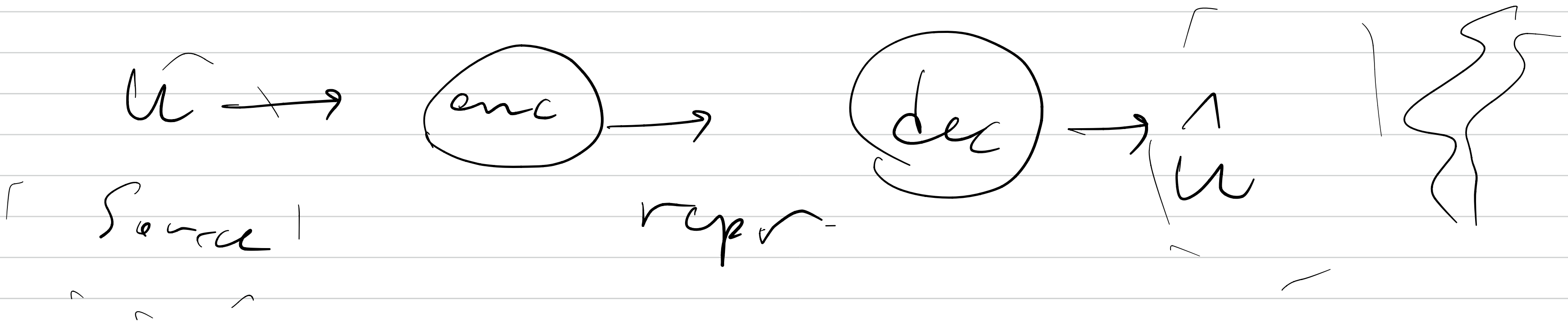
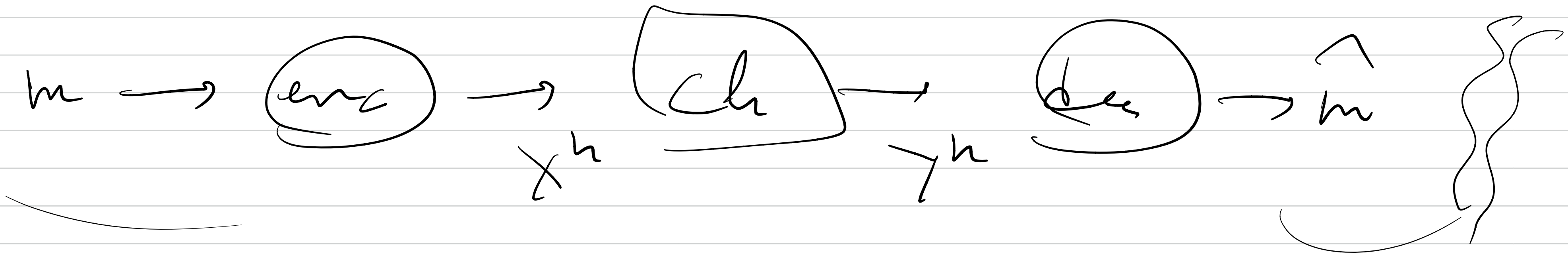
Information Theory & Coding

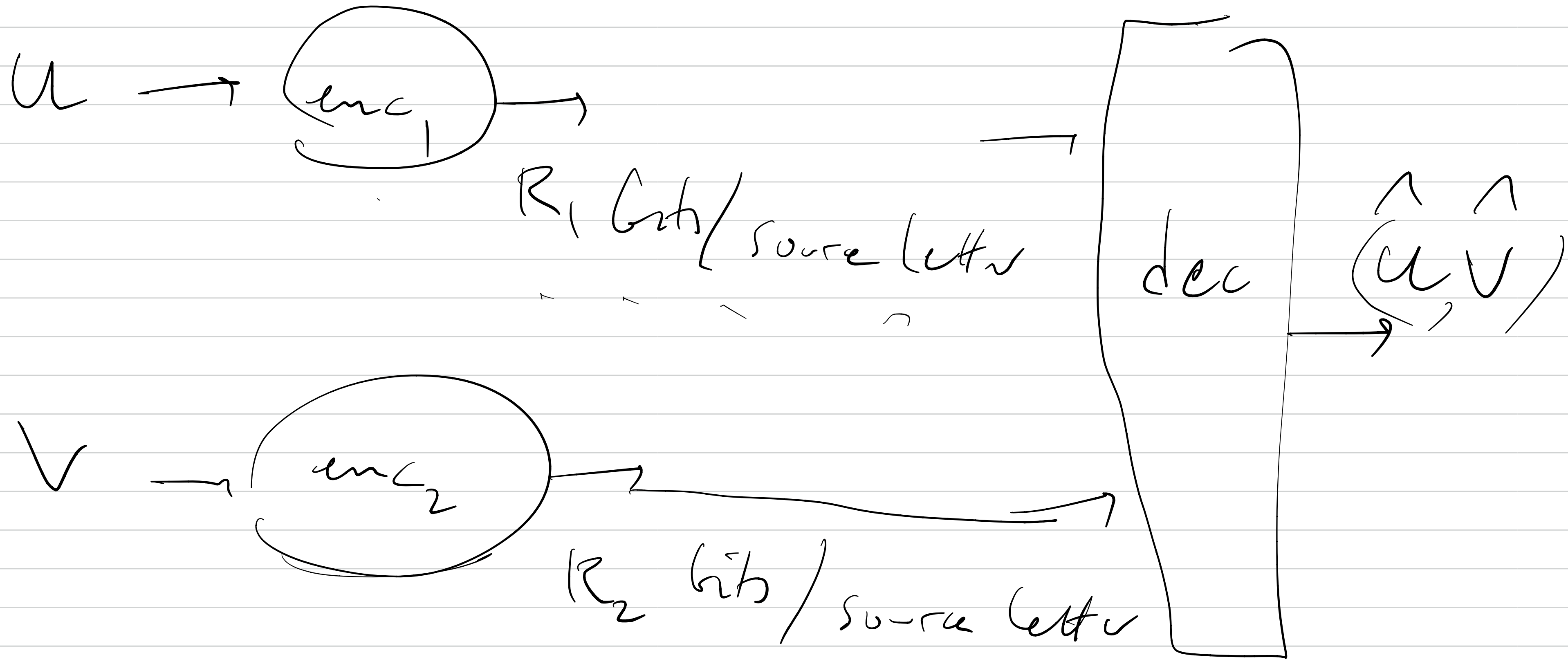
Dec 14th, 2020

Until now:





"Network" extensions:



clearly $\{R_1 > H(U), R_2 > H(V)\}$

will allow perfect recovery.

Q: Can we do better, taking advantage of any dependence between U & V?

Setup: we are given P_{UV}, the model of the source is $(U_1 V_1), (U_2 V_2), \ldots$ i.i.d $\sim P_{UV}$

we have encoders enc_1, enc_2; dec

$$enc_1 : \mathcal{U}^n \to \{1, \ldots, M_1\}, \quad R_1 = \frac{1}{n}\log_2 M_1$$

$$enc_2 : \mathcal{V}^n \to \{1, \ldots M_2\} \quad R_2 = \frac{1}{n}\log M_2$$

$$dec: \{1, .. M_1\} \times \{1 .. M_2\} \to (\mathcal{U}^n \times \mathcal{V}^n)$$

We also have:

$$P_e = \Pr(dec(enc_1(U^n), enc_2(V^n)) \neq (U^n, V^n))$$

error probability.

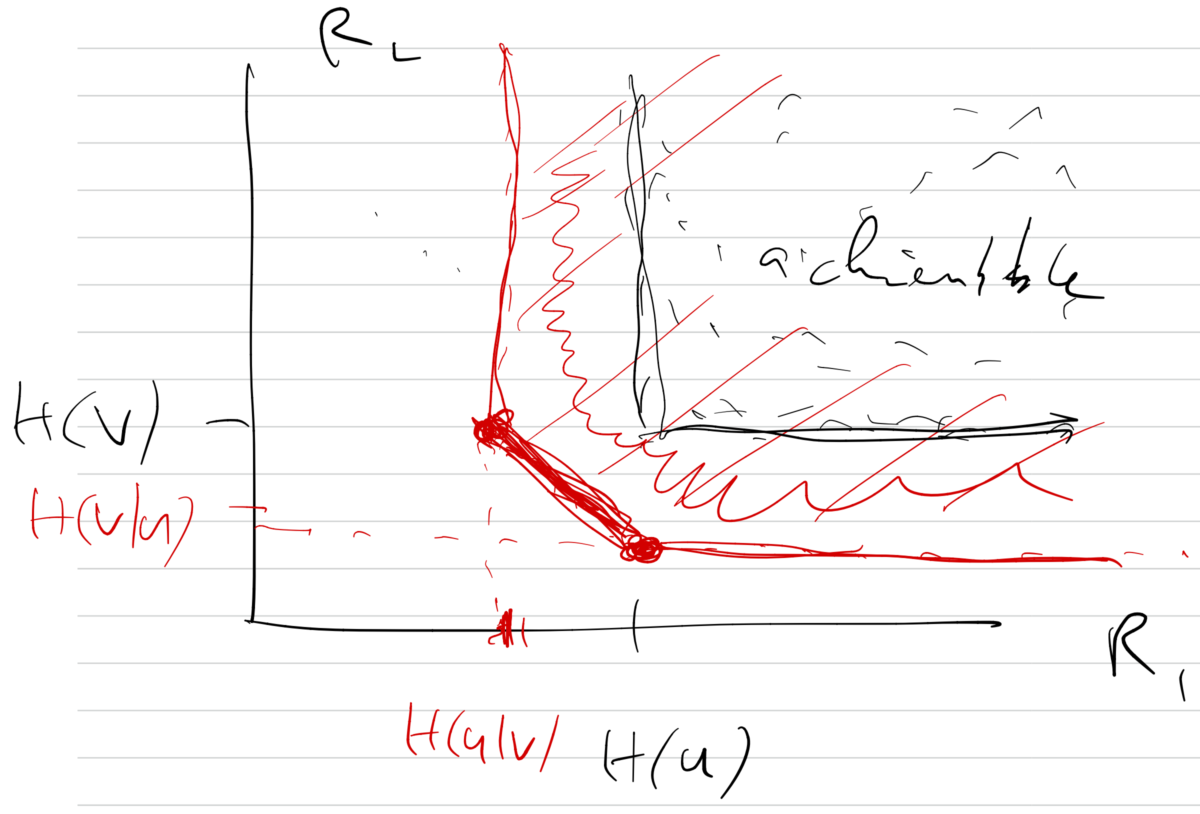
Def: given P_{UV} we will say (R_1, R_2) is achievable if

$\forall \varepsilon > 0 \ \exists \ enc_1, enc_2, dec$ s.t

rate $(enc_1) \leq R_1$, rate$(enc_2) \leq R_2$

$P_{error} < \varepsilon$.

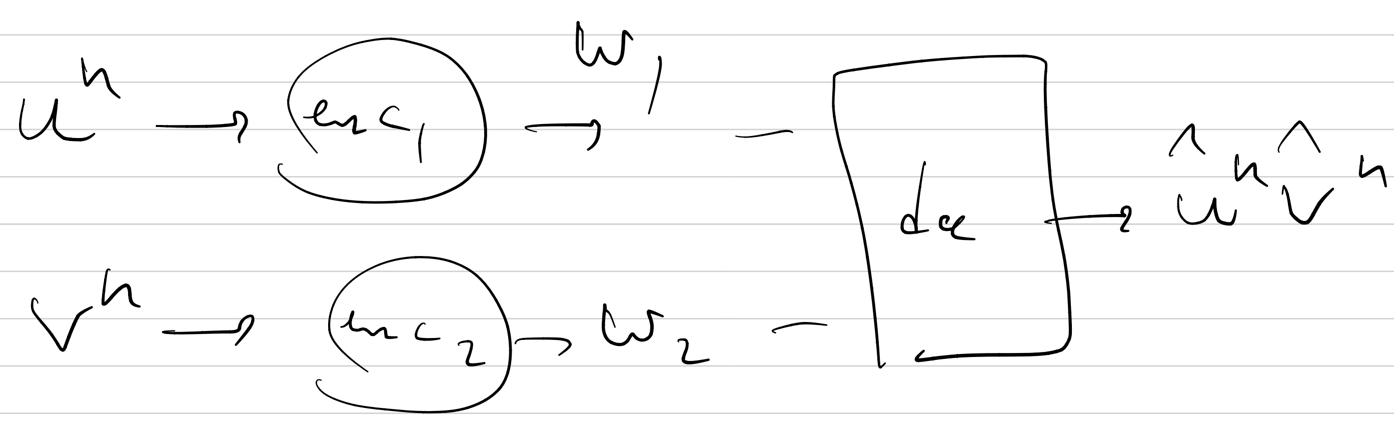
we know: if $R_1 > H(U)$, $R_2 > H(V)$ then (R_1, R_2) is achievable



<u>Thm</u>: (Bad news): if (R_1, R_2) is achievable, then

$$\left(\begin{array}{l} R_1 \geq H(U|V) \\ R_2 \geq H(V|U) \\ R_1 + R_2 \geq H(UV) = H(U) + H(V|U) \\ \qquad\qquad\qquad\quad\; = H(V) + H(U|V) \end{array} \right)$$

<u>Pf</u>: Suppose $(R_1 R_2)$ is achievable, we know that for every $\varepsilon > 0$, we have:



with $\Pr(\hat{U}^n \hat{V}^n \neq U^n V^n) < \varepsilon$

$$W_1 \in \{1, \dots, M_1\} \quad \text{with} \quad \underbrace{M_1 \leq 2^{nR_1}}$$
$$W_2 \in \{1, \dots, M_2\} \quad \text{"} \quad M_2 \leq 2^{nR_2}$$

Recall Fano's inequality :

$$\Pr(U \neq \hat{U}) < \varepsilon \Rightarrow$$

$$H(U|\hat{U}) \leq h_2(\varepsilon) + \varepsilon \log(|\mathcal{U}|-1).$$

because of this

$$H(\underbrace{U^n V^n}|\underbrace{\hat{U}^n \hat{V}^n}) \leq h_2(\varepsilon) + \varepsilon \log(|\mathcal{U}^n||\mathcal{V}^n|)$$

$$= h_2(\varepsilon) + n\varepsilon \log(|\mathcal{U}||\mathcal{V}|)$$

$$= n\varepsilon' \qquad (\varepsilon \approx 0 \Rightarrow \varepsilon' \approx 0).$$

note also, that $\hat{U}^n \hat{V}^n$ is a function of $(W_1 W_2)$

$$\Rightarrow \underbrace{H(U^n V^n | W_1 W_2)}$$

$$= H(U^n V^n | W_1 W_2 \hat{U}^n \hat{V}^n)$$

$$\leq H(U^n V^n | \hat{U}^n \hat{V}^n) \leq \underbrace{n\varepsilon'}$$

$$n R_1 \geq H(W_1) \geq H(W_1 | W_2)$$

$$= \underbrace{H(W_1 U^n V^n | W_2)} - \underbrace{H(U^n V^n | W_1 W_2)}$$

$$\geq H(W_1 U^n V^n | W_2) - n\varepsilon'$$

$$\geq H(U^n V^n | W_2) - n\varepsilon'$$

$$= H(V^n | W_2) + H(U^n | V^n W_2) - n\varepsilon'$$

$$\geq H(U^n | V^n W_2) - n\varepsilon' = H(U^n | V^n) - n\varepsilon'$$

$n R_1 \geq n H(U|V) - n \varepsilon'$

$\Rightarrow R_1 \geq H(U|V) - \varepsilon'$

$\Rightarrow R_1 \geq H(U|V)$ by letting $\varepsilon \downarrow 0$.

by symmetry we also obtain

$R_2 \geq H(V|U)$.

Also note:

$n(R_1 + R_2) \geq H(W_1 W_2)$

$= H(W_1 W_2 U^n V^n) - H(U^n V^n | W_1 W_2)$

$\geq H(W_1 W_2 U^n V^n) - n \varepsilon'$

$\geq H(U^n V^n) - n \varepsilon'$

$= n\left[H(UV) - \varepsilon'\right]$.

$\Rightarrow R_1 + R_2 \geq H(UV) - \varepsilon'$

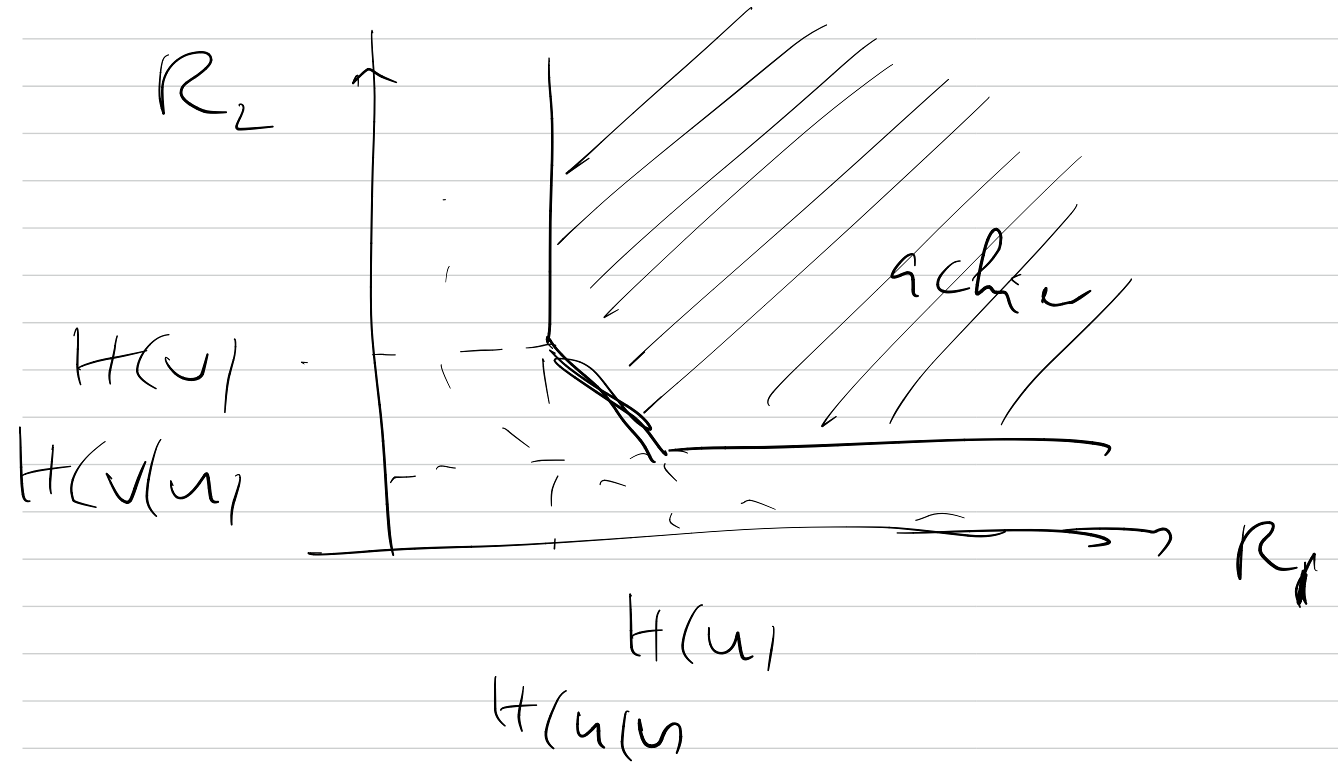
$\Rightarrow R_1 + R_2 \geq H(UV)$. ///

<u>Thm</u>: (Good news): Given P_{UV}. if (R_1, R_2)

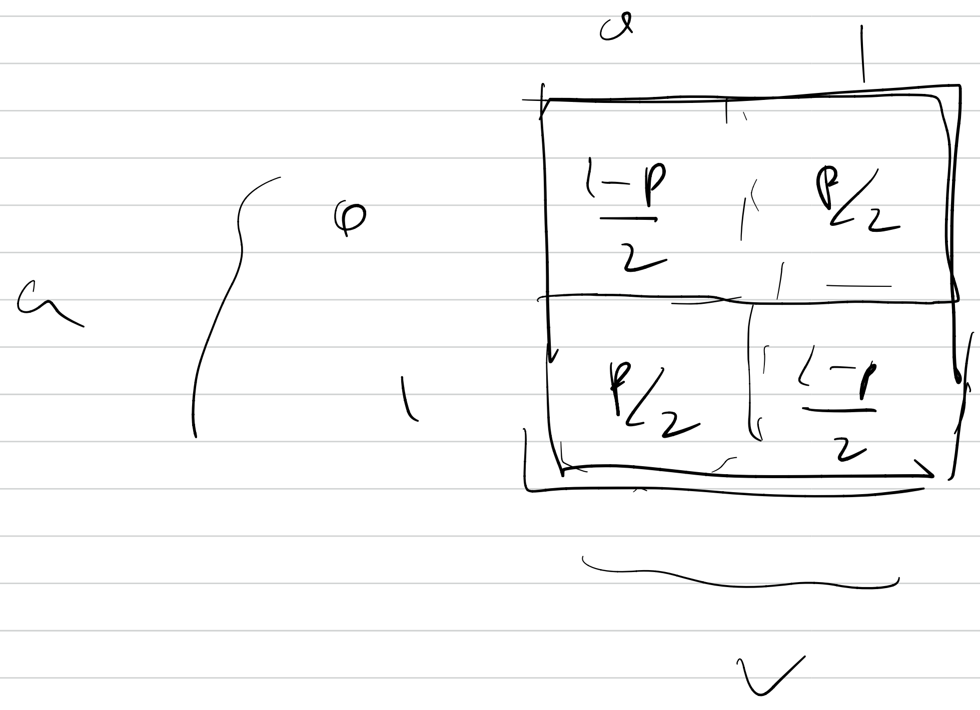
satisfies:

$R_1 > H(U|V)$, $R_2 > H(V|U)$, $R_1 + R_2 > H(UV)$

then (R_1, R_2) is achievable.



Example: U, V are binary:



$p(U=0) = p(U=1) = \frac{1}{2}$

$P(V=0) = p(V=1) = \frac{1}{2}$

$P(U=0 \mid V=1) = p$

$P(U=1 \mid V=0) = p$

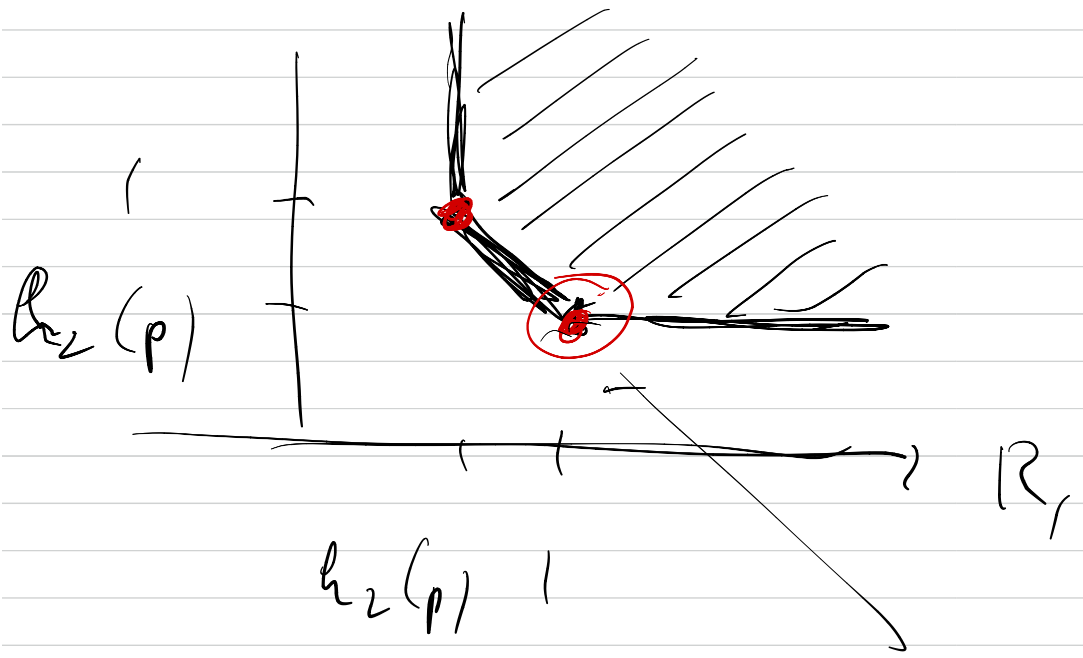
$U = 0, 1$ with equal probability

$V = U \oplus Z$ with $Z \perp\!\!\!\perp U$, $Z = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p. \end{cases}$

$H(U) = H(V) = 1$

$H(U|V) = H(V|U) = h_2(p)$

$H(UV) = 1 + h_2(p).$



$\underline{R_1 = 1}, \quad \underline{R_2 = h_2(p)}.$

<u>Pf of the Slepian–Wolf theorem:</u>

given P_{UV}, R_1, R_2, we will see that if R_1, R_2 satisfy the three inequalities, there exists a (enc_1, enc_2, dec) with small P_e.

Pf by random construction: pick n (large)

for each $u^n \in \mathcal{U}^n$ assign $enc_1(u^n)$ as a uniformly chosen integer in $\{1, \ldots, M_1\}$ $(M_1 = 2^{nR_1})$, independently over $\mathcal{U}^n$.

do the same for enc_2.

for the decoder: pick $\delta > 0$, and consider the typical set $T(p_{UV}, n, \delta)$.

$dec(i,j) = \underbrace{\hat{u}^n, \hat{v}^n}$ if

$1 \le i \le M_1$

$1 \le j \le M_2$

among all $\underbrace{(u^n, v^n) \in T(p_{UV}, n, \delta)}$ $\underbrace{(\hat{u}^n, \hat{v}^n)}$ is the only one with $enc_1(\hat{u}^n) = i$

$enc_2(\hat{v}^n) = j$.

if no such $(\hat{u}^n, \hat{v}^n)$ then set $d(i,j)$ to be an arbitrary value in $\mathcal{U}^n \times \mathcal{V}^n$.

Analysis of this random construction:

$$p(\text{error}) = \Pr((U^n, V^n) \in T \text{ and error}) + \underbrace{\Pr((U^n, V^n) \notin T \text{ and error})}$$

$$\le \Pr((U^n, V^n) \notin T) + \underbrace{\Pr((U^n, V^n) \in T \text{ \& error})}.$$

$$\Pr((U^n, V^n) \in T, \text{ error})$$

$$= \sum_{(u^n, v^n) \in T} \Pr((U^n, V^n) = (u^n, v^n)) \cdot \mathbb{I}\{\text{enc}_1(\tilde{u}^n) = \text{enc}(u^n) \text{ and } \text{enc}_2(\tilde{v}^n) = \text{enc}_2(v^n)\}$$

$$\left\{ \begin{array}{l} (\tilde{u}^n, \tilde{v}^n) \in T \\ (\tilde{u}^n, \tilde{v}^n) \neq (u^n, v^n) \end{array} \right.$$

$$= \sum_{(u^n, v^n) \in T} \left\{ \sum_{\substack{(\tilde{u}^n, \tilde{v}^n) \in T: \\ \tilde{u}^n \neq u^n \\ \tilde{v}^n \neq v^n}} \cdots + \sum_{\substack{(\tilde{u}^n, \tilde{v}^n) \in T \\ \tilde{u}^n = u^n \\ \tilde{v}^n \neq v^n}} \cdots + \sum_{\substack{(\tilde{u}^n, \tilde{v}^n) \in T \\ \tilde{u}^n \neq u^n \\ \tilde{v}^n = v^n}} \cdots \right.$$

In the first sum:

$$\Pr(\text{enc}_1(\tilde{u}^n) = \text{enc}_1(u^n) \ \& \ \text{enc}_2(\tilde{v}^n) = \text{enc}_2(v^n))$$

$$= \frac{1}{M_1} \cdot \frac{1}{M_2}$$

In the second:

$$\Pr\Big(\underbrace{\mathrm{enc}_1(\tilde{u}^n) = \mathrm{enc}_1(u^n)}_{=\frac{1}{M_2}},\ \&\ \mathrm{enc}_2(\tilde{v}^n) = \mathrm{enc}_2(v^n)\Big)$$

In the third sum ... $= \frac{1}{M_1}$

$$E(P_{error}) \leq \Pr((u^n, v^n) \notin T)$$

$$+ \sum_{(u^n,v^n)}' \Pr((u^n, v^n) = (u^n, v^n)) \cdot \left[|T| \cdot (M_1 M_2)^{-1} + \#\{\tilde{v}^n : (u^n, \tilde{v}^n) \in T\} \, M_2^{-1} + \#\{\tilde{u}^n : (\tilde{u}^n, v^n) \in T\} \cdot M_1^{-1} \right]$$

$$|T| \le 2^{nH(UV)(1+\delta)}$$

$$\#\{\tilde{v}^n : (u^n, \tilde{v}^n) \in T\} \le 2^{nH(V|U)(1+2\delta)}$$

when $u^n \in T$

$$\#\{\tilde{u} : (\tilde{u}^n, v^n) \in T\} \le 2^{nH(U|V)(1+2\delta)}$$

for $v^n \in T$

$$\Rightarrow E(P_{error}) \le \Pr((U^n, V^n) \notin T) \searrow 0$$

$$+ 2^{-n(R_1+R_2)} 2^{nH(UV)(1+\delta)} \searrow 0$$

$$+ 2^{-nR_2} 2^{nH(V|U)(1+2\delta)} \searrow 0$$

$$+ 2^{-nR_1} 2^{nH(U|V)(1+2\delta)}. \searrow 0.$$

$$\left.\begin{aligned} R_1+R_2 &> H(UV) \\ R_2 &> H(V|U) \\ R_1 &> H(U|V) \end{aligned}\right\}$$

when these are satisfied we can find small enough $\delta > 0$ s.t. all the four terms above decay to 0 as $n \to \infty$.

//

Regarding the claim: "for a $u^n \in T(p_u, \delta, n)$ the

$$\#\underbrace{\{v^n : (u^n, v^n) \in T(p_{uv}, \delta, n)\}}_{A(u^n)} \leq 2^{nH(V|U)(1+2\delta)}."$$

Consider the following: let V^n to be iid $\sim p_V$.

$$\underline{\Pr((u^n, V^n) \in T(p_{uv}, \delta, n))}$$

$$= \Pr(V^n \in A(u^n)).$$

when $v^n \in A(u^n)$, $v^n \in T(p_V, \delta, n)$.

$$\Rightarrow \Pr(V^n = v^n) = 2^{-nH(V)(1\pm\delta)}.$$

So $\Pr((u^n, V^n) \in T) = \sum\limits_{v^n \in A(u^n)}' 2^{-nH(V)(1\pm\delta)}$.

$$\geq |A(u^n)| \cdot 2^{-nH(V)(1+\delta)}.$$

$$\Rightarrow |A(u^n)| \leq 2^{nH(V)(1+\delta)} \cdot \underbrace{\Pr((u^n, V^n) \in T)}$$

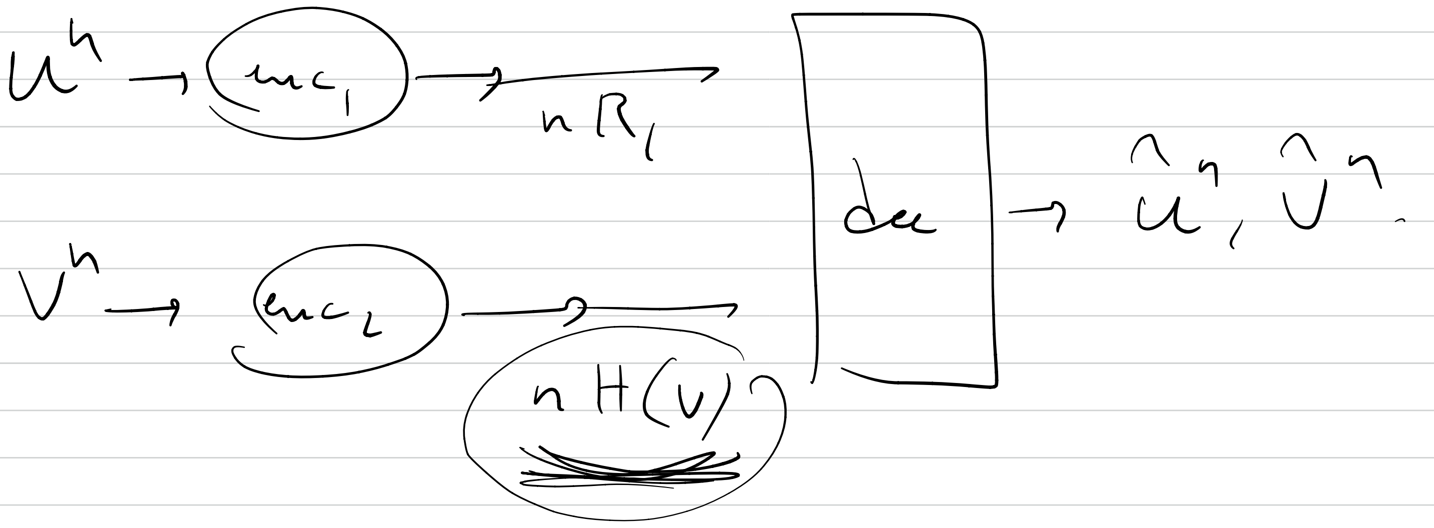
$$\Pr((u^n, V^n) \in T) \leq 2^{-nI(U;V)(1-\delta)}.$$

$$\Rightarrow |A(u^n)| \leq 2^{n\{\underbrace{H(V) - I(U;V)}\} + \delta(\ldots)}$$

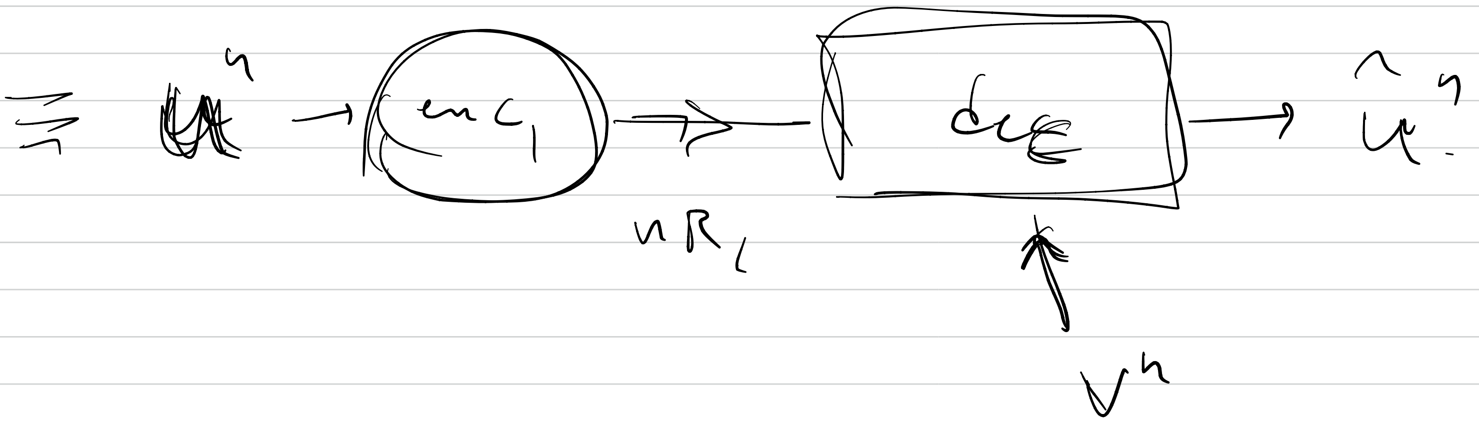
$$= 2^{nH(V|U) + \delta(\ldots)}.$$

//

Special case:

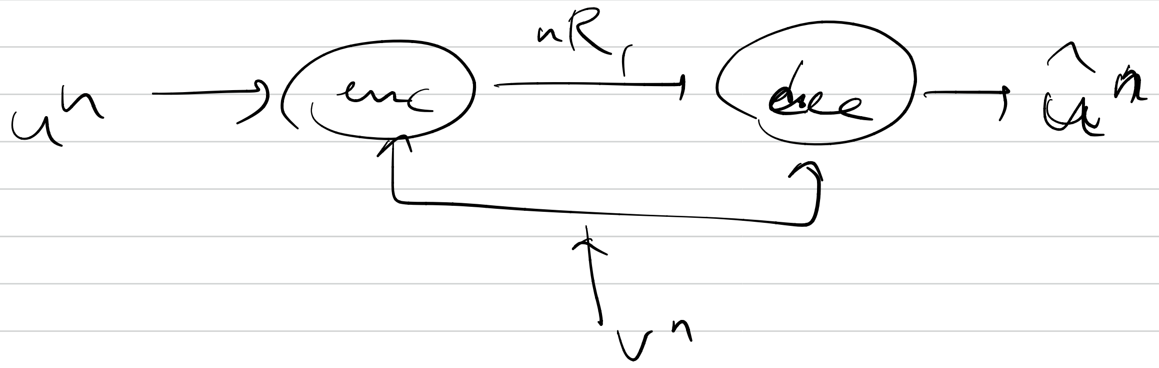


Slepian-Wolf source coding



Required R_1 is $H(U|V)$ even though the enc$_1$ is not aware of V^n.

Alternate problem:



Conditional source coding

required R_1 is again $H(U|V)$.

Tomorrow: Multiple-access channel