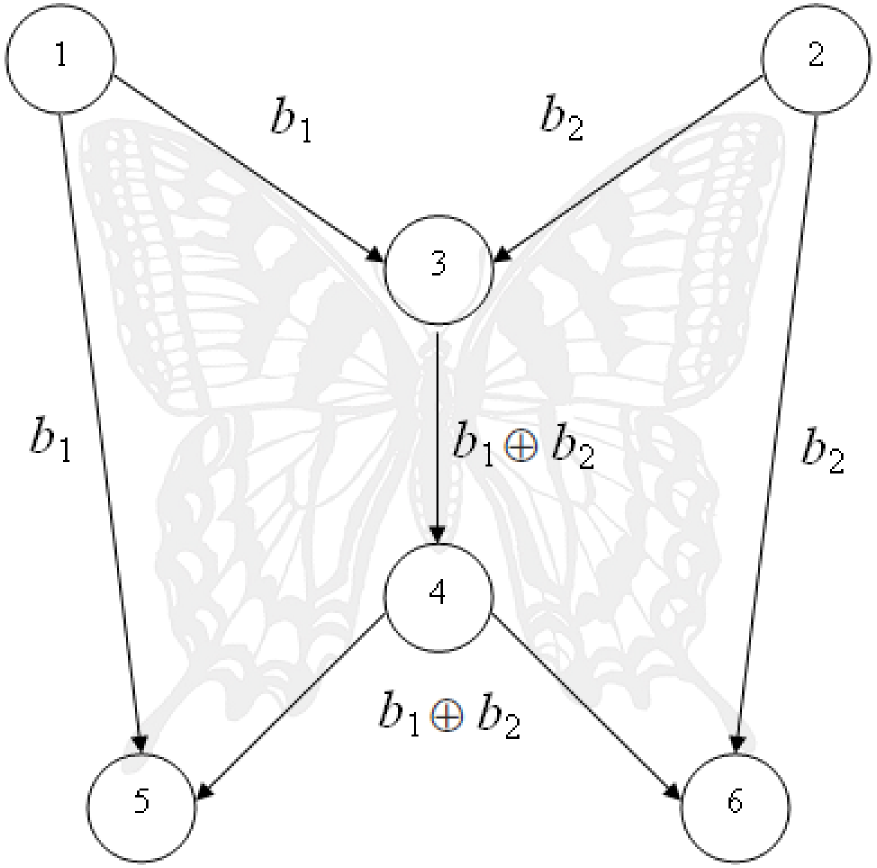
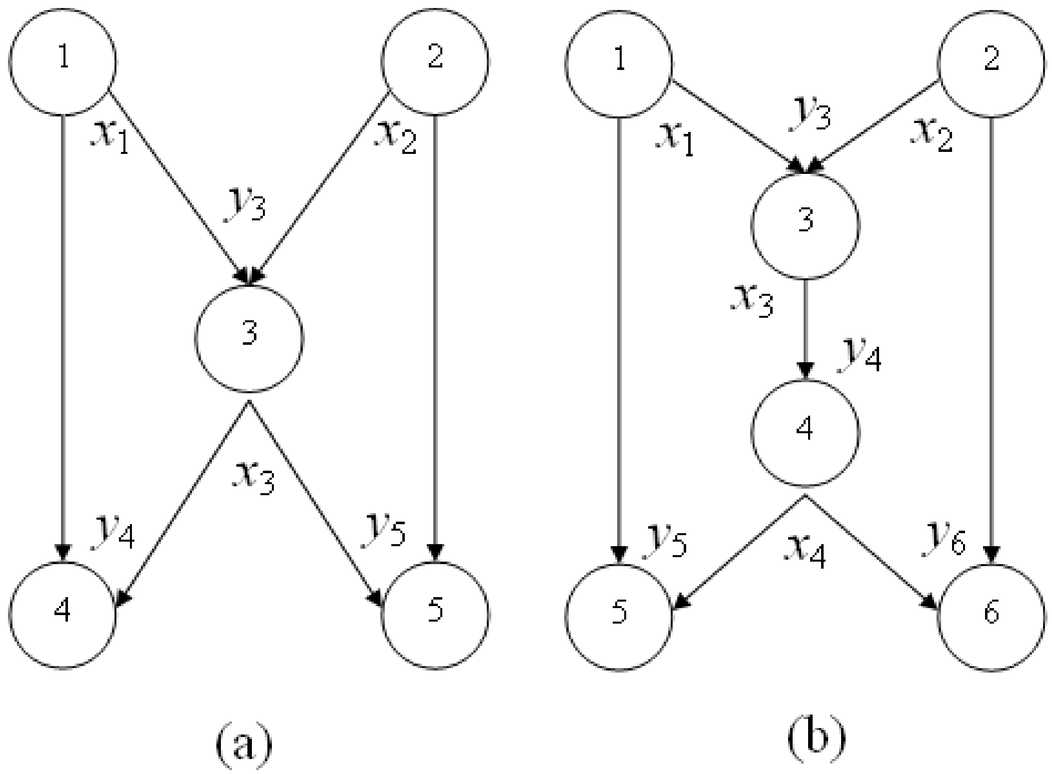
Butterfly Channel

Liang Chen
University of Maryland

The Butterfly Network



The Butterfly Channels



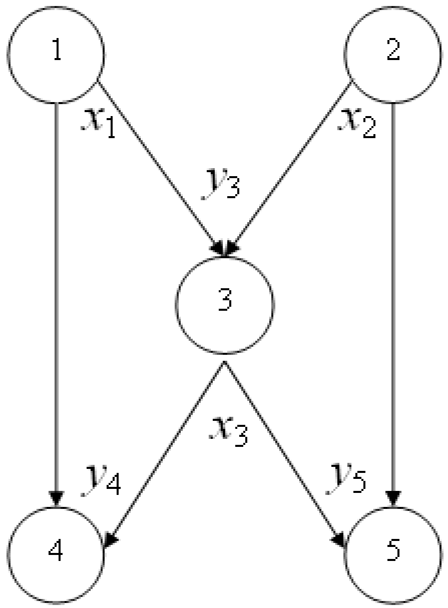
(a) Type I (b) Type II

System Model of Butterfly Channel

Definition: The butterfly channel (Type I) consists of:

- two independent channel input alphabets $\mathcal{X}_1$, $\mathcal{X}_2$
- a relay input alphabet $\mathcal{X}_3$
- a relay output alphabet $\mathcal{Y}_3$
- two channel output alphabets $\mathcal{Y}_4$, $\mathcal{Y}_5$
- a probability transition function

$$p(y_3, y_4, y_5 \mid x_1, x_2, x_3)$$



Channel Code for Butterfly Channel

Definition: A $(2^{nR_1}, 2^{nR_2}, n)$ code consists of:

- *Two message sets*: $\mathcal{W}_k = \{1, 2, \ldots, 2^{nR_k}\}, k = 1, 2$
- *Two encoders*: at node k, $\mathcal{W}_k \to \mathcal{X}_k^n$, which maps each message $W_k \in \mathcal{W}_k$ to a codeword $\mathbf{x}_k \in \mathcal{X}_k^n, k = 1, 2.$
- *Three decoders*: at node k, $\mathcal{Y}_k^n \to \mathcal{W}_1 \times \mathcal{W}_2, k = 3, 4, 5,$ which maps a received sequence y_k^n to a message pair $(\hat{W}_1, \hat{W}_2)$.

Network Coding at Node 3

- Message-by-message Network Coding
- Bit-by-bit Network Coding

Message-by-message Network Coding

- In block j, node 3 sends a message Z to both node 4 and node 5.

$$z_j = f(w_{1,j-1}, w_{2,j-1})$$

f: function of network coding

Z: outcome of network coding

- After computing Z, node 3 encodes it via $\mathcal{Z} \to \mathcal{X}_3^n$ and forwards the signal $\mathbf{x}_3$ (*n* symbols) to node 4 and node 5.

Message-by-message Network Coding

- In block j, node 3 sends a message Z to both node 4 and node 5.

$$z_j = f(w_{1,j-1}, w_{2,j-1})$$

f: function of network coding

Z: outcome of network coding

- After computing Z, node 3 encodes it via $\mathcal{Z} \rightarrow \mathcal{X}_3^n$ and forwards the signal $\mathbf{x}_3$ (n symbols) to node 4 and node 5.

Message-by-message Network Coding (cont.)

- Given $w_{1,j-1}$, $z_j = f_1(w_{2,j-1})$.
  Given $w_{2,j-1}$, $z_j = f_2(w_{1,j-1})$.
- If both functions f_1 and f_2 are bijective, at node 4, after having known $w_{1,j-1}$,
  $$w_{2,j-1} = f_1^{-1}(z_j),$$
  at node 5, after having known $w_{2,j-1}$,
  $$w_{1,j-1} = f_2^{-1}(z_j).$$

Bit-by-bit Network Coding

- Node 3 generates its relayed signal $\mathbf{x}_3$ bit-by-bit through a set of relay function such that

$$x_{3,i} = g_i(y_{3,1}, y_{3,2}, ..., y_{3,i-1}),\ 1 \le i \le n$$

- In any block, node 3 sends the message correlated with the messages sent from node 1 and node 2 in the same block.

Bit-by-bit Network Coding

- Node 3 generates its relayed signal $\mathbf{x}_3$ bit-by-bit through a set of relay function such that

$$x_{3,i} = g_i(y_{3,1}, y_{3,2}, \ldots, y_{3,i-1}),\ 1 \le i \le n$$

- In any block, node 3 sends the message correlated with the messages sent from node 1 and node 2 in the same block.

Comparison of Two Network Coding Schemes

- Message-by-message Network Coding

$$\mathbf{y}_3(j-1) \Rightarrow \begin{matrix} \hat{w}_{1,j-1} \\ \hat{w}_{2,j-1} \end{matrix} \Rightarrow z_j \Rightarrow \mathbf{x}_3(z_j)$$

Comparison of Two Network Coding Schemes

- Message-by-message Network Coding

$$\mathbf{y}_3(j-1) \Rightarrow \begin{matrix}\hat{w}_{1,j-1} \\ \hat{w}_{2,j-1}\end{matrix} \Rightarrow z_j \Rightarrow \mathbf{x}_3(z_j)$$

- Bit-by-bit Network Coding

  No need to decode $\hat{w}_1$ and $\hat{w}_2$ at node 3.

  Directly generate $\mathbf{x}_3$ through

$$x_{3,i} = g_i(y_{3,1}, y_{3,2}, \ldots, y_{3,i-1}), \; 1 \le i \le n.$$

Capacity Region with Message-by-message Network Coding

- The capacity of the butterfly channel (Type I) with message-by-message network coding is the closure of the convex hull of all (R_1, R_2) satisfying

$$R_1 < \min\{I(X_1;Y_3 \mid X_2, X_3), I(X_1;Y_4 \mid X_3), I(X_3;Y_5 \mid X_2)\}$$

$$R_2 < \min\{I(X_2;Y_3 \mid X_1, X_3), I(X_2;Y_5 \mid X_3), I(X_3;Y_4 \mid X_1)\}$$

$$R_1 + R_2 < \min\{I(X_1, X_2;Y_3 \mid X_3), I(X_1, X_3;Y_4), I(X_2, X_3;Y_5)\}$$

for some product distribution $p(x_1)p(x_2)p(x_3)$.

Random Codebook Generation

- 1. Generate 2^{nR_1} i.i.d. $\mathbf{x}_1$ each with $\prod_{i=1}^{n} p(x_{1,i})$.

  Index each $\mathbf{x}_1(w_1), w_1 \in [1, 2^{nR_1}]$.

- 2. Generate 2^{nR_2} i.i.d. $\mathbf{x}_2$ each with $\prod_{i=1}^{n} p(x_{2,i})$.

  Index each $\mathbf{x}_2(w_2), w_2 \in [1, 2^{nR_2}]$.

- 3. Generate 2^{nR_z} i.i.d. $\mathbf{x}_3$ each with $\prod_{i=1}^{n} p(x_{3,i})$.

  Index each $\mathbf{x}_3(z), z \in [1, 2^{nR_z}]$.

Encoding

- In block j,

node 1: $w_{1,j}$

Encoding

- In block j,

node 1: $w_{1,j}$ ⇨ $\mathbf{x}_1(w_{1,j})$

Encoding

- In block j,

node 1: $w_{1,j} \Rightarrow \mathbf{x}_1(w_{1,j})$

node 2: $w_{2,j} \Rightarrow \mathbf{x}_2(w_{2,j})$

Encoding

- In block j,

node 1: $w_{1,j} \Rightarrow \mathbf{x}_1(w_{1,j})$

node 2: $w_{2,j} \Rightarrow \mathbf{x}_2(w_{2,j})$

- At the beginning of block j, node 3:

Encoding

- In block j,

node 1: $w_{1,j} \Rightarrow \mathbf{x}_1(w_{1,j})$

node 2: $w_{2,j} \Rightarrow \mathbf{x}_2(w_{2,j})$

- At the beginning of block j, node 3:

$\hat{w}_{1,j-1}$

$\hat{w}_{2,j-1}$

Encoding

- In block j,

node 1: $w_{1,j} \Rightarrow \mathbf{x}_1(w_{1,j})$

node 2: $w_{2,j} \Rightarrow \mathbf{x}_2(w_{2,j})$

- At the beginning of block j, node 3:

$$\hat{w}_{1,j-1}, \hat{w}_{2,j-1} \Rightarrow z_j = f(\hat{w}_{1,j-1}, \hat{w}_{2,j-1})$$

Encoding

- In block j,

  node 1: $w_{1,j} \Rightarrow \mathbf{x}_1(w_{1,j})$

  node 2: $w_{2,j} \Rightarrow \mathbf{x}_2(w_{2,j})$

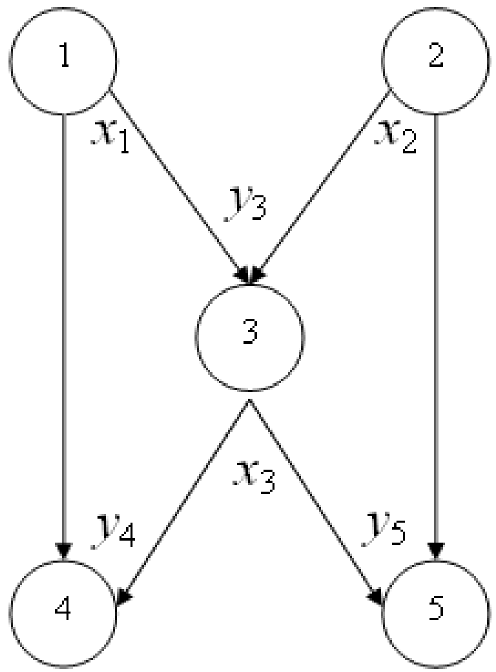
- At the beginning of block j, node 3:

$$\hat{w}_{1,j-1},\ \hat{w}_{2,j-1} \Rightarrow z_j = f(\hat{w}_{1,j-1}, \hat{w}_{2,j-1}) \Rightarrow \mathbf{x}_3(z_j)$$

Decoding

- Node 3:

$$\mathbf{y}_3(j) \Rightarrow \begin{matrix} \hat{w}_{1,j} \\ \hat{w}_{2,j} \end{matrix}$$



Decoding

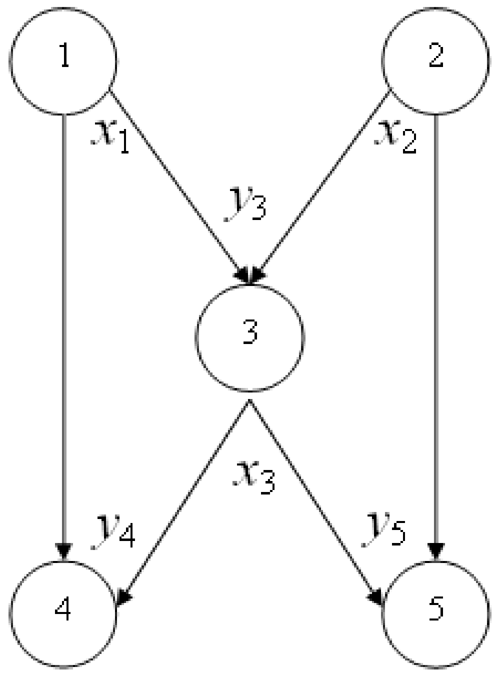
- Node 3:

$$\mathbf{y}_3(j) \Rightarrow \begin{matrix} \hat{w}_{1,j} \\ \hat{w}_{2,j} \end{matrix}$$

$$R_1 < I(X_1;Y_3 \mid X_2, X_3)$$

$$R_2 < I(X_2;Y_3 \mid X_1, X_3)$$

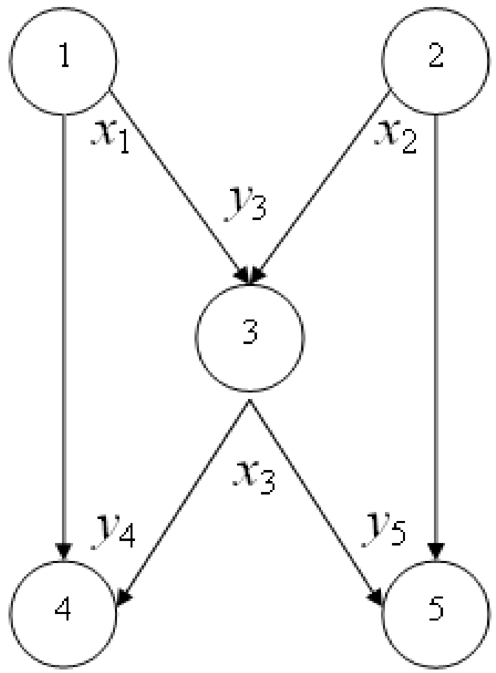
$$R_1 + R_2 < I(X_1, X_2;Y_3 \mid X_3)$$



Decoding (cont.)

- Node 4:

$$\mathbf{y}_4(j) \Rightarrow \begin{matrix} \hat{w}_{1,j} \\ \hat{z}_j \end{matrix}$$



$$R_1 < I(X_1;Y_4 \mid X_3)$$

$$R_z < I(X_3;Y_4 \mid X_1)$$

$$R_1 + R_z < I(X_1, X_3;Y_4)$$

Decoding (cont.)

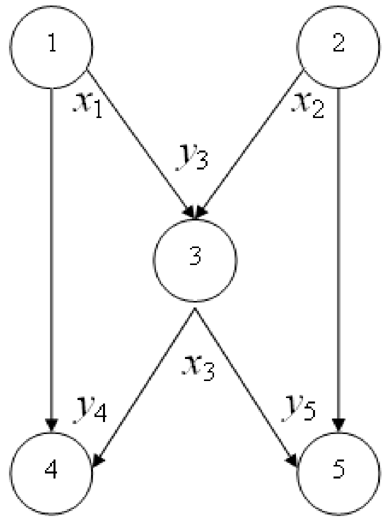
- Having known $w_{1,j-1}$, for each $\hat{z}_j$, node 4 can derive a $\hat{w}_{2,j-1}$ through $\hat{w}_{2,j-1} = f_1^{-1}(\hat{z}_j)$.

$$
\begin{aligned}
R_1 &< I(X_1;Y_4 \mid X_3) \\
R_z &< I(X_3;Y_4 \mid X_1) \\
R_1 + R_z &< I(X_1, X_3;Y_4)
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
R_1 &< I(X_1;Y_4 \mid X_3) \\
R_2 &< I(X_3;Y_4 \mid X_1) \\
R_1 + R_2 &< I(X_1, X_3;Y_4)
\end{aligned}
$$

Decoding (cont.)

- Node 5:

$$\mathbf{y}_5(j) \Rightarrow \begin{matrix} \hat{w}_{2,j} \\ \hat{z}_j \end{matrix} \Rightarrow \hat{w}_{1,j-1}$$



$$\begin{aligned} R_1 &< I(X_1;Y_4 \mid X_3) \\ R_z &< I(X_3;Y_4 \mid X_1) \\ R_1 + R_z &< I(X_1, X_3;Y_4) \end{aligned} \Rightarrow \begin{aligned} R_1 &< I(X_3;Y_5 \mid X_2) \\ R_2 &< I(X_2;Y_5 \mid X_3) \\ R_1 + R_2 &< I(X_2, X_3;Y_5) \end{aligned}$$

Achievable Rate Region with Message-by-message Network Coding

- The rate pair (R_1, R_2) satisfying

$$R_1 < \min\{I(X_1;Y_3 \mid X_2, X_3), I(X_1;Y_4 \mid X_3), I(X_3;Y_5 \mid X_2)\}$$

$$R_2 < \min\{I(X_2;Y_3 \mid X_1, X_3), I(X_2;Y_5 \mid X_3), I(X_3;Y_4 \mid X_1)\}$$

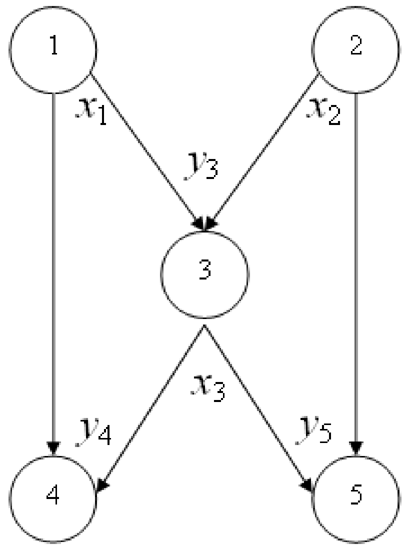
$$R_1 + R_2 < \min\{I(X_1, X_2;Y_3 \mid X_3), I(X_1, X_3;Y_4), I(X_2, X_3;Y_5)\}$$

are achievable for any distribution $p(x_1)p(x_2)p(x_3)$.

Bit-by-bit Network Coding

- z_j is not independent of $w_{1,j}$ and $w_{2,j}$.
- The codebook of $\mathbf{x}_3$ is generated through distribution

$$\prod_{i=1}^{n} p(x_{3,i} \mid x_{1,i}(w_1), x_{2,i}(w_2)).$$



$$R_1 < I(X_1;Y_4 \mid X_3) \qquad R_2 < I(X_2;Y_5 \mid X_3)$$

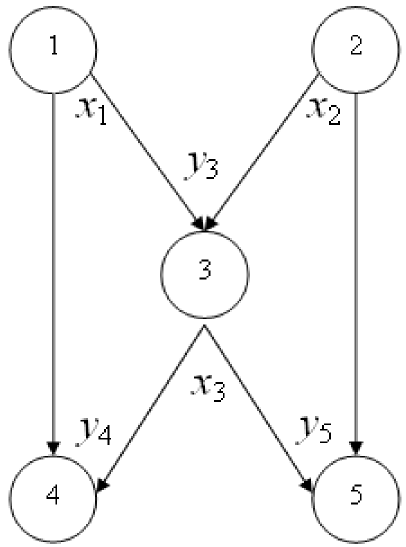
$$R_z < I(X_3;Y_4 \mid X_1) \qquad R_z < I(X_3;Y_5 \mid X_2)$$

$$R_1 + R_z < I(X_1, X_3;Y_4) \qquad R_2 + R_z < I(X_2, X_3;Y_5)$$

Bit-by-bit Network Coding

- z_j is not independent of $w_{1,j}$ and $w_{2,j}$.
- The codebook of $\mathbf{x}_3$ is generated through distribution

$$\prod_{i=1}^{n} p(x_{3,i} \mid x_{1,i}(w_1), x_{2,i}(w_2)).$$



$$R_1 < I(X_1;Y_4 \mid X_3) \qquad R_2 < I(X_2;Y_5 \mid X_3)$$

$$R_z < I(X_3;Y_4 \mid X_1) \qquad R_z < I(X_3;Y_5 \mid X_2)$$

$$R_1 + R_z < I(X_1, X_3;Y_4) \qquad R_2 + R_z < I(X_2, X_3;Y_5)$$

Achievable Rate Region with Bit-by-bit Network Coding

- The rate pair (R_1, R_2) satisfying

$$R_1 < \min\{I(X_1;Y_3 \mid X_2, X_3), I(X_1;Y_4)\}$$

$$R_2 < \min\{I(X_2;Y_3 \mid X_1, X_3), I(X_2;Y_5)\}$$

$$R_1 + R_2 < I(X_1, X_2;Y_3 \mid X_3)$$

are achievable for any distribution

$p(x_1)p(x_2)p(x_3 \mid x_1, x_2)$,

subject to…

Achievable Rate Region with Bit-by-bit Network Coding (cont.)

subject to

$$H(W_1 \mid Z) < I(X_1; Y_4 \mid X_3, Z)$$

$$H(Z \mid W_1) < I(X_3; Y_4 \mid X_1, W_1)$$

$$H(W_1, Z) < I(X_1, X_3; Y_4)$$

$$H(W_2 \mid Z) < I(X_2; Y_5 \mid X_3, Z)$$

$$H(Z \mid W_2) < I(X_3; Y_5 \mid X_2, W_2)$$

$$H(W_2, Z) < I(X_2, X_3; Y_5).$$

Converse for the Butterfly Channel with Message-by-message Network Coding

Given any sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes with $P_e^{(n)} \to 0$, the rates satisfy

$$R_1 \le \min\{I(X_1;Y_3 \mid X_2, X_3, Q), I(X_1;Y_4 \mid X_3, Q), I(X_3;Y_5 \mid X_2, Q)\}$$

$$R_2 \le \min\{I(X_2;Y_3 \mid X_1, X_3, Q), I(X_2;Y_5 \mid X_3, Q), I(X_3;Y_4 \mid X_1, Q)\}$$

$$R_1 + R_2 \le \min\{I(X_1, X_2;Y_3 \mid X_3, Q), I(X_1, X_3;Y_4, Q), I(X_2, X_3;Y_5, Q)\}$$

for some choice of random variable Q and joint distribution $p(x_1 \mid q)\, p(x_2 \mid q)\, p(x_3 \mid q)\, p(y_3, y_4, y_5 \mid x_1, x_2, x_3)$.

Outer Bounds on the Rate Region of the Butterfly Channel with Bit-by-bit Network Coding

$$R_1 \le \min\{I(X_1;Y_3 \mid X_2, X_3), I(X_1;Y_4)\}$$

$$R_2 \le \min\{I(X_2;Y_3 \mid X_1, X_3), I(X_2;Y_5)\}$$

$$R_1 + R_2 \le I(X_1, X_2;Y_3 \mid X_3)$$

with

$$H(W_1 \mid Z) \le I(X_1;Y_4 \mid X_3, Z^m, Q)$$

$$H(Z \mid W_1) \le I(X_3;Y_4 \mid X_1, W_1^m, Q)$$

$$H(W_1, Z) \le I(X_1, X_3;Y_4, Q)$$

$$H(W_2 \mid Z) \le I(X_2;Y_5 \mid X_3, Z^m, Q)$$

$$H(Z \mid W_2) \le I(X_3;Y_5 \mid X_2, W_2^m, Q)$$

$$H(W_2, Z) \le I(X_2, X_3;Y_5, Q)$$

Type II Butterfly Channel

- The capacity of the butterfly channel (Type II) is the closure of the convex hull of all (R_1, R_2) satisfying

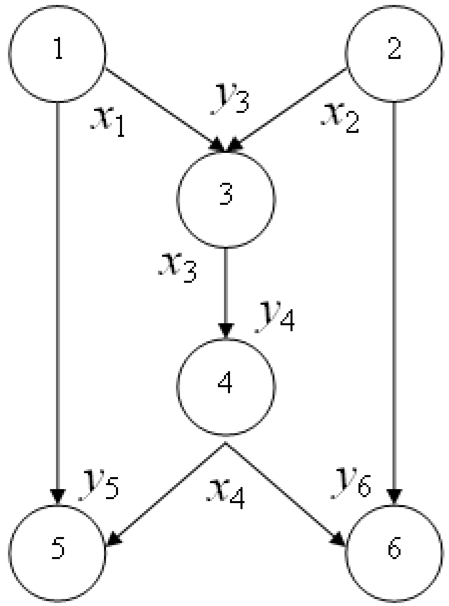
$$R_1 < \min\{I(X_1;Y_3 \mid X_2, X_3), I(X_3;Y_4 \mid X_4),$$
$$I(X_1;Y_5 \mid X_4), I(X_4;Y_6 \mid X_2)\}$$

- $$R_2 < \min\{I(X_2;Y_3 \mid X_1, X_3), I(X_3;Y_4 \mid X_4),$$
$$I(X_2;Y_6 \mid X_4), I(X_4;Y_5 \mid X_1)\}$$

$$R_1 + R_2 < \min\{I(X_1, X_2;Y_3 \mid X_3), I(X_1, X_4;Y_5),$$
$$I(X_2, X_4;Y_6)\}$$

for some product distribution

$$p(x_1)\, p(x_2)\, p(x_3 \mid x_1, x_2)\, p(x_4).$$



Thank you!