Multi-User Information Theory		November 29th, 2011
	Lecture 5	
Lecturer: Haim Permuter		*Scribe: Lior Dikstein*

I. Relay Channel- Decode & Forward via Backward Decoding

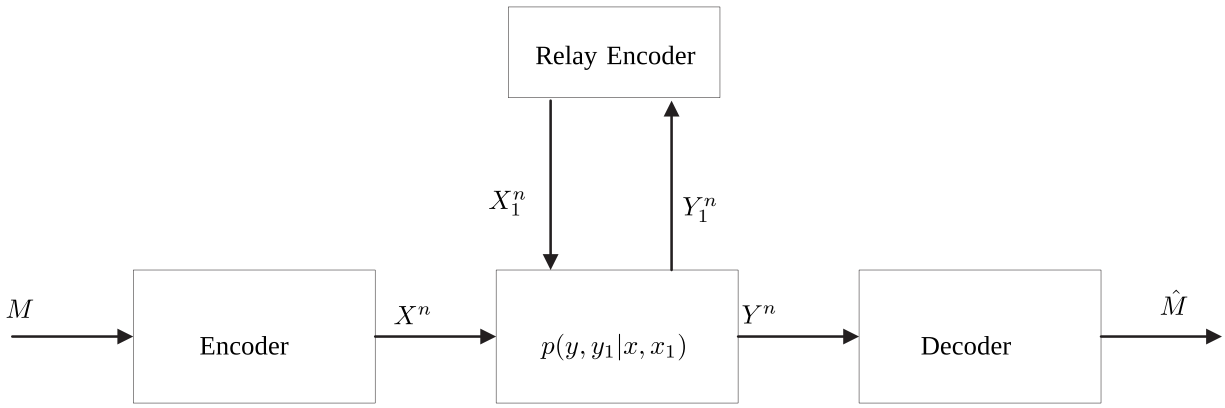

Fig. 1. Relay Channel

Last lecture we saw an upper bound on the capacity of the relay channel, shown in Fig. 1:

$$C < \max_{P(x,x_1)} \min \left\{ I(X;Y,Y_1|X_1), I(X,X_1;Y) \right\}. \tag{1}$$

This lecture we will show the following theorem:

Theorem 1 (Decode & Forward via Backward Decoding Rate) If:

$$R < \min \left\{ I(X;Y,Y_1|X_1), I(X,X_1;Y) \right\}, \tag{2}$$

for some $p(x,x_1)$, then R is achievable.

Proof: We will use *Block Markov Coding*. The idea is to take an N long block, and divide it into B smaller blocks, where $N = nB$ (see Fig 2). Another new method we will use is called *Backward Decoding*. Here the idea is to decode block b, using block $b+1$.

Let us denote m_b as the message send in block b.

Code design (for block b): Fix $p(x,x_1)$ that achieves the lower bound. Randomly and independently generate 2^{nR} sequences $x_1^n(m_{b-1}) \sim p(x_1)$. For each $x_1^n(m_{b-1})$, generate 2^{nR} sequences $x^n(m_b|m_{b-1})$ according to i.i.d. $\sim p(x|x_1)$. The code design is illustrated in Fig 3.

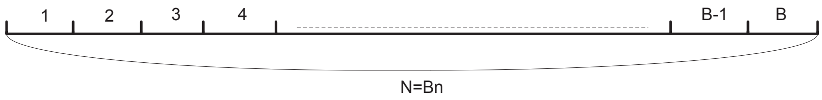

Fig. 2. Superblock Bn separated into B blocks

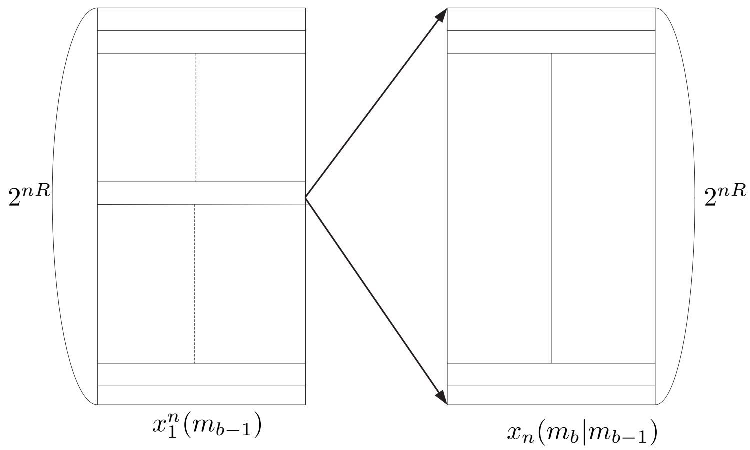

Fig. 3. Coding scheme of Block Markov Coding, Decode & Forward via Backward Decoding. For each codeword $x_1^n(m_{b-1})$, we generate 2^{nR} sequences $x^n(m_b|m_{b-1})$

Encoder: In block b sends $x^n(m_b|m_{b-1})$.

Relay Encoder: At the end of block $b-1$ it decodes message $\hat{m}_{b-1}$ and in block b transmitts the message in block b.

Relay Decoder: At the end of block b the relay needs to decode the message m_b. The relay knows $\hat{m}_{b-1}$ and it looks for:

$$\Big(X^n(m_b, \hat{m}_{b-1}), X_1^n(\hat{m}_{b-1}), Y_1^n\Big) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, Y_1). \tag{3}$$

Decoder: First, we assume the decoder knows $\hat{m}_b$ and wants to decode $\hat{m}_{b-1}$. We also assume that $m_B = 1$.

The decoder waits until the end of the block, and starts decoding backwards. Therefore it looks for:

$$\Big(X^n(\hat{m}_b, m_{b-1}), X_1^n(m_{b-1}), Y^n\Big) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, Y). \tag{4}$$

Analysis of probability of error:

With out loss of generality, we can assume that messages $(\hat{m}_b, \hat{m}_{b-1}) = (1, 1)$ where sent.

An error occurs in the following cases. Define the events:

$$E_1 \quad = \quad \Big\{(X^n(1,1), X_1^n(1), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\Big\}, \tag{5}$$

$$
\begin{aligned}
E_2 &= \left\{(X^n(1,1), X_1^n(1), Y^n) \notin \mathcal{T}_\epsilon^{(n)}\right\}, && (6)\\
E_{3,j} &= \left\{\exists \hat{m}_b = j, j \neq 1 : (X^n(j,1), X_1^n(1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\right\}, && (7)\\
E_{4,j} &= \left\{\exists \hat{m}_{b-1} = j, j \neq 1 : \Big(X^n(1,j), X_1^n(j), Y^n) \in \mathcal{T}_\epsilon^{(n)}\right\}. && (8)\\
& && (9)
\end{aligned}
$$

Then by the union of events bound:

$$
\begin{aligned}
P_e^{(n)} &= Pr(E_1 \cup E_2 \cup E_3 \cup E_4)\\
&\leq P(E_1) + P(E_2) + P(E_3) + P(E_4).
\end{aligned}
$$

Now, let us find the probability of each event:

- For the first two terms, $P(E_1) \to 0$ and $P(E_2) \to 0$ as $n \to \infty$ from L.L.N.
- For the third term we look at the probability that Y_1^n, which is generated according to $\sim p(y_1|x_1)$, is jointly typical with x^n which is generated according to $\sim p(x|x_1)$ where $x_1^n \in \mathcal{T}_\epsilon^{(n)}$. The probability of this event is bounded by:

$$
\begin{aligned}
Pr(\cup_j E_{3,j}) &\leq \sum_{j=2}^{2^{nR}} P(E_{3,j})\\
&\leq \sum_{j=2}^{2^{nR}} 2^{-nI(X;Y_1|X_1)}\\
&= 2^{nR} 2^{-nI(X;Y_1|X_1)}.
\end{aligned}
$$

- For the forth term we look at the probability that Y^n, which is generated according to $\sim p(y)$, is jointly typical with x^n which is generated according to $\sim p(x|x_1)$ and x_1^n which is generated according to $\sim p(x_1)$. The probability of this event is bounded by:

$$
\begin{aligned}
Pr(\cup_j E_{4,j}) &\leq \sum_{j=2}^{2^{nR}} P(E_{4,j})\\
&\leq \sum_{j=2}^{2^{nR}} 2^{-nI(X,X_1;Y)}\\
&= 2^{nR} 2^{-nI(X,X_1;Y)}.
\end{aligned}
$$

Now, to complete the proof, the following lemma will enable us to bound the probability of error of the super-block nB by bounding the probability of error of each block.

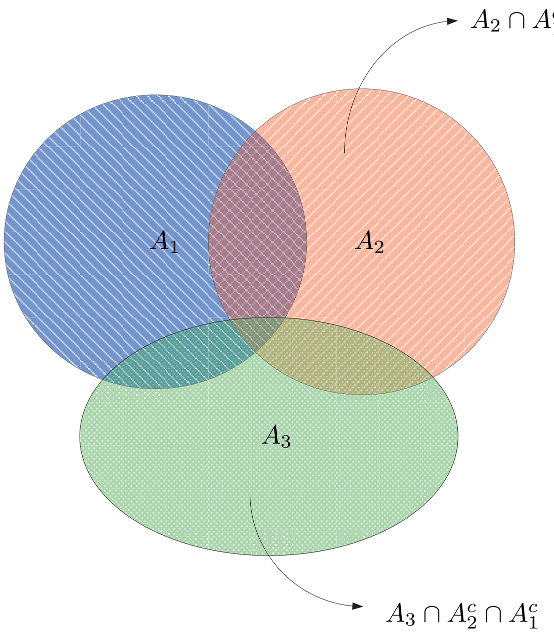

Fig. 4. Graphical display of Lemma 1. The red only area is $A_2 \cap A_1^c$ and the green only area is $A_3 \cap A_2^c \cap A_1^c$

Lemma 1 Let $\{A_j\}_{j=1}^J$ be a set of events and let A_j^c denotes the complement of the event A_j. Then

$$P(\bigcup_{j=1}^{J} A_j) \leq \sum_{j=1}^{n} P(A_j | \bigcap_{i=1}^{j-1} A_i^c) = \sum_{j=1}^{n} P(A_j | A_1^c, A_2^c, ..., A_{j-1}^c). \tag{10}$$

Proof: For simplicity let us assume that $J = 3$. In a straightforward manner the proof extends to any number of sets J. For any three sets of events A_1, A_2, A_3 we have

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P\left(A_1 \cup (A_2 \cap A_1^c) \cup (A_3 \cap A_1^c \cap A_2^c)\right) \\ &= P(A_1) + P(A_2 \cap A_1^c) + P(A_3 \cap A_1^c \cap A_2^c) \\ &\leq P(A_1) + \frac{P(A_2 \cap A_1^c)}{P(A_1^c)} + \frac{P(A_3 \cap A_1^c \cap A_2^c)}{P(A_1^c \cap A_2^c)} \\ &= P(A_1) + P(A_2 | A_1^c) + P(A_3 | A_1^c \cap A_2^c) \\ &= P(A_1) + P(A_2 | A_1^c) + P(A_3 | A_1^c, A_2^c). \end{aligned} \tag{11}$$

Fig. 4 illustrates the lemma for $J = 3$. ■

Using Lemma 1 we bound the probability of error in the supper block Bn by the sum of the probability of having an error in each block b given that in previous blocks $(b+1,...,B)$ the messages were decoded correctly.

Let us bound the probability that for some b. Using Lemma 1 it suffices to show that the probability of error-decoding in each block b goes to zero, assuming that all previous messages in block $(1,2,...,b-1)$ were decoded correctly. ■

II. PARTIAL DECODE & FORWARD

In this coding scheme the relay will decode only part of the message. This provides a better lower bound on capacity.

Theorem 2 (Partial Decode & Forward Rate) If:

$$R < \min\Big\{I(U;Y_1|X_1) + I(X;Y|X_1,U), I(X,X_1;Y)\Big\}, \tag{12}$$

for some $p(x,x_1,u)$, then R is achievable.

Where U is an auxiliary random variable.

Note: If we substitute $U = X$, the above lower bound reduces the decode-and forward lower bound, and if we substitute $U = \emptyset$, it reduces to the direct transmission lower bound.

Outline of achievability: Again, we will use *Block Markov Coding*. Divide the N long block into B smaller blocks, where $N = nB$, as illustrated in Fig 5.

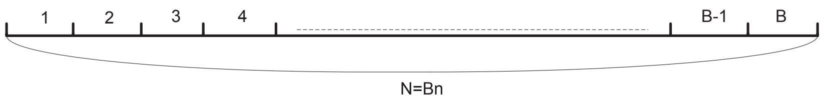

Fig. 5. Separation into B blocks

Proof: Split the message m into two independent messages (m_b', m_b'') with rates R' and R''.Thus $R = R' + R''$.

$$\begin{aligned} m_b &\in \{1,2,...,2^{nR}\}, \\ m_b' &\in \{1,2,...,2^{nR'}\}, \\ m_b'' &\in \{1,2,...,2^{nR''}\}. \end{aligned} \tag{13}$$

$$\tag{14}$$

The idea is to have the relay decode only m_b'.

Code design (for block b): Fix $p(x,x_1)$ that achieves the lower bound. The Relay decodes m'_b so randomly and independently generate 2^{nR} sequences $x_1^n(m'_{b-1}) \sim p(x_1)$. For each $x_1^n(m'_{b-1})$, generate $2^{nR'}$ sequences $u^n(m'_b|m'_{b-1})$ according to i.i.d. $\sim p(u|x_1)$. Now, for every (m'_b, m''_b) generate $2^{nR''}$ sequences $x^n(m''_b|m'_b, m'_{b-1})$ according to i.i.d. $\sim p(x|u,x_1)$.The code design is illustrated in Fig 6.

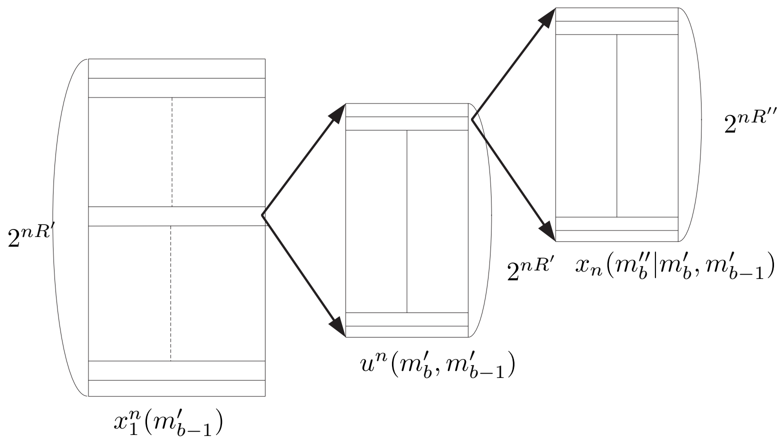

Fig. 6. Coding scheme of Block Markov Coding, Partial Decode & Forward. For each $x_1^n(m'_{b-1})$, we generate $2^{nR'}$ sequences $u^n(m'_b|m'_{b-1})$ and for every (m'_b, m''_b) we generate $2^{nR''}$ sequences $x^n(m''_b|m'_b, m'_{b-1})$

Encoder: Sends $x^n(m''_b|m'_b, m'_{b-1})$.

Relay Decoder: At the end of block b the relay needs to decode the message m'_b. The relay knows $\hat{m}'_{b-1}$ and it looks for:

$$\left(U^n(m'_b, \hat{m}'_{b-1}), X_1^n(\hat{m}'_{b-1}), Y_1^n\right) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, Y_1). \tag{15}$$

Decoder: First, we assume the decoder knows $\hat{m}'_{b+1}$ and wants to decode $\hat{m}'_b$ and $\hat{m}''_b$. We also assume that $m_B = 1$.

The decoder waits until the end of the block, and starts decoding backwards. Therefore it looks for:

$$\left(U^n(\hat{m}'_{b+1}, m'_b), X^n(\hat{m}''_{b+1}|m'_{b+1}, m'_b), X_1^n(m'_b), Y^n\right) \in \mathcal{T}_\epsilon^{(n)}(X, X_1, Y). \tag{16}$$

Analysis of probability of error:

With out loss of generality, we can assume that messages $(\hat{m}'_b, \hat{m}''_{b+1} = (1,1)$ where sent.

An error occurs in the following cases. Define the events:

$$E_1 = \left\{U^n(1,1), X_1^n(1), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\right\}, \tag{17}$$

$$E_2 = \left\{U^n(1,1), X^n(1,1), X_1^n(1), Y^n) \notin \mathcal{T}_\epsilon^{(n)}\right\}, \tag{18}$$

$$E_{3,j} = \left\{\exists \hat{m}'_b = j, j \neq 1 : (U^n(\hat{m}'_b, 1), X_1^n(1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\right\}, \tag{19}$$

$$E_{4,j} = \left\{\exists \hat{m}'_b = j, j \neq 1 : (U^n(1, m'_b), X^n(1|1, m'_b), X_1^n(m'_b), Y^n) \in \mathcal{T}_\epsilon^{(n)}\right\}, \quad (20)$$

$$E_{5,j} = \left\{\exists \hat{m}''_{b+1} = j, j \neq 1 : (U^n(1, 1), X^n(\hat{m}''_{b+1}|1, 1), X_1^n(1), Y^n) \in \mathcal{T}_\epsilon^{(n)}\right\}, \quad (21)$$

$$E_{6,j,i} = \left\{\exists \hat{m}'_b \neq 1 = j, \hat{m}''_{b+1} = i, i, j \neq 1 : (U^n(1, m'_b), X^n(\hat{m}''_{b+1}|1, m'_b), X_1^n(m'_b), Y^n) \in \mathcal{T}_\epsilon^{(n)}\right\}. \quad (22)$$

Then by the union of events bound:

$$\begin{aligned} P_e^{(n)} &= Pr(E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6) \\ &\leq P(E_1) + P(E_2) + P(E_3) + P(E_4) + P(E_5) + P(E_6). \end{aligned}$$

Now, let us find the probability of each event:

- For the first two terms, $P(E_1) \to 0$ and $P(E_2) \to 0$ as $n \to \infty$ from L.L.N.

- For the third term we look at the probability that Y_1^n, which is generated according to $\sim p(y_1|x_1)$, is jointly typical with u^n which is generated according to $\sim p(u|x_1)$ where $x_1^n \in \mathcal{T}_\epsilon^{(n)}$. The probability of this event is bounded by:

$$\begin{aligned} Pr(\cup_j E_{3,j}) &\leq \sum_{j=2}^{2^{nR'}} P(E_{3,j}) \\ &\leq \sum_{j=2}^{2^{nR'}} 2^{-nI(U;Y_1|X_1)} \\ &= 2^{nR'} 2^{-nI(U;Y_1|X_1)}. \end{aligned}$$

- For the forth term we look at the probability that Y^n, which is generated according to $\sim p(y)$, is jointly typical with x^n which is generated according to $\sim p(x)$, u^n which is generated according to $\sim p(u)$, and x_1^n which is generated according to $\sim p(x_1)$. The probability of this event is bounded by:

$$\begin{aligned} Pr(\cup_j E_{4,j}) &\leq \sum_{j=2}^{2^{nR'}} P(E_{4,j}) \\ &\leq \sum_{j=2}^{2^{nR'}} 2^{-nI(U,X,X_1;Y)} \\ &= 2^{nR'} 2^{-nI(U,X,X_1;Y)}. \end{aligned}$$

- For the fifth term we look at the probability that Y^n, which is generated according to $\sim p(y|x_1)$, is jointly typical with x^n which is generated according to $\sim p(x|x_1, u)$ where $(u^n, x_1^n) \in \mathcal{T}_\epsilon^{(n)}$. The probability of this event is bounded by:

$$\begin{aligned} Pr(\cup_j E_{5,j}) &\leq \sum_{j=2}^{2^{nR''}} P(E_{5,j}) \\ &\leq \sum_{j=2}^{2^{nR''}} 2^{-nI(X;Y|U,X_1)} \\ &= 2^{nR''} 2^{-nI(X;Y|U,X_1)}. \end{aligned}$$

- For the last term we look at the probability that Y^n, which is generated according to $\sim p(y)$, is jointly typical with x^n which is generated according to $\sim p(x)$, u^n which is generated according to $\sim p(u)$, and x_1^n which is generated according to $\sim p(x_1)$. The probability of this event is bounded by:

$$\begin{aligned} Pr(\cup_j \cup_i E_{6,j,i}) &\leq \sum_{j=2}^{2^{nR'}} \sum_{i=2}^{2^{nR''}} P(E_{6,j,i}) \\ &\leq \sum_{j=2}^{2^{nR'}} \sum_{i'=2}^{2^{nR''}} 2^{-nI(U,X,X_1;Y)} \\ &= 2^{nR'} 2^{nR''} 2^{-nI(U,X,X_1;Y)}. \end{aligned}$$

Therefore we have the bounds:

$$\begin{aligned} R' &\leq I(U;Y|X_1), \\ R' &\leq I(U,X,X_1;Y), \\ R &\leq I(U,X,X_1;Y), \\ R'' &\leq I(X;Y|U,X_1). \end{aligned}$$

However, we want to find the bound on R alone. To do so we will use *Fourier-Mutskin elimination* which is a mathematical algorithm for eliminating variables from a system of linear inequalities.

Example 1

$$x_1 \leq 2 + x_2, \tag{23}$$

$$x_1 \geq 3 - x_2. \tag{24}$$

For each x_2 there exists x_1 such that (23) and (24) are satisfied if:

$$3 - x_2 \leq 2 + x_2, \tag{25}$$

so we can reach an inequality for x_2 alone:

$$x_2 \geq \frac{1}{2}. \tag{26}$$

In our setting, where $R'' = R - R'$ we have

$$R' \leq I(U;Y|X_1),$$
$$R' \leq R - I(X;Y|U,X_1).$$

So using the Fourier-Mutskin elimination we get

$$I(U;Y|X_1) \geq R - I(X;Y|U,X_1),$$

therfore,

$$R \leq I(U;Y|X_1) + I(X;Y|U,X_1).$$

■

References

[1] T. M. Cover and J. A. Thomas, ‘*Elements of Information Theory.*‘. Wiley, New York, 2nd edition 2006

[2] Abbas El Gamal, Young-Han Kim, ’*Lecture Notes on Network Information Theory*’.