Achievable Rate Regions of the Butterfly Network with Noisy Links and End-to-End Error Correction

Lawrence Ong, Sarah J. Johnson, and Christopher M. Kellett
 School of Electrical Engineering and Computer Science, The University of Newcastle
 Email: lawrence.ong@cantab.net, {Sarah.Johnson, Chris.Kellett}@newcastle.edu.au

Abstract—We characterize achievable rate regions of the butterfly network with noisy links and end-to-end error correction. We consider low-complexity relays, i.e., relays without the processing power to decode received signals. Decoding is done only at the destinations. Three decoding strategies are studied: independent, serial, and joint decoding.

I. INTRODUCTION

In this paper, we characterize achievable rate regions of different decoding strategies on the butterfly network [1] with noisy links and end-to-end error correction. It has been shown [2], [3] by simulations with low-density parity-check codes that joint decoding schemes can outperform serial decoding, which in turn outperforms independent decoding. In this paper, we derive the information-theoretic achievable rates of these three decoding strategies in the Shannon sense.

We consider relays with minimal processing power. They are unable to decode received signals and can only perform *simple* operations on the received signals, e.g., adding and scaling, before relaying. This can be used to model low-complexity networks where nodes with limited capability are used to meet size, cost, and power constraints.

We examine the butterfly network, which has been thoroughly studied in the case of noiseless links. Consider the butterfly network in Fig. 1 where node 1 aims to transmit messages (A, B) to nodes 6 and 7, and each arrow represents a directed noiseless link of unit capacity. It has been shown [1] that optimal transmission across the network sends A on links $1 \rightarrow 2$, $2 \rightarrow 4$, and $2 \rightarrow 6$, B on links $1 \rightarrow 3$, $3 \rightarrow 4$, and $3 \rightarrow 7$, and $A \oplus B$ on links $4 \rightarrow 5$, $5 \rightarrow 6$, and $5 \rightarrow 7$, where $\oplus$ denotes bit-wise XOR. This way of combining messages at node 4 is called network coding.

When the links are noisy and independent, it can be shown [4] that it is optimal to use channel coding on the links to remove the noise and to create virtual noiseless channels, and network coding is then performed on the virtual channels (separation theorem). However, if the relays are not capable of decoding the received signals, we cannot use channel coding to create virtual noiseless channels. Hence, the achievable rates depend heavily on the decoding scheme implemented at the destinations. In this paper, we investigate how three different decoding schemes at the destinations perform. We consider:

- 1) independent decoding of A and B ,

This work is supported by the Australian Research Council under grant DP0877258.

Fig. 1. The butterfly network

- 2) serial decoding: channel decoding of A , network decoding, then channel decoding of B ,
- 3) joint decoding: joint network and channel decoding.

II. CHANNEL MODEL

We consider the message flow from node 1 to node 6 on the links marked by solid lines in Fig. 1. Without loss of generality, we assume links $1 \rightarrow 2$ and $1 \rightarrow 3$ are noiseless as the effect of these links can be included on links $2 \rightarrow 6$, $2 \rightarrow 4$, and $3 \rightarrow 4$. We also combine links $4 \rightarrow 5$ and $5 \rightarrow 6$ to form link $4 \rightarrow 6$. The equivalent simplified network is shown in Fig. 2. We denote by $X_{i,j}$ and $Y_{i,j}$ the channel input from node i onto link $i \rightarrow j$ and the channel output from link $i \rightarrow j$ to node j respectively. Each channel is specified by a probability function $p(y_{i,j}|x_{i,j})$. We assume the links are time invariant, memoryless, and independent, modeling wired networks or wireless networks with orthogonal links for all node pairs.

In the equivalent network, we take nodes 2 and 3 as sources for message A and B respectively, but with the constraint that $X_{2,4} = X_{2,6}$ as node 2 merely forwards its received signal from node 1, without decoding and re-encoding. Let A and B be independently, randomly, and uniformly chosen from the message alphabets $\{1, 2, \dots, \lfloor 2^{nR_A} \rfloor\}$ and $\{1, 2, \dots, \lfloor 2^{nR_B} \rfloor\}$ respectively, where n is the block length of the channel codes for all channels. The aim is to send both A and B to node 6 in n simultaneous channel uses on each link. We use $\hat{A}$ and $\hat{B}$ to denote the estimates for A and B respectively at node 6. The rate pair (R_A, R_B) is *achievable* if $\Pr\{(\hat{A}, \hat{B}) \neq (A, B)\}$ can be made arbitrarily small. Node 6 can *reliably* decode A (or B) iff R_A (or R_B) is achievable. The *capacity* is the set of all achievable rate. We consider two types of channels:

- 1) Additive white Gaussian noise (AWGN) channel, in which nodes sum and scale received signals.

Fig. 2. Channel inputs, channel outputs, and channel transition probabilities

- 2) Binary symmetric channel (BSC), in which nodes XOR received bits.

III. ACHIEVABLE RATE REGIONS FOR AWGN CHANNELS

First, we consider AWGN channels, in which

$$Y_{i,j} = X_{i,j} + Z_{i,j}, \quad (1)$$

where the channel input, $X_{i,j} \in \mathcal{R}$, is subject to a fixed per-node per-link transmit power $E[X_{i,j}^2] = P_{i,j}$ ¹, and the noise, $Z_{i,j} \in \mathcal{R}$, is a zero-mean, independent Gaussian random variable with variance $\sigma_{i,j}^2$.

Random code construction: For a link with independent additive white Gaussian noise, the capacity can be attained by a Gaussian distributed channel input [5]. In this section, we choose Gaussian distributed codewords.

Let $p(u)$ be a Gaussian distribution with zero mean and unit variance. We construct the codebook $\mathcal{C}_A$ by independently generating $\lfloor 2^{nR_A} \rfloor$ codewords in $\mathcal{R}^n$ accordingly $p(u)^n$. We denote these codewords as $\mathbf{U}(1), \mathbf{U}(2), \dots, \mathbf{U}(\lfloor 2^{nR_A} \rfloor)$. Here, $\mathbf{U} = (U[1], U[2], \dots, U[n])$, where $U[k]$ is the k -th independently generated Gaussian code letter. Independently, we construct the codebook $\mathcal{C}_B$ by independently generating $\lfloor 2^{nR_B} \rfloor$ codewords in $\mathcal{R}^n$ accordingly $p(u)^n$. We denote these codewords as $\mathbf{V}(1), \mathbf{V}(2), \dots, \mathbf{V}(\lfloor 2^{nR_B} \rfloor)$.

Properties of relays: The relays send the summed-and-scaled version of the received signals.

Source transmission: Node 2 sends A using codebook $\mathcal{C}_A$, i.e., it sends codewords $\mathbf{X}_{2,6}(A) = \sqrt{P_{2,6}}\mathbf{U}(A)$ and $\mathbf{X}_{2,4}(A) = \sqrt{P_{2,4}}\mathbf{U}(A)$ down links $2 \rightarrow 6$ and $2 \rightarrow 4$ respectively. Here, $\mathbf{X}_{i,j} = (X_{i,j}[1], X_{i,j}[2], \dots, X_{i,j}[n])$ is a sequence of n channel inputs to link $i \rightarrow j$, where $X_{i,j}[k]$ is the k -th channel input. Similarly, node 3 sends $\mathbf{X}_{3,4}(B) = \sqrt{P_{3,4}}\mathbf{V}(B)$.

Relay Transmission: Upon receiving channel output sequences $\mathbf{Y}_{2,4}$ and $\mathbf{Y}_{3,4}$, node 4 simply sums, scales, and sends

$$\mathbf{X}_{4,6} = \alpha (\mathbf{Y}_{2,4} + \mathbf{Y}_{3,4}) \quad (2a)$$

$$= \alpha (X_{2,4} + Z_{2,4} + X_{3,4} + Z_{3,4}), \quad (2b)$$

¹We assume fixed transmit power for simplicity, using an average power constraint, it is possible to achieve higher rates.

where the multiplication and the addition are done element-wise, and $\alpha = \sqrt{\frac{P_{4,6}}{P_{2,4} + \sigma_{2,4}^2 + P_{3,4} + \sigma_{3,4}^2}}$ so that the power constraint of node 4 is met. Note that α is fixed for fixed transmit power considered in this paper, and we assume that it is *hard-coded* at node 4.

Node 6 receives $Y_{2,6} = X_{2,6} + Z_{2,6}$ on link $2 \rightarrow 6$, and

$$Y_{4,6} = X_{4,6} + Z_{4,6} \quad (3a)$$

$$= \alpha \left(X_{2,4} + X_{3,4} + Z_{2,4} + Z_{3,4} + \frac{1}{\alpha} Z_{4,6} \right) \quad (3b)$$

$$= \alpha (X_{2,4} + X_{3,4} + Z'), \quad (3c)$$

on link $4 \rightarrow 6$, where $Z' = Z_{2,4} + Z_{3,4} + \frac{1}{\alpha} Z_{4,6}$ is a zero-mean independent Gaussian variable with variance $\sigma'^2 = \sigma_{2,4}^2 + \sigma_{3,4}^2 + \frac{1}{\alpha^2} \sigma_{4,6}^2$.

First, we derive an achievable rate region when the decoders on the links do not cooperate.

A. Independent Decoding

Strategy: Node 6 can reliably decode A from $\mathbf{Y}_{2,6}$ and (A, B) from $\mathbf{Y}_{4,6}$. Note that the decoding of each link is done independently, i.e., no information is shared during decoding.

Rates: Node 6 can reliably decode A if [5]

$$R_A \leq I(X_{2,6}; Y_{2,6}) = L \left(\frac{P_{2,6}}{\sigma_{2,6}^2} \right), \quad (4)$$

where $L(x) = \frac{1}{2} \log(1 + x)$.

Eq. (3c) is a multiple-access channel [6], [7] from $X_{2,4}$ and $X_{3,4}$ to $Y_{4,6}$. Node 6 can reliably decode A and B if

$$R_A \leq I(X_{2,4}; Y_{4,6} | X_{3,4}) = L \left(\frac{P_{2,4}}{\sigma'^2} \right) \quad (5a)$$

$$R_B \leq I(X_{3,4}; Y_{4,6} | X_{2,4}) = L \left(\frac{P_{3,4}}{\sigma'^2} \right) \quad (5b)$$

$$R_A + R_B \leq I(X_{2,4}, X_{3,4}; Y_{4,6}) = L \left(\frac{P_{2,4} + P_{3,4}}{\sigma'^2} \right). \quad (5c)$$

Remark 1: Using independent decoding, no information is shared between the links: node 6 decodes A solely from $\mathbf{Y}_{2,6}$ and B solely from $\mathbf{Y}_{4,6}$. However, when decoding B from $\mathbf{Y}_{4,6}$, A is also decoded on the same link (hence R_A is restricted by (5a), (5c)). Otherwise, A acts as an interference on B , and R_B in (5b) reduces to $R_B \leq I(X_{3,4}; Y_{4,6}) = L(\frac{P_{3,4}}{\sigma'^2 + P_{2,4}})$, but constraints (5a) and (5c) can be removed.

If information can be shared between the links, the decoded A from $\mathbf{Y}_{2,6}$ can be used to remove the interference of A in $\mathbf{Y}_{4,6}$ without requiring A to be decoded from $\mathbf{Y}_{4,6}$. This strategy is studied next.

B. Serial Decoding

Strategy: Node 6 first decodes message A from link $2 \rightarrow 6$. It then reconstructs $\mathbf{X}_{2,4}(A)$ and subtracts it from $\mathbf{Y}_{4,6}$ before decoding message B .

Rates: So, on link $2 \rightarrow 6$, we have

$$R_A \leq I(X_{2,6}; Y_{2,6}) = L \left(\frac{P_{2,6}}{\sigma_{2,6}^2} \right), \quad (6)$$

and on link $4 \rightarrow 6$, we have

$$R_B \leq I(X_{3,4}; Y_{4,6} | X_{2,4}) = L \left(\frac{P_{3,4}}{\sigma'^2} \right). \quad (7)$$

Remark 2: If link $4 \rightarrow 6$ is extremely good to the extent that it is better to decode both messages on it, then the achievable rate region is given by (5a)–(5c), i.e., constraint (6) can be removed. This also applies to independent decoding.

In serial decoding, though information is shared, decoding is still done independently. Next, we study a strategy where decoding on both links is done jointly, i.e., information is shared during decoding.

C. Joint Decoding

Strategy: Node 6 decodes messages A and B together from the received signals $\mathbf{Y}_{2,6}$ and $\mathbf{Y}_{4,6}$.

Rates: We see a multiple-access channel from X_2 and $X_{3,4}$ to Y_6 , where $X_2 = X_{2,4} = X_{2,6}$ and $Y_6 = (Y_{2,6}, Y_{4,6})$. Hence, we have the following capacity region.

$$R_A \leq I(X_2; Y_6 | X_{3,4}) \quad (8a)$$

$$R_B \leq I(X_{3,4}; Y_6 | X_2) = I(X_{3,4}; Y_{4,6} | X_{2,6}) = L \left(\frac{P_{3,4}}{\sigma'^2} \right) \quad (8b)$$

$$R_A + R_B \leq I(X_2, X_{3,4}; Y_6). \quad (8c)$$

The RHS of (8a) can be evaluated as

$$H(Y_{2,6}, Y_{4,6} | X_{3,4}) - H(Y_{2,6}, Y_{4,6} | X_{2,4}, X_{2,6}, X_{3,4}) \quad (9a)$$

$$= H(X_{2,6} + Z_{2,6}, X_{2,4} + Z') - H(Z_{2,6}, Z') \quad (9b)$$

$$= \frac{1}{2} \log \left[(2\pi e)^2 \begin{vmatrix} P_{2,6} + \sigma_{2,6}^2 & \sqrt{P_{2,6}P_{2,4}} \\ \sqrt{P_{2,6}P_{2,4}} & P_{2,4} + \sigma'^2 \end{vmatrix} \right] \quad (9c)$$

$$- \frac{1}{2} \log \left[(2\pi e)^2 \begin{vmatrix} \sigma_{2,6}^2 & 0 \\ 0 & \sigma'^2 \end{vmatrix} \right] \quad (9d)$$

$$= L \left(\frac{P_{2,6}}{\sigma_{2,6}^2} + \frac{P_{2,4}}{\sigma'^2} \right), \quad (9d)$$

where $|M|$ is the determinant of the matrix M , and the RHS of (8c) can be evaluated as

$$H(Y_{2,6}, Y_{4,6}) - H(Y_{2,6}, Y_{4,6} | X_{2,4}, X_{2,6}, X_{3,4}) \quad (10a)$$

$$= \frac{1}{2} \log \left[(2\pi e)^2 \begin{vmatrix} P_{2,6} + \sigma_{2,6}^2 & \sqrt{P_{2,6}P_{2,4}} \\ \sqrt{P_{2,6}P_{2,4}} & P_{2,4} + P_{3,4} + \sigma'^2 \end{vmatrix} \right] \quad (10b)$$

$$- \frac{1}{2} \log \left[(2\pi e)^2 \sigma_{2,6}^2 \sigma'^2 \right] \quad (10c)$$

$$= L \left(\frac{P_{2,6}}{\sigma_{2,6}^2} + \frac{P_{2,4}}{\sigma'^2} + \frac{P_{3,4}}{\sigma'^2} + \frac{P_{2,6}P_{3,4}}{\sigma_{2,6}^2 \sigma'^2} \right). \quad (10c)$$

It can be shown that the corner points of the above capacity is achievable by successive decoding, i.e., by decoding A from $(\mathbf{Y}_{2,6}, \mathbf{Y}_{4,6})$, subtracting $\mathbf{X}_{2,6}(A)$ and $\mathbf{X}_{2,4}(A)$ from $(\mathbf{Y}_{2,6}, \mathbf{Y}_{4,6})$, and then decoding B ; or vice versa, i.e., decoding B first, subtracting B , and followed by decoding A .

Fig. 3. Achievable rate regions for different decoding schemes for AWGN channels

D. Comparison of the Decoding Schemes

Lemma 1: Let $\Lambda_{\text{indep}} \in \mathcal{R}^2$, $\Lambda_{\text{serial}} \in \mathcal{R}^2$, and $\Lambda_{\text{joint}} \in \mathcal{R}^2$ denote the achievable rate regions of the independent, serial, and joint decoding schemes respectively on the AWGN butterfly network with low-complexity relays. We have

$$\Lambda_{\text{indep}} \subseteq \Lambda_{\text{serial}} \subseteq \Lambda_{\text{joint}}. \quad (11)$$

Proof of Lemma 1: There are two more constraints for independent decoding compared to serial decoding. Hence $\Lambda_{\text{indep}} \subseteq \Lambda_{\text{serial}}$. Constraint (8a) is more relaxed than constraint (6), and constraint (8b) is the same as constraint (7). Furthermore, combining (6) and (7), we get $R_A + R_B \leq L (P_{2,6}/\sigma_{2,6}^2 + P_{3,4}/\sigma'^2 + P_{2,6}P_{3,4}/(\sigma_{2,6}^2 \sigma'^2))$, which is stricter than (8c). Hence, $\Lambda_{\text{serial}} \subseteq \Lambda_{\text{joint}}$. ■

As an example, if $P_{i,j} = 10, \forall i, j$, $\sigma_{mn}^2 = 1, \forall m, n$, then the achievable rate regions for the three decoding schemes are as shown in Fig. 3. Operating at $R_A = 0.8$ bits/channel use, we can increase R_B from independent decoding to serial decoding. On the other hand, operating at $R_B = 0.8$ bits/channel use, we can increase R_A from independent decoding to serial decoding, and further improve it with joint decoding.

IV. ACHIEVABLE RATE REGIONS FOR BINARY SYMMETRIC CHANNELS

Next, we consider the binary symmetric channel (BSC), where the channel input $X_{i,j}$ and the channel output $Y_{i,j}$ both take values in $\{0, 1\}$. The BSC $i \rightarrow j$ is defined by the following conditional probabilities:

$$\Pr\{Y_{i,j} = m | X_{i,j} = m\} = 1 - p_{i,j}, \quad m \in \{0, 1\}$$

$$\Pr\{Y_{i,j} = n | X_{i,j} = m\} = p_{i,j}, \quad m, n \in \{0, 1\}, n \neq m.$$

In words, with probability $p_{i,j}$ the input symbol is received in error. $p_{i,j}$ is often called the cross-over probability. Equivalently, we can write

$$Y_{i,j} = X_{i,j} \oplus E_{i,j}, \quad (13)$$

where $E_{i,j} \in \{0, 1\}$, $\Pr\{E_{i,j} = 1\} = p_{i,j}$, and $\oplus$ is bit-wise XOR. It can be easily shown that the capacity of the BSC $i \rightarrow j$ is $1 - H(p_{i,j})$, where $H(p_{i,j}) = -p_{i,j} \log p_{i,j} - (1 - p_{i,j}) \log(1 - p_{i,j})$. The capacity is achieved with equiprobable

channel input distribution. Now, we study the network in Fig. 2 where all links are replaced by independent BSCs.

Properties of relays: The relays send the bit-wise XOR of received symbols on incoming links.

Source transmission: Nodes 2 and 3 send codewords for messages A and B respectively: node 2 sends $\mathbf{X}_{2,4}(A) \in \{0,1\}^n$ and $\mathbf{X}_{2,6}(A) \in \{0,1\}^n$ on links $2 \rightarrow 4$ and $2 \rightarrow 6$ respectively, where $\mathbf{X}_{2,4}(A) = \mathbf{X}_{2,6}(A)$; and node 3 sends $\mathbf{X}_{3,4}(B) \in \{0,1\}^n$ on link $3 \rightarrow 4$.

Linear codes: It has been shown by Elias [8] that the capacity of the BSC is achievable by linear codes. Hence for the BSC network, we study two cases that possibly lead to different results:

- 1) $\mathbf{X}_{2,4}(A)$ (or equivalently $\mathbf{X}_{2,6}(A)$) and $\mathbf{X}_{3,4}(B)$ are codewords from independently generated codebooks, $\mathcal{C}_A$ and $\mathcal{C}_B$ respectively. $\mathcal{C}_A = \{\mathbf{X}_{2,4}(1), \mathbf{X}_{2,4}(2), \dots, \mathbf{X}_{2,4}(\lfloor 2^{nR_A} \rfloor)\}$, is a set of length n codewords in $\{0,1\}^n$, one for each source letter. Similarly, $\mathcal{C}_B = \{\mathbf{X}_{3,4}(1), \mathbf{X}_{3,4}(2), \dots, \mathbf{X}_{3,4}(\lfloor 2^{nR_B} \rfloor)\}$
- 2) $\mathbf{X}_{2,4}(A)$ and $\mathbf{X}_{3,4}(B)$ are codewords from the same codebook, $\mathcal{C}_A$. We use linear codes and hence $\mathbf{X}_{2,4}(A) \oplus \mathbf{X}_{3,4}(B) \in \mathcal{C}_A$. In this case, since each source letter of A is mapped to a unique codeword, as is B , and hence $|\mathcal{C}_A| = \lfloor 2^{n \max\{R_A, R_B\}} \rfloor$.

Note that in both constructions, the marginal distributions of $X_{2,6}$, $X_{2,4}$, and $X_{3,4}$ are made equiprobable binary (capacity-achieving). Furthermore, $X_{2,4}$ and $X_{3,4}$ are statistically independent as messages A and B are independent.

Relay transmission: Node 4 sends $X_{4,6} = Y_{2,4} \oplus Y_{3,4}$.

An equivalent channel: Note that $Y_{4,6} = Y_{2,4} \oplus Y_{3,4} \oplus E_{4,6} = X' \oplus E'$, where $X' = X_{2,4} \oplus X_{3,4}$ and $E' = E_{2,4} \oplus E_{3,4} \oplus E_{4,6}$. So, we can think of an equivalent channel from X' to $Y_{4,6}$ as a BSC with the following cross-over probability.

$$\Pr\{E' = 1\} = p' = [(1 - p_{2,4})(1 - p_{3,4}) + p_{2,4}p_{3,4}](p_{4,6}) + [p_{2,4}(1 - p_{3,4}) + (1 - p_{2,4})p_{3,4}](1 - p_{4,6}). \quad (14)$$

The equivalent channel is depicted in Fig. 4. Note that $(X_{2,4}, X_{3,4}) - X' - X_{4,6} - Y_{4,6}$ forms a Markov chain.

A. Independent Decoding

Using independent decoding, node 6 decodes $\mathbf{X}_{2,6}(A)$ from $\mathbf{Y}_{2,6}$ and $\mathbf{X}'(D) \in \mathcal{C}_{AB}$ from $\mathbf{Y}_{4,6}$, where $\mathcal{C}_{AB}$ is the set of all “combined” codewords $\mathbf{X}'(D) = \mathbf{X}_{2,4}(A) \oplus \mathbf{X}_{3,4}(B)$ and D is the index of the codewords in $\mathcal{C}_{AB}$. Let $\mathcal{C}_{AB} = \{\mathbf{X}'(1), \mathbf{X}'(2), \dots, \mathbf{X}'(\lfloor 2^{nR_D} \rfloor)\}$.

Node 6 can reliably decode A if

$$R_A \leq C_{2,6} = 1 - H(p_{2,6}), \quad (15)$$

where $C_{2,6}$ is the capacity of link $2 \rightarrow 6$. Similarly, node 6 can reliably decode D if

$$R_D \leq C' = 1 - H(p'), \quad (16)$$

where $C' = 1 - H(p')$ is the capacity of the equivalent BSC $X' \rightarrow Y_{4,6}$, and p' is given in (14).

After decoding D and A , B can be obtained.

Fig. 4. An equivalent channel for BSC network

1) *Different Codebooks:* We assume that the codebooks $\mathcal{C}_A$ and $\mathcal{C}_B$ are independently generated and each pair of codewords $\mathbf{X}_{2,4}(A)$ and $\mathbf{X}_{3,4}(B)$ maps to a unique $\mathbf{X}'(D)$, such that message B can be obtained from A and D . In this case, $\lfloor 2^{nR_D} \rfloor = \lfloor 2^{nR_A} \rfloor \times \lfloor 2^{nR_B} \rfloor \leq \lfloor 2^{n(R_A + R_B)} \rfloor$. So if

$$R_A + R_B \leq C', \quad (17)$$

then (16) is satisfied and D can be reliably decoded.

2) *Same Codebook:* Recall that using the same codebook, $\mathbf{X}'(D) = \mathbf{X}_{2,4}(A) \oplus \mathbf{X}_{3,4}(B) \in \mathcal{C}_{AB} \subseteq \mathcal{C}_A$. If B is to be uniquely decodable from D for any given A , there must be a one-to-one mapping from D to B . So, $\lfloor 2^{nR_D} \rfloor \geq \lfloor 2^{nR_B} \rfloor$. For a fixed B , a unique $\mathbf{X}_{2,4}(A)$ can result in a unique $\mathbf{X}'(D)$. So, $\lfloor 2^{nR_D} \rfloor \geq \lfloor 2^{nR_A} \rfloor$. Fixing $\lfloor 2^{nR_D} \rfloor = \lfloor 2^{n \max\{R_A, R_B\}} \rfloor$, if

$$\max\{R_A, R_B\} \leq C', \quad (18)$$

then (16) is satisfied and D can be reliably decoded.

B. Serial Decoding

In serial decoding, message A is first decoded from $\mathbf{Y}_{2,6}$, and hence we get the rate constraint (15). The decoded message A is then used to generate $\mathbf{X}_{3,4}(A)$ and to subtract it from $\mathbf{Y}_{4,6}$. We have the following:

$$\begin{aligned} \mathbf{Y}_{4,6} \oplus \mathbf{X}_{2,4}(A) &= \mathbf{X}_{4,6} \oplus \mathbf{E}_{4,6} \oplus \mathbf{X}_{2,4}(A) \\ &= (\mathbf{X}' \oplus \mathbf{E}_{2,4} \oplus \mathbf{E}_{3,4}) \oplus \mathbf{E}_{4,6} \oplus \mathbf{X}_{2,4}(A) \\ &= \mathbf{X}_{3,4}(B) \oplus \mathbf{E}'. \end{aligned} \quad (19)$$

By subtracting $\mathbf{X}_{2,4}(A)$ from $\mathbf{Y}_{4,6}$, we get an equivalent BSC with cross-over probability p' as depicted in Fig. 4(ii), with X' replaced by $X_{3,4}$ and $Y_{4,6}$ replaced by $(Y_{4,6} \oplus X_{2,4})$. So, node 6 can reliably decode message B if

$$R_B \leq C'. \quad (20)$$

C. Joint Decoding

With joint decoding, messages A and B are jointly decoded from the received messages $\mathbf{Y}_{2,6}$ and $\mathbf{Y}_{4,6}$. We see a multiple-access channel from X_2 and $X_{3,4}$ to Y_6 , where $X_2 = X_{2,4} =$

Fig. 5. Achievable rate regions for different decoding schemes for BSC, $\rho_{i,j} = 0.05$ for all links

$X_{2,6}$ and $Y_6 = (Y_{2,6}, Y_{4,6})$. Hence, we have the following capacity region:

$$R_A \leq I(X_2; Y_6 | X_{3,4}) \quad (21a)$$

$$R_B \leq I(X_{3,4}; Y_6 | X_2) \quad (21b)$$

$$R_A + R_B \leq I(X_2, X_{3,4}; Y_6), \quad (21c)$$

maximized over all possible $p(x_2, x_{3,4})$. The capacity region can be attained by independent and equiprobable X_2 and $X_{3,4}$.

Evaluating $I(X_2; Y_6 | X_{3,4})$ gives

$$\begin{aligned} I(X_{2,4}, X_{2,6}; Y_{2,6}, Y_{4,6} | X_{3,4}) \\ = H(Y_{2,6}, Y_{4,6} | X_{3,4}) - H(Y_{2,6}, Y_{4,6} | X_{3,4}, X_{2,4}, X_{2,6}) \end{aligned} \quad (22a)$$

$$= (1 + H(\tilde{p})) - (H(p_{2,6}) + H(p')) \quad (22b)$$

$$= C_{2,6} + C' - \tilde{C} \quad (22c)$$

where $\tilde{C} = 1 - H(\tilde{p})$ is the capacity of a BSC with cross-over probability $\tilde{p} = p_{2,6}(1 - p') + (1 - p_{2,6})p'$.

Evaluating $I(X_{3,4}; Y_6 | X_2)$ gives $I(X_{3,4}; Y_{4,6}, Y_{2,6} | X_{2,6}, X_{2,4}) = I(X_{3,4}; Y_{4,6} | X_{2,4}) = C'$. The first equality is because $Y_{2,6} - X_{2,6} - (X_{2,4}, X_{3,4})$ forms a Markov chain, and the second equality is because of (19).

Finally, evaluating $I(X_2, X_{3,4}; Y_6)$ gives

$$\begin{aligned} I(X_{2,6}, X_{2,4}, X_{3,4}; Y_{2,6}, Y_{4,6}) \\ = H(Y_{2,6}, Y_{4,6}) - H(Y_{2,6}, Y_{4,6} | X_{2,6}, X_{2,4}, X_{3,4}) \end{aligned} \quad (23a)$$

$$= 2 - (H(p_{2,6}) + H(p')) \quad (23b)$$

$$= C_{2,6} + C'. \quad (23c)$$

D. Comparison

Lemma 2: Let $\Lambda_{\text{indep,same}}$, $\Lambda_{\text{indep,diff}}$, Λ_{serial} , and Λ_{joint} denote the achievable rate regions of independent decoding with the same codebook, independent decoding with different codebooks, serial decoding, and joint decoding on the BSC butterfly network with low-complexity relays respectively. We have

$$\Lambda_{\text{indep,same}} \subseteq \Lambda_{\text{indep,diff}} \subseteq \Lambda_{\text{serial}} \subseteq \Lambda_{\text{joint}}. \quad (24)$$

Proof of Lemma 2: The constraint (17) is stricter than the constraint (18), so $\Lambda_{\text{indep,same}} \subseteq \Lambda_{\text{indep,diff}}$. The constraint

(18) is stricter than the constraint (20), so $\Lambda_{\text{indep,diff}} \subseteq \Lambda_{\text{serial}}$. The constraint (15) is stricter than the constraint (21a) because $C' \geq \tilde{C}$. Summing (15) and (20) gives $R_A + R_B \leq C_{2,6} + C'$ which is the constraint (21c). So, $\Lambda_{\text{serial}} \subseteq \Lambda_{\text{joint}}$. ■

Figure 5 shows the achievable rates for the four decoding strategies. We see that the achievable regions are similar to those for AWGN channels.

Independent decoding achieves a larger rate region using the same codebook compared to different codebooks. This is because message D is a complete description of messages A and B in the latter but not in the former (where we only need to meet the requirement that B can be decoded from D and A). Hence, we get some “bandwidth” savings using the same codebook, similar to that in traditional network coding.

With serial decoding (or joint decoding), using different codebooks or the same codebook for messages A and B achieves the same rate region.

V. REFLECTION

While it is not surprising that joint decoding outperforms the other schemes, this paper provides theoretical analyses of the different decoding schemes and helps us to understand their potential independently of the chosen channel codes.

Compared to independent decoding, serial decoding can (i) improve R_B as it subtracts $\mathbf{X}_{2,4}(A)$ from the received message $\mathbf{Y}_{4,6}$, and thus cancels the interference from message A before decoding message B , and (ii) improve R_A as it does not need to decode the entire combined message D (which includes A) from $\mathbf{Y}_{4,6}$, and thus reduces a constraint on R_A .

Compared to serial decoding, joint decoding can improve R_A because node 6 decodes message A from both $\mathbf{Y}_{2,6}$ and $\mathbf{Y}_{4,6}$ in the latter, but solely from $\mathbf{Y}_{2,6}$ in former. Joint decoding does not improve R_B over serial decoding as only $\mathbf{Y}_{4,6}$ carries information about B . Upon canceling the interference by A when decoding B , serial decoding is already doing the best for decoding B .

REFERENCES

- [1] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, “Network information flow,” *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [2] S. J. Johnson and C. Kellett, “Joint network and channel coding for co-operative networks,” in *Proc. Australasian Telecommunication Networks and Applications Conference 2007 (ATNAC)*, Christchurch, New Zealand, Dec. 2-5 2007, pp. 435–440.
- [3] J. Kang, B. Zhou, Z. Ding, and S. Lin, “LDPC coding schemes for error control in a multicast network,” in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Toronto, Canada, Jul. 6-11 2008, pp. 822–826.
- [4] L. Song, R. W. Yeung, and N. Cai, “A separation theorem for single-source network coding,” *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1861–1871, May 2006.
- [5] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, Jul. and Oct. 1948.
- [6] H. Liao, “Multiple access channel,” Ph.D. dissertation, Univ. Hawaii Honolulu, HI, 1972.
- [7] R. Ahlswede, “The capacity of a channel with two senders and two receivers,” *Ann. Probab.*, vol. 2, pp. 805–814, Oct. 1974.
- [8] P. Elias, “Coding for noisy channels,” *IRE Conv. Rec.*, vol. 4, pp. 37–46, Mar. 1955.
- [9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, 1991.