,1] \rightarrow \mathbb{R}$ as $f_{y_{1}}(h)=h_{2}\left(p_{1}(y) * h_{2}^{-1}(h)\right)$. Part (c) implies that $f_{y_{1}}$ is convex for every $y_{1} \in \mathcal{Y}_{1}$. We have

$$
\begin{aligned}
H\left(X_{1} \oplus X_{2} \mid Y_{1}, Y_{2}\right) & =\sum_{y_{1} \in \mathcal{Y}_{1}} \sum_{y_{2} \in \mathcal{Y}_{2}} h_{2}\left(p_{1}\left(y_{1}\right) * p_{2}\left(y_{2}\right)\right) \mathbb{P}_{Y_{1}, Y_{2}}\left(y_{1}, y_{2}\right) \\
& =\sum_{y_{1} \in \mathcal{Y}_{1}} \sum_{y_{2} \in \mathcal{Y}_{2}} h_{2}\left(p_{1}\left(y_{1}\right) * p_{2}\left(y_{2}\right)\right) q_{1}\left(y_{1}\right) q_{2}\left(y_{2}\right) \\
& =\sum_{y_{1} \in \mathcal{Y}_{1}} q_{1}\left(y_{1}\right) \sum_{y_{2} \in \mathcal{Y}_{2}} h_{2}\left(p_{1}\left(y_{1}\right) * h_{2}^{-1}\left(H\left(X_{2} \mid Y_{2}=y_{2}\right)\right)\right) q_{2}\left(y_{2}\right) \\
& =\sum_{y_{1} \in \mathcal{Y}_{1}} q_{1}\left(y_{1}\right) \sum_{y_{2} \in \mathcal{Y}_{2}} f_{y_{1}}\left(H\left(X_{2} \mid Y_{2}=y_{2}\right)\right) q_{2}\left(y_{2}\right) \\
& \stackrel{(*)}{\geq} \sum_{y_{1} \in \mathcal{Y}_{1}} q_{1}\left(y_{1}\right) f_{y_{1}}\left(\sum_{y_{2} \in \mathcal{Y}_{2}} H\left(X_{2} \mid Y_{2}=y_{2}\right) q_{2}\left(y_{2}\right)\right) \\
& =\sum_{y_{1} \in \mathcal{Y}_{1}} q_{1}\left(y_{1}\right) f_{y_{1}}\left(H\left(X_{2} \mid Y_{2}\right)\right)=\sum_{y_{1} \in \mathcal{Y}_{1}} q_{1}\left(y_{1}\right) h_{2}\left(p_{1}\left(y_{1}\right) * h_{2}^{-1}\left(H\left(X_{2} \mid Y_{2}\right)\right)\right) \\
& =\sum_{y_{1} \in \mathcal{Y}_{1}} q_{1}\left(y_{1}\right) h_{2}\left(p_{1}\left(y_{1}\right) * p_{2}\right)=\sum_{y_{1} \in \mathcal{Y}_{1}} h_{2}\left(p_{2} * h_{2}^{-1}\left(H\left(X_{1} \mid Y_{1}=y_{1}\right)\right)\right) q_{1}\left(y_{1}\right) \\
& =\sum_{y_{1} \in \mathcal{Y}_{1}} f\left(H\left(X_{1} \mid Y_{1}=y_{1}\right)\right) q_{1}\left(y_{1}\right) \stackrel{(* *)}{\geq} f\left(\sum_{y_{1} \in \mathcal{Y}_{1}} H\left(X_{1} \mid Y_{1}=y_{1}\right) q\left(y_{1}\right)\right) \\
& =f\left(H\left(X_{1} \mid Y_{1}\right)\right)=h_{2}\left(p_{2} * h_{2}^{-1}\left(H\left(X_{1} \mid Y_{1}\right)\right)\right)=h_{2}\left(p_{2} * p_{1}\right)=h_{2}\left(p_{1} * p_{2}\right),
\end{aligned}
$$

where $(*)$ follows from the convexity of the functions $\left\{f_{y_{1}}: y_{1} \in \mathcal{Y}_{1}\right\}$ and ( $* *$ ) follows from the convexity of $f$.

## Problem 3.

(a) Since $u^{n} \in \mathcal{T}_{\delta}^{n}\left(p_{U}\right)$, we have $n \mathbb{P}_{U}(a)(1-\delta) \leq n_{a}\left(u^{n}\right) \leq n \mathbb{P}_{U}(a)(1+\delta)$. Therefore, we have:

$$
\begin{aligned}
n_{a, b}\left(u^{n}, v^{n}\right) & \leq n_{a}\left(u^{n}\right) \mathbb{P}_{V \mid U}(b \mid a)(1+\delta) \leq n \mathbb{P}_{U}(a)(1+\delta) \mathbb{P}_{V \mid U}(b \mid a)(1+\delta) \\
& =n \mathbb{P}_{U, V}(a, b)\left(1+2 \delta+\delta^{2}\right) \leq n \mathbb{P}_{U, V}(a, b)(1+3 \delta)
\end{aligned}
$$

and

$$
\begin{aligned}
n_{a, b}\left(u^{n}, v^{n}\right) & \geq n_{a}\left(u^{n}\right) \mathbb{P}_{V \mid U}(b \mid a)(1-\delta) \geq n \mathbb{P}_{U}(a)(1-\delta) \mathbb{P}_{V \mid U}(b \mid a)(1-\delta) \\
& =n \mathbb{P}_{U, V}(a, b)\left(1-2 \delta+\delta^{2}\right) \geq n \mathbb{P}_{U, V}(a, b)(1-3 \delta) .
\end{aligned}
$$

Therefore, $\left(u^{n}, v^{n}\right) \in \mathcal{T}_{3 \delta}^{n}\left(p_{U, V}\right)$.
(b) For every $a \in \mathcal{U}$, let $v_{a}^{n}$ be the subsequence of $v^{n}$ corresponding to the indices $1 \leq i \leq n$ where $u_{i}=a$. Note that the size of the sequence $v_{a}^{n}$ is $n_{a}\left(u^{n}\right)$ (i.e., $v_{a}^{n} \in \mathcal{V}^{n_{a}\left(u^{n}\right)}$ ). Moreover, for every $b \in \mathcal{V}$ we have $n_{b}\left(v_{a}^{n}\right)=n_{a, b}\left(u^{n}, v^{n}\right)$. Therefore, the condition

$$
\begin{equation*}
n_{a}\left(u^{n}\right) \mathbb{P}_{V \mid U}(b \mid a)(1-\delta) \leq n_{a, b}\left(u^{n}, v^{n}\right) \leq n_{a}\left(u^{n}\right) \mathbb{P}_{V \mid U}(b \mid a)(1+\delta) \quad \forall a \in \mathcal{U}, \forall b \in \mathcal{V} \tag{3}
\end{equation*}
$$

is equivalent to the the condition " $v_{a}^{n} \in \mathcal{T}_{\delta}^{n_{a}\left(u^{n}\right)}\left(p_{V \mid U=a}\right)$ for every $a \in \mathcal{V}$ ". Now since $\left|\mathcal{T}_{\delta}^{n_{a}\left(u^{n}\right)}\left(p_{V \mid U=a}\right)\right| \geq(1-\delta) 2^{n_{a}\left(u^{n}\right) H(V \mid U=a)(1-\delta)}$ for every $a \in \mathcal{U}$, and since the correspondence $v^{n} \leftrightarrow\left(v_{a}^{n}\right)_{a \in \mathcal{U}}$ is a one-to-one correspondence, we conclude that there are at least $\prod_{a \in \mathcal{U}}\left[(1-\delta) 2^{n_{a}\left(u^{n}\right) H(V \mid U=a)(1-\delta)}\right]$ sequences $v^{n} \in \mathcal{V}^{n}$ satisfying (3).
(c) Part (b) shows that there are at least $\prod_{a \in \mathcal{U}}\left[(1-\delta) 2^{n_{a}\left(u^{n}\right) H(V \mid U=a)(1-\delta)}\right]$ sequences $v^{n} \in \mathcal{V}^{n}$ satisfying (3). We have

$$
\begin{aligned}
\prod_{a \in \mathcal{U}}\left[(1-\delta) 2^{n_{a}\left(u^{n}\right) H(V \mid U=a)(1-\delta)}\right] & =(1-\delta)^{|\mathcal{U}|} \prod_{a \in \mathcal{U}} 2^{n_{a}\left(u^{n}\right) H(V \mid U=a)(1-\delta)} \\
& =(1-\delta)^{|\mathcal{U}|} 2^{\sum_{a \in \mathcal{U}} n_{a}\left(u^{n}\right) H(V \mid U=a)(1-\delta)} \\
& \geq(1-\delta)^{|\mathcal{U}|} 2^{\sum_{a \in \mathcal{U}} n \mathbb{P}_{U}(a)(1-\delta) H(V \mid U=a)(1-\delta)} \\
& =(1-\delta)^{|\mathcal{U}|} 2^{n(1-\delta)^{2} \sum_{a \in \mathcal{U}} \mathbb{P}_{U}(a) H(V \mid U=a)} \\
& =(1-\delta)^{|\mathcal{U}|} 2^{n\left(1-2 \delta+\delta^{2}\right) H(V \mid U)} \\
& \geq(1-\delta)^{|\mathcal{U}|} 2^{n(1-2 \delta) H(V \mid U)} .
\end{aligned}
$$

Hence, there are at least $(1-\delta)^{|\mathcal{U}|} 2^{n H(V \mid U)(1-2 \delta)}$ sequences $v^{n} \in \mathcal{V}^{n}$ satisfying (3). On the other hand, Part (a) shows that the condition (3) implies that $\left(u^{n}, v^{n}\right) \in \mathcal{T}_{3 \delta}^{n}\left(p_{U, V}\right)$. We conclude that there are at least $(1-\delta)^{|\mathcal{U}|} 2^{n H(V \mid U)(1-2 \delta)}$ sequences $v^{n} \in \mathcal{V}^{n}$ satisfying $\left(u^{n}, v^{n}\right) \in \mathcal{T}_{3 \delta}^{n}\left(p_{U, V}\right)$.
(d) If $\left(u^{n}, v^{n}\right) \in \mathcal{T}_{3 \delta}^{n}\left(p_{U, V}\right)$ then $v^{n} \in \mathcal{T}_{3 \delta}^{n}\left(p_{V}\right)$ which implies that $\mathbb{P}_{V^{n}}\left(v^{n}\right) \geq 2^{-n H(V)(1+3 \delta)}$.
(e) We have

$$
\begin{aligned}
\mathbb{P}\left[\left(u^{n}, V^{n}\right) \in \mathcal{T}_{3 \delta}^{n}\left(p_{U, V}\right)\right] & =\sum_{\substack{v^{n} \in \mathcal{V}^{n} ; \\
\left(u^{n}, v^{n}\right) \in \mathcal{T}_{3 \delta}^{n}}} \mathbb{P}_{V^{n}}\left(v^{n}\right) \stackrel{(*)}{\geq} \sum_{\substack{\left.v_{U, V}\right)}} 2^{-n H(V)(1+3 \delta)} \\
& \left.=\mid\left\{v^{n} \in \mathcal{V}^{n}:\left(u^{n}, v^{n}\right) \in \mathcal{T}_{3 \delta}^{n}\left(u^{n}, v^{n}\right) \in \mathcal{T}_{3 \delta}^{n}\right)\right\} \mid \cdot 2^{-n H(V)(1+3 \delta)} \\
& \stackrel{(* *)}{\geq}(1-\delta)^{|\mathcal{U}|} 2^{n H(V \mid U)(1-2 \delta)} 2^{-n H(V)(1+3 \delta)} \\
& =(1-\delta)^{|\mathcal{U}|} 2^{n H(V \mid U)-2 \delta H(V \mid U)-n H(V)-3 \delta H(V)} \\
& \stackrel{(* * *)}{\geq}(1-\delta)^{|\mathcal{U}|} 2^{-n I(U ; V)-2 \delta \log |\mathcal{V}|-3 \delta \log |\mathcal{V}|} \\
& =(1-\delta)^{|\mathcal{U}|} 2^{-n[I(U ; V)+5 \delta \log \mid \mathcal{V}]},
\end{aligned}
$$

where ( $*$ ) follows from Part (d), ( $* *$ ) follows from Part (c) and $(* * *)$ follows from the fact that $I(U ; V)=H(V)-H(V \mid U)$ and $H(V \mid U) \leq H(V) \leq \log |\mathcal{V}|$.

