Layered Noisy Network Coding

Sung Hoon Lim
Department of EE
KAIST
Daejeon, Korea
sunghlim@kaist.ac.kr

Young-Han Kim
Department of ECE
UCSD
La Jolla, CA 92093, USA
yhk@ucsd.edu

Abbas El Gamal
Department of EE
Stanford University
Standord, CA 94305, USA
abbas@ee.stanford.edu

Sae-Young Chung
Department of EE
KAIST
Daejeon, Korea
sychung@ee.kaist.ac.kr

***Abstract*—Noisy network coding naturally combines two parallel lines of work on network information flow — network coding over noiseless networks by Ahlswede, Cai, Li, and Yeung, and compress–forward coding for the relay channel by Cover and El Gamal and extends both results to general discrete memoryless and Gaussian networks. In particular, it achieves the best known capacity inner bounds for several multi-source multicast networks including deterministic networks by Avestimehr, Diggavi, and Tse and erasure networks by Dana, Gowaikar, Palanki, Hassibi, and Effros.**

In this paper we further improve noisy network coding for the two-way relay channel. In the new scheme, instead of the relay node just sending a common compression index to both destinations, the relay compresses its observation into a description of two layers. The common layer is used at both decoders, while the refinement layer is used only at one of the decoders. The advantage of the new scheme is demonstrated through an example of the Gaussian two-way relay channel.

I. Introduction

Consider a general N-node discrete memoryless network (DMN)

$$(\mathcal{X}_1\times\cdots\times\mathcal{X}_N, p(y_1,\ldots,y_N|x_1,\ldots,x_N), \mathcal{Y}_1\times\cdots\times\mathcal{Y}_N) \quad (1)$$

in which each node $k\in\{1,\ldots,N\} =: [1:N]$ wishes to send a message M_k to a set $\mathcal{D}_k\subset[1:N]$ of destination nodes. Formally, a $(2^{nR_1},\ldots,2^{nR_N},n)$ code for a DMN consists of N message sets $[1:2^{nR_1}],\ldots,[1:2^{nR_N}]$, a set of encoders with encoder $k\in[1:N]$ that assigns an input symbol x_{ki} to each pair (m_k, y_k^{i-1}) for $i\in[1:n]$, and a set of decoders with decoder $d\in\cup_{k=1}^N\mathcal{D}_k$ that assigns message estimates $\{\hat{m}_{kd} : k\in\mathcal{S}_d\}$ to each (y_d^n, m_d), where $\mathcal{S}_d := \{k : d\in\mathcal{D}_k\}$ is the set of nodes that send messages to destination d.

We assume that messages M_k, $k\in[1:N]$, are each uniformly distributed over $[1:2^{nR_k}]$, and are independent of each other. The average probability of error is defined by

$$P_e^{(n)} = \mathsf{P}\{\hat{M}_{kd}\neq M_k \text{ for some } d\in\mathcal{D}_k, k\in[1:N]\}.$$

A rate tuple $(R_1,\ldots,R_N)$ is said to be achievable if there exists a sequence of $(2^{nR_1},\ldots,2^{nR_N},n)$ codes with $P_e^{(n)}\to 0$ as $n\to\infty$. The capacity region of the DMN is the closure of the set of achievable rate tuples.

Recently, noisy network coding was proposed [1], which achieves the following rate tuples for the multicast setting, i.e., $\mathcal{D}_k=\mathcal{D}$ for $k\in[1:N]$.

Theorem 1 ([1]): Suppose $\mathcal{D}_k=\mathcal{D}$ for $k\in[1:N]$. A rate tuple $(R_1,\ldots,R_N)$ is achievable if there exists some joint pmf $p(q)\prod_{k=1}^N p(x_k|q)p(\hat{y}_k|y_k,x_k,q)$ such that

$$\begin{aligned} R(\mathcal{S}) < \min_{d\in\mathcal{S}^c\cap\mathcal{D}} &(I(X(\mathcal{S});\hat{Y}(\mathcal{S}^c),Y_d|X(\mathcal{S}^c),Q) \\ &- I(Y(\mathcal{S});\hat{Y}(\mathcal{S})|X^N,\hat{Y}(\mathcal{S}^c),Y_d,Q)) \end{aligned} \quad (2)$$

for all *cutsets* $\mathcal{S}\subseteq[1:N]$ with $\mathcal{S}^c\cap\mathcal{D}\neq\emptyset$, where $R(\mathcal{S})=\sum_{k\in\mathcal{S}}R_k$.

By evaluating Theorem 1 for specific networks, it can be shown that noisy network coding recovers many previously known results, such as the results on deterministic networks and Gaussian relay networks by Avestimehr, Diggavi, and Tse [2], and the results on wireless erasure networks by Dana, Gowaikar, Palanki, Hassibi, and Effros [3]. In fact, noisy network coding naturally combines, network coding over noiseless networks by Ahlswede, Cai, Li, and Yeung [4], and compress–forward coding for the relay channel by Cover and El Gamal [5], extending both results to general discrete memoryless and Gaussian networks.

In a way, the noisy network coding scheme can be viewed as transforming a multi-hop relay network into a single-hop interference network where the channel outputs are compressed versions of the received signals. For non-multicast networks with general message demand, two additional coding schemes has been proposed, which combines the underlying network compress–forward coding scheme for multicast networks with decoding techniques for interference channels. At one extreme, noisy network coding is combined with simultaneous decoding [1, Theorem 2], while at the other extreme interference is treated as noise [1, Theorem 3].

Noisy network coding involves three key ingredients: message repetition coding, simultaneous decoding, and relay signal compression. Message repetition coding, unlike previous block Markov schemes where independent messages are sent over multiple blocks, operates by sending the *same* message through multiple blocks using independent codebooks. Furthermore, the compression indices are sent directly without using Wyner–Ziv binning as done in previous compress–forward schemes. For the decoding procedure, each decoder performs simultaneous joint typicality decoding on the received signals from all the blocks *without* explicitly decoding the compression indices. As a consequence of the scheme, arbitrary delays due to multiple hops in the network can be

978-1-4244-7980-1/10/$26.00 ©2010 IEEE

handled in a much simpler manner without the two-stage approach which resorts to time expansion for extending the results of acyclic networks to networks with cycles. Additionally, the approach differs from the classical compress–forward scheme [5], where there are two decoding stages involving the compression index. In the classical compress–forward scheme, the decoder first decodes the bin index and the compression index is decoded sequentially. The two decoding stages imply additional constraints compared to noisy network coding, and it was shown in [1] by considering two examples, the two-way relay channel and the interference relay channel, that noisy network coding outperforms other compress–forward variants.

In this paper, we focus on the two-way relay channel as a canonical model for short-range multiple user networks, and further improve noisy network coding for this specific channel model. Various coding schemes for the two-way relay channel have been investigated in [6], [7], [8]. Here, we exploit the simple structure of the network to develop more sophisticated extensions of noisy network coding that may be tedious to analyze in large scale networks.

For the two-way relay channel, the original noisy network coding scheme in [1] involves the relay node to use a *single* compression codebook for both destination nodes. However, the quality of the relay–destination links as well as the side information depending on the channel conditions may differ at the destination nodes. Accordingly, in the new scheme, instead of just sending a common compression index to both destinations, the relay node compresses its observation into two layers: the common layer and the refinement layer. The common layer is used at both decoders, while the refinement layer is used only at one of the decoders.

Throughout the paper, we follow the notation in [9]. In particular, a sequence of random variables with node index k and time index $i \in [1:n]$ is denoted as $X_k^n = (X_{k1}, \ldots, X_{kn})$. A set of random variables is denoted as $X(\mathcal{A}) = \{X_k : k \in \mathcal{A}\}$.

II. Two-way relay channel

The two-way relay channel $p(y_1, y_2, y_3|x_1, x_2, x_3)$ is a DMN with two source nodes 1 and 2 that wish to exchange independent messages reliably with the help of relay node 3, i.e., $N = 3$, $R_3 = 0$, $\mathcal{D}_1 = \{2\}$, and $\mathcal{D}_2 = \{1\}$ in the general DMN model (1).

A. Main result

We denote $\mathcal{R}_1$ as the set of rate pairs (R_1, R_2) that satisfy

$$\begin{aligned}
R_1 &< I(X_1; \tilde{Y}_3, Y_2|X_2, U, Q),\\
R_1 &< I(X_1, U; Y_2|X_2, Q) - I(\tilde{Y}_3; Y_3, X_3|X_1, X_2, Y_2, U, Q),\\
R_2 &< I(X_2; \hat{Y}_3, \tilde{Y}_3, Y_1|X_1, X_3, U, Q),\\
R_2 &< I(X_2, X_3; Y_1|X_1, Q)\\
&\quad - I(\hat{Y}_3, \tilde{Y}_3; Y_3|X_1, X_2, X_3, U, Y_1, Q),\\
R_2 &< I(X_2, X_3; \tilde{Y}_3, Y_1|X_1, U, Q)\\
&\quad - I(\hat{Y}_3; Y_3|X_1, X_2, X_3, U, Y_1, \tilde{Y}_3, Q)
\end{aligned}$$

for some

$$p(q)p(x_1|q)p(x_2|q)p(x_3, u|q)p(\hat{y}_3, \tilde{y}_3|y_3, x_3, u, q),$$

and denote $\mathcal{R}_2$ as the set of rate pairs (R_1, R_2) that satisfy

$$\begin{aligned}
R_1 &< I(X_1; \hat{Y}_3, \tilde{Y}_3, Y_2|X_2, X_3, U, Q),\\
R_1 &< I(X_1, X_3; Y_2|X_2, Q)\\
&\quad - I(\hat{Y}_3, \tilde{Y}_3; Y_3|X_1, X_2, X_3, U, Y_2, Q),\\
R_1 &< I(X_1, X_3; \tilde{Y}_3, Y_2|X_2, U, Q)\\
&\quad - I(\hat{Y}_3; Y_3|X_1, X_2, X_3, U, Y_1, \tilde{Y}_3, Q)\\
R_2 &< I(X_2; \tilde{Y}_3, Y_1|X_1, U, Q),\\
R_2 &< I(X_2, U; Y_1|X_1, Q) - I(\tilde{Y}_3; Y_3, X_3|X_1, X_2, Y_1, U, Q)
\end{aligned}$$

for some

$$p(q)p(x_1|q)p(x_2|q)p(x_3, u|q)p(\hat{y}_3, \tilde{y}_3|y_3, x_3, u, q).$$

Theorem 2: For the two-way relay channel, a rate pair (R_1, R_2) is achievable if it is in the convex hull of the rate region $\mathcal{R}_1 \cup \mathcal{R}_2$.

As a special case, by choosing $U = X_3$ and $\tilde{Y} = \hat{Y}$, we recover the noisy network coding inner bound in [1] that consists of rate pairs (R_1, R_2) satisfying the rate region

$$\begin{aligned}
R_1 &< I(X_1; \hat{Y}_3, Y_2|X_2, X_3, Q),\\
R_1 &< I(X_1, X_3; Y_2|X_2, Q) - I(\hat{Y}_3; Y_3|X_1, X_2, X_3, Y_2, Q),\\
R_2 &< I(X_2; \hat{Y}_3, Y_1|X_1, X_3, Q),\\
R_2 &< I(X_2, X_3; Y_1|X_1, Q) - I(\hat{Y}_3; Y_3|X_1, X_2, X_3, Y_1, Q)
\end{aligned}$$

for some $p(q)p(x_1|q)p(x_2|q)p(x_3|q)p(\hat{y}_3|y_3, x_3, q)$.

B. Proof of achievability

In the following, we prove the achievability of the rate region $\mathcal{R}_1$. The region $\mathcal{R}_2$ can be achieved by switching the role of user 1 and 2. By time sharing, the convex hull of $\mathcal{R}_1 \cup \mathcal{R}_2$ is achievable. For simplicity of notation, we consider the case $Q = \emptyset$. Achievability for an arbitrary time-sharing random variable Q can be proved using the coded time sharing technique [9].

Let $\mathbf{x}_{kj}$ denote $(x_{k,(j-1)n+1}, \ldots, x_{k,jn})$, $j \in [1:b]$; thus $x_k^{bn} = (x_{k1}, \ldots, x_{k,nb}) = (\mathbf{x}_{k1}, \ldots, \mathbf{x}_{kb}) = \mathbf{x}_k^b$. To send a message $m_k \in [1 : 2^{nbR_k}]$, the source node $k \in \{1, 2\}$ transmits $\mathbf{x}_{kj}(m_k)$ for each block $j \in [1 : b]$. In block j, the relay finds the "compressed" pair $(\hat{\mathbf{y}}_{3j}, \tilde{\mathbf{y}}_{3j})$ of the relay output $\mathbf{y}_{3j}$ conditioned on $(\mathbf{x}_{3j}, \mathbf{u}_j)$ where $\tilde{\mathbf{y}}_{3j}$ represents the common layer and $\hat{\mathbf{y}}_{3j}$ represents the refinement layer. The relay node then transmits a codeword $\mathbf{x}_{3,j+1}(\hat{\mathbf{y}}_{3j}, \tilde{\mathbf{y}}_{3j})$ which is superimposed on $\mathbf{u}_{j+1}(\tilde{\mathbf{y}}_{3j})$ in the next block. After b block transmissions, decoder 1 finds the correct message $m_2 \in [1 : 2^{nbR_2}]$ using the codebook information of the common layer to joint typical decode for each of b blocks simultaneously. Decoder 2 finds the correct message $m_1 \in [1 : 2^{nbR_1}]$ similar to decoder 1 with additionally utilizing the codebook information of the refinement layer. The details are given in the following.

Block	1	2	3	$\cdots$	b
X_1	$\mathbf{x}_{11}(m_1)$	$\mathbf{x}_{12}(m_1)$	$\mathbf{x}_{13}(m_1)$	$\ldots$	$\mathbf{x}_{1b}(m_1)$
X_2	$\mathbf{x}_{21}(m_2)$	$\mathbf{x}_{22}(m_2)$	$\mathbf{x}_{23}(m_2)$	$\ldots$	$\mathbf{x}_{2b}(m_2)$
Y_3	$\tilde{\mathbf{y}}_{31}(t_1\|1)$, $\hat{\mathbf{y}}_{31}(l_1\|t_1,1,1)$	$\tilde{\mathbf{y}}_{32}(l_2\|l_1)$, $\hat{\mathbf{y}}_{32}(l_2\|t_2,l_1,t_1)$	$\tilde{\mathbf{y}}_{33}(l_3\|l_2)$, $\hat{\mathbf{y}}_{33}(l_3\|t_3,l_2,t_2)$	$\ldots$	$\tilde{\mathbf{y}}_{3b}(l_b\|l_{b-1})$, $\hat{\mathbf{y}}_{3b}(l_b\|t_b,l_{b-1},t_{b-1})$
X_3	$\mathbf{x}_{31}(1\|1)$, $\mathbf{u}_1(1)$	$\mathbf{x}_{32}(l_1\|t_1)$, $\mathbf{u}_2(t_1)$	$\mathbf{x}_{33}(l_2\|t_2)$, $\mathbf{u}_2(t_2)$	$\ldots$	$\mathbf{x}_{3b}(l_{b-1}\|t_{b-1})$, $\mathbf{u}_b(t_{b-1})$
Y_1	$\emptyset$	$\emptyset$	$\emptyset$	$\ldots$	$\hat{m}_2$
Y_2	$\emptyset$	$\emptyset$	$\emptyset$	$\ldots$	$\hat{m}_1$

TABLE I
LAYERED NOISY NETWORK CODING FOR THE TWO-WAY RELAY CHANNEL.

Codebook generatioin: Fix

$$p(x_1)p(x_2)p(x_3,u)p(\hat{y}_3,\tilde{y}_3|x_3,u,y_3).$$

Randomly and independently generate 2^{nbR_1} sequences $\mathbf{x}_{1j}(m_1)$, $m_1 \in [1 : 2^{nbR_1}]$, each according to $\prod_{i=1}^n p_{X_1}(x_{1(j-1)n+i})$. Similarly, randomly and independently generate 2^{nbR_2} sequences $\mathbf{x}_{2j}(m_2)$, $m_2 \in [1 : 2^{nbR_2}]$, each according to $\prod_{i=1}^n p_{X_2}(x_{2(j-1)n+i})$. Randomly and independently generate $2^{n\tilde{R}}$ sequences $\mathbf{u}_j(t_{j-1})$, $t_{j-1} \in [1 : 2^{n\tilde{R}}]$, each according to $\prod_{i=1}^n p_U(u_{(j-1)n+i})$. For each $t_{j-1} \in [1 : 2^{n\tilde{R}}]$, generate $2^{n\hat{R}}$ sequences $\mathbf{x}_{3j}(l_{j-1}|t_{j-1})$, $l_{j-1} \in [1 : 2^{n\hat{R}}]$, each according to $\prod_{i=1}^n p_{X_3|U}(x_{3,(j-1)n+i}|u_{(j-1)n+i}(t_{j-1}))$. For each $t_{j-1} \in [1 : 2^{n\tilde{R}}]$, randomly and conditionally independently generate $2^{n\tilde{R}}$ sequences $\tilde{\mathbf{y}}_{3j}(t_j|t_{j-1})$ according to $\prod_{i=1}^n p_{\tilde{Y}_3|U}(\tilde{y}_{3,(j-1)n+i}|u_{(j-1)n+i}(t_{j-1}))$. For each $t_j, t_{j-1} \in [1 : 2^{n\tilde{R}}]$ and each $l_{j-1} \in [1 : 2^{n\hat{R}}]$, randomly and conditionally independently generate $2^{n\hat{R}}$ sequences $\hat{\mathbf{y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1})$ according to

$$\prod_{i=1}^n p_{\hat{Y}_3|\tilde{Y}_3,X_3,U}(\hat{y}_{3,(j-1)n+i}|\tilde{y}_{3,(j-1)n+i}(t_j|t_{j-1}), x_{3,(j-1)n+i}(l_{j-1}|t_{j-1}),u_{(j-1)n+i}(t_{j-1})).$$

This procedure generates the codebook,

$$\mathcal{C}_j\{\mathbf{x}_{1j}(m_1),\mathbf{x}_{2j}(m_2),\mathbf{x}_{3j}(l_{j-1}|t_{j-1}),\mathbf{u}_j(t_{j-1}),\tilde{\mathbf{y}}_{3j}(t_j|t_{j-1}), \hat{\mathbf{y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1}) : m_1 \in [1:2^{nbR_1}], m_2 \in [1:2^{nbR_2}], t_j,t_{j-1} \in [1:2^{n\tilde{R}}], l_j,l_{j-1} \in [1:2^{n\hat{R}}]\}$$

for $j \in [1:b]$.

Encoding and decoding are explained with the help of Table I.

Encoding: Let (m_1,m_2) be the messages to be sent. Then, the source nodes each transmits the codeword $\mathbf{x}_{1j}(m_1)$ and $\mathbf{x}_{2j}(m_2)$ in block $j \in [1:b]$. The relay node upon receiving $\mathbf{y}_{3j}$ at the end of block $j \in [1:b]$, finds an index pair (t_j,l_j) such that

$$(\hat{\mathbf{y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1}),\tilde{\mathbf{y}}_{3j}(t_j|t_{j-1}), \mathbf{x}_{3j}(l_{j-1}|t_{j-1}),\mathbf{u}_j(t_{j-1}),\mathbf{y}_{3j}) \in \mathcal{T}_{\epsilon'}^{(n)},$$

where $(l_0,t_0)=(1,1)$ by convention.

Decoding: Let $\epsilon > \epsilon'$. At the end of block b, decoder 2 looks for a unique message $\hat{m}_1 \in [1:2^{nbR_1}]$ such that

$$(\mathbf{x}_{1j}(\hat{m}_1),\mathbf{x}_{2j}(m_2),\mathbf{u}_j(\hat{t}_{j-1}), \tilde{\mathbf{y}}_{3j}(\hat{t}_j|\hat{t}_{j-1}),\mathbf{y}_{2j}) \in \mathcal{T}_\epsilon^{(n)} \text{ for all } j \in [1:b]$$

for some $\hat{t}_{j-1},\hat{t}_j \in [1:2^{n\tilde{R}}]$. On the other hand, decoder 1 looks for a unique message $\hat{m}_2 \in [1:2^{nbR_2}]$ such that

$$(\mathbf{x}_{1j}(m_1),\mathbf{x}_{2j}(\hat{m}_2),\mathbf{x}_{3j}(\hat{l}_{j-1}|\hat{t}_{j-1}),\mathbf{u}_j(\hat{t}_{j-1}), \tilde{\mathbf{y}}_{3j}(\hat{t}_j|\hat{t}_{j-1}),\hat{\mathbf{y}}_{3j}(\hat{l}_j|\hat{t}_j,\hat{l}_{j-1},\hat{t}_{j-1}),\mathbf{y}_{1j}) \in \mathcal{T}_\epsilon^{(n)},$$

for all $j \in [1:b]$, for some $\hat{t}_{j-1},\hat{t}_j \in [1:2^{n\tilde{R}}]$ and $\hat{l}_{j-1},\hat{l}_j \in [1:2^{n\hat{R}}]$.

Probability of error analysis: Let (M_1,M_2) denote the messages sent by the source nodes, and (T_j,L_j), $j \in [1:b]$ denote the index pair chosen by the relay node at block j. To bound the probability of error, assume without loss of generality that $(M_1,M_2)=(1,1)$.

Define the events:

$$\mathcal{E}_0 := \cup_{j=1}^b\{(\hat{\mathbf{Y}}_{3j}(l_j|t_j,L_{j-1},T_{j-1}),\tilde{\mathbf{Y}}_{3j}(t_j|T_{j-1}), \mathbf{U}_j(T_{j-1}),\mathbf{Y}_{3j}) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ for all } (l_j,t_j)\},$$

$$\mathcal{E}_{1,m_1} := \{(\mathbf{X}_{1j}(m_1),\mathbf{X}_{2j}(1),\mathbf{U}_j(t_{j-1}),\tilde{\mathbf{Y}}_{3j}(t_j|t_{j-1}) \mathbf{Y}_{2j}) \in \mathcal{T}_\epsilon^{(n)},\ j \in [1:b] \text{ for some } t_{j-1},t_j\},$$

$$\mathcal{E}_{2,m_2} := \{(\mathbf{X}_{1j}(1),\mathbf{X}_{2j}(m_2),\mathbf{X}_{3j}(l_{j-1}|t_{j-1}),\mathbf{U}_j(t_{j-1}), \tilde{\mathbf{Y}}_{3j}(t_j|t_{j-1}),\hat{\mathbf{Y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1}),\mathbf{Y}_{1j}) \in \mathcal{T}_\epsilon^{(n)}, j \in [1:b] \text{ for some } t_{j-1},t_j,l_{j-1},l_j\}.$$

Then, the probability of error of decoder 1 ($\mathsf{P}(\mathcal{E}_1)$) and decoder 2 ($\mathsf{P}(\mathcal{E}_2)$) are upper bounded by

$$\mathsf{P}(\mathcal{E}_1) \le \mathsf{P}(\mathcal{E}_0) + \mathsf{P}(\mathcal{E}_0^c \cap \mathcal{E}_{1,1}^c) + \mathsf{P}(\cup_{m_1\neq 1}\mathcal{E}_{1,m_1}),$$
$$\mathsf{P}(\mathcal{E}_2) \le \mathsf{P}(\mathcal{E}_0) + \mathsf{P}(\mathcal{E}_0^c \cap \mathcal{E}_{2,1}^c) + \mathsf{P}(\cup_{m_2\neq 1}\mathcal{E}_{2,m_2}).$$

Lemma 1: There exists $\delta(\epsilon')$ which tends to zero as $\epsilon' \to 0$ such that $\mathsf{P}(\mathcal{E}_0) \to 0$, as $n \to \infty$ if

$$\tilde{R} > I(\tilde{Y}_3;X_3,Y_3|U) - \delta(\epsilon'), \quad (3)$$
$$\tilde{R}+\hat{R} > I(\tilde{Y}_3;X_3,Y_3|U) + I(\hat{Y}_3;Y_3|X_3,U,\tilde{Y}_3) - \delta(\epsilon'). \quad (4)$$

Proof: See the appendix. ■

By the conditional typicality lemma [9], $\mathsf{P}(\mathcal{E}_0^c \cap \mathcal{E}_{1,1}^c) \to 0$ and $\mathsf{P}(\mathcal{E}_0^c \cap \mathcal{E}_{2,1}^c) \to 0$ as $n \to \infty$. For the remaining terms, we first present an upper bound on $\mathsf{P}(\cup_{m_2\neq 1}\mathcal{E}_{2,m_2})$. Assume without loss of generality that $(L_1,\ldots,L_b)=(1,\ldots,1)$; recall the symmetry of the codebook construction. For $j\in[1:b]$, $m_2\in[1:2^{nbR_2}]$, $t_{j-1},t_j\in[1:2^{n\tilde{R}}]$, and $l_{j-1},l_j\in[1:2^{n\hat{R}}]$, define the events

$$\begin{aligned}&\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j):=\\&\quad\big\{(\mathbf{X}_{1j}(1),\mathbf{X}_{2j}(m_2),\mathbf{X}_{3j}(l_{j-1}|t_{j-1}),\mathbf{U}_j(t_{j-1}),\\&\qquad\tilde{\mathbf{Y}}_{3j}(t_j|t_{j-1}),\hat{\mathbf{Y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1}),\mathbf{Y}_{1j})\in\mathcal{T}_\epsilon^{(n)}\big\},\end{aligned}$$

where $t_{j-1}^j=(t_{j-1},t_j)$ and $l_{j-1}^j=(l_{j-1},l_j)$. Then,

$$\begin{aligned}\mathsf{P}(\mathcal{E}_{2,m_2})&=\mathsf{P}(\cup_{l^b}\cup_{t^b}\cap_{j=1}^b\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\\&\le\sum_{t^b}\sum_{l^b}\mathsf{P}(\cap_{j=1}^b\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\\&=\sum_{t^b}\sum_{l^b}\prod_{j=1}^b\mathsf{P}(\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\qquad(5)\\&\le\sum_{t^b}\sum_{l^b}\prod_{j=2}^b\mathsf{P}(\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j)),\end{aligned}$$

where equality (5) follows since the codebook is generated independently for each block j and the channel is memoryless. Note that if $m_2\neq 1$ and $t_{j-1}\neq 1$, then

$$\begin{aligned}&(\mathbf{X}_{2j}(m_2),\mathbf{X}_{3j}(l_{j-1}|t_{j-1}),\\&\qquad\mathbf{U}_j(t_{j-1}),\tilde{\mathbf{Y}}_{3j}(t_j|t_{j-1}),\hat{\mathbf{Y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1}))\\&\sim\prod_{i=1}^n p_{X_2}(x_{2,(j-1)n+i})\,p_{X_3,U}(x_{3,(j-1)n+i},u_{(j-1)n+i})\\&\qquad\cdot p_{\tilde{Y}_3|U}(\tilde{y}_{3,(j-1)n+i}|u_{(j-1)n+i})\\&\qquad\cdot p_{\hat{Y}_3|\tilde{Y}_3,X_3,U}(\hat{y}_{3,(j-1)n+i}|\tilde{y}_{3,(j-1)n+i},\\&\qquad\qquad x_{3,(j-1)n+i},u_{(j-1)n+i})\end{aligned}$$

is independent of $(\mathbf{X}_{1j},\mathbf{Y}_{1j})$ (given $T_{j-1}=T_j=L_{j-1}=L_j=1$). Hence, by the joint typicality lemma [9], $\mathsf{P}(\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\le 2^{-n(I_1-\delta(\epsilon))}$ where

$$\begin{aligned}I_1:=&\,I(X_2,X_3;Y_1|X_1)+I(\tilde{Y}_3;X_1,X_2,X_3,Y_1|U)\\&+I(\hat{Y}_3;X_1,X_2,Y_1|X_3,U,\tilde{Y}_3).\end{aligned}$$

Similarly, if $m_2\neq 1$, $t_{j-1}=1$, and $l_{j-1}\neq 1$, then

$$\begin{aligned}&(\mathbf{X}_{2j}(m_2),\mathbf{X}_{3j}(l_{j-1}|t_{j-1}),\hat{\mathbf{Y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1}))\\&\sim\prod_{i=1}^n p_{X_2}(x_{2,(j-1)n+i})\,p_{X_3|U}(x_{3,(j-1)n+i}|u_{(j-1)n+i})\\&\qquad\cdot p_{\hat{Y}_3|\tilde{Y}_3,X_3,U}(\hat{y}_{3,(j-1)n+i}|\tilde{y}_{3,(j-1)n+i},\\&\qquad\qquad x_{3,(j-1)n+i},u_{(j-1)n+i})\end{aligned}$$

conditionally independent of $(\mathbf{U}_j,\tilde{\mathbf{Y}}_{3j},\mathbf{X}_{1j},\mathbf{Y}_{1j})$. Hence, by the joint typicality lemma, $\mathsf{P}(\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\le 2^{-n(I_2-\delta(\epsilon))}$ where

$$\begin{aligned}I_2:=&\,I(X_2,X_3;\tilde{Y}_3,Y_1|X_1,U)\\&+I(\hat{Y}_3;X_1,X_2,Y_1|X_3,U,\tilde{Y}_3).\end{aligned}$$

Finally, if $m_2\neq 1$, $t_{j-1}=1$, and $l_{j-1}=1$, then

$$\mathbf{X}_{2j}(m_2)\sim\prod_{i=1}^n p_{X_2}(x_{2,(j-1)n+i}),$$

independent of $(\mathbf{U}_j,\mathbf{X}_{3j},\hat{\mathbf{Y}}_{3j},\tilde{\mathbf{Y}}_{3j},\mathbf{X}_{1j},\mathbf{Y}_{1j})$. Again, by the joint typicality lemma, $\mathsf{P}(\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\le 2^{-n(I_3-\delta(\epsilon))}$ where $I_3:=I(X_2;\hat{Y}_3,\tilde{Y}_3,Y_1|X_1,X_3,U)$.

Let b_1, b_2, b_3, and b_4 denote the number of $(t_j\neq 1,l_j\neq 1)$, $(t_j\neq 1,l_j=1)$, $(t_j=1,l_j\neq 1)$, and $(t_j=1,l_j=1)$ in the sequence of index pairs $(l_1,t_1),\ldots,(l_{b-1},t_{b-1})$, respectively. Note that $\mathbf{b}:=b_1,\ldots,b_4$ depends on the value of (t_j,l_j), $j\in[1:b-1]$, however, for notational convenience the relation is suppressed. Therefore

$$\begin{aligned}&\sum_{t^b,l^b}\prod_{j=2}^b\mathsf{P}(\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\\&=\sum_{t_b,l_b}\sum_{t^{b-1},l^{b-1}}\prod_{j=2}^b\mathsf{P}(\mathcal{A}_j(m_2,t_{j-1}^j,l_{j-1}^j))\\&\le\sum_{t_b,l_b}\sum_{\mathbf{b}}K(\mathbf{b})\,2^{b_1n(\hat{R}+\tilde{R})}\,2^{b_2n\tilde{R}}\,2^{b_3n\hat{R}}\\&\qquad\cdot 2^{-n(b_1I_1+b_2I_1+b_3I_2+b_4I_3-(b-1)\delta(\epsilon))}\\&=\sum_{\mathbf{b}}K(\mathbf{b})\sum_{t_b,l_b}2^{-n\left(b_1(I_1-\hat{R}-\tilde{R})+b_2(I_1-\tilde{R})+b_3(I_2-\hat{R})\right)}\\&\qquad\cdot 2^{-n(b_4I_3-(b-1)\delta(\epsilon))}\\&\le 4^{b-1}2^{n(\hat{R}+\tilde{R})}2^{-n(b-1)\left(\min\{I_1-\hat{R}-\tilde{R},\,I_1-\tilde{R},\,I_2-\hat{R},\,I_3\}-\delta(\epsilon)\right)}\\&=4^{b-1}2^{n(\hat{R}+\tilde{R})}2^{-n(b-1)\left(\min\{I_1-\hat{R}-\tilde{R},\,I_2-\hat{R},\,I_3\}-\delta(\epsilon)\right)}\end{aligned}$$

where the summation of positive integers $\mathbf{b}$ are over those such that $\sum_{i=1}^4 b_i=b-1$ and $K(\mathbf{b})$ is the number of all possible sequences of index pairs (t_j,l_j), $j\in[1:b-1]$ with the number of each corresponding element equal to $\mathbf{b}$. Thus, by the union of events bound, $\mathsf{P}(\cup_{m_2\neq 1}\mathcal{E}_{2,m_2})\to 0$ as $n\to\infty$, provided that

$$\begin{aligned}R_2<&\frac{b-1}{b}(\min\{I_1-\hat{R}-\tilde{R},\,I_2-\hat{R},\,I_3\}-\delta(\epsilon))\\&-\frac{1}{b}(\hat{R}+\tilde{R}).\end{aligned}$$

Finally, by letting $b\to\infty$, we have

$$\begin{aligned}R_2<&\,I(X_2,X_3;Y_1|X_1)+I(\tilde{Y}_3;X_1,X_2,X_3,Y_1|U)\\&+I(\hat{Y}_3;X_1,X_2,Y_1|X_3,U,\tilde{Y}_3)-\tilde{R}-\hat{R}-\delta(\epsilon),\qquad(6)\\R_2<&\,I(X_2,X_3;Y_1,\tilde{Y}_3|X_1,U)\\&+I(\hat{Y}_3;X_1,X_2,Y_1|X_3,U,\tilde{Y}_3)-\hat{R}-\delta(\epsilon),\qquad(7)\\R_2<&\,I(X_2;\hat{Y}_3,\tilde{Y}_3,Y_1|X_1,X_3,U)-\delta(\epsilon).\qquad(8)\end{aligned}$$

By exchanging the user indices and neglecting l_{j-1}, l_j, $\mathbf{X}_{3j}(l_{j-1}|t_{j-1})$ and $\hat{\mathbf{Y}}_{3j}(l_j|t_j,l_{j-1},t_{j-1})$ from the previous analysis of decoder 1, we can further show that $\mathsf{P}(\cup_{m_1\neq 1}\mathcal{E}_{1,m_1})\to 0$ as $n\to\infty$ provided that

$$\begin{aligned}R_1<&\,I(X_1,U;Y_2|X_2)\\&+I(\tilde{Y}_3;X_1,X_2,Y_2|U)-\tilde{R}-\delta(\epsilon),\qquad(9)\\R_1<&\,I(X_1;\tilde{Y}_3,Y_2|X_2,U)-\delta(\epsilon).\qquad(10)\end{aligned}$$

Finally, gathering inequalities (3), (4), (6), (7), (8), (9), and (10), using the Fourier–Motzkin procedure to eliminate $\hat{R}$ and $\tilde{R}$, getting rid of inactive inequalities, and time sharing, we obtain $\mathcal{R}_1$ which concludes the proof. ∎

III. Gaussian Two-Way Relay Channel

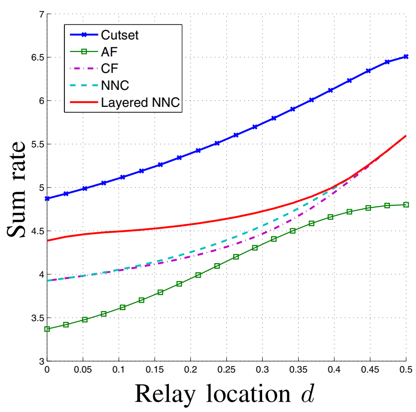

Fig. 1. Comparison of coding schemes for $g_{12} = g_{21} = 1$ and $P = 10$.

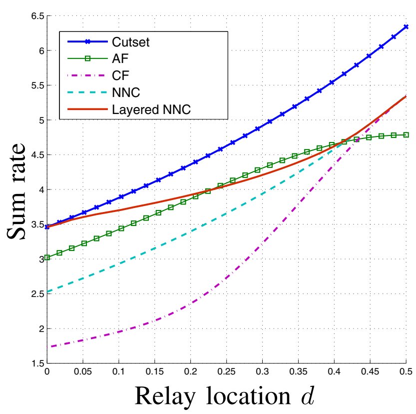

Fig. 2. Comparison of coding schemes for $g_{12} = g_{21} = 0$ and $P = 10$.

In the following, we evaluate Theorem 2 for the Gaussian two-way relay channel

$$\begin{aligned} Y_1 &= g_{21}X_2 + g_{31}X_3 + Z_1,\\ Y_2 &= g_{12}X_1 + g_{32}X_3 + Z_2,\\ Y_3 &= g_{13}X_1 + g_{23}X_2 + Z_3, \end{aligned} \tag{11}$$

where Z_k, $k \in \{1,2,3\}$ are independent Gaussian random variables with zero mean and unit variance and g_{12}, g_{21}, g_{13}, g_{23}, g_{31}, and g_{32} are the channel gains. We further assume average power constraint P on each sender X_k, $k \in \{1,2,3\}$.

Let $Q = \emptyset$, X_1 and X_2 be i.i.d. Gaussian with zero mean and variance P. Let $U \sim \mathrm{N}(0, \alpha P)$ and $X_3 = V + U$ where $V \sim \mathrm{N}(0, \bar{\alpha}P)$ independent of U. Further, let $\hat{Y}_3 = Y_3 + \hat{Z}_3$ and $\tilde{Y}_3 = \hat{Y}_3 + \tilde{Z}_3$, where $\hat{Z}_3$ and $\tilde{Z}_3$ are independent Gaussians with $\hat{Z} \sim \mathrm{N}(0, \hat{\sigma}^2)$ and $\tilde{Z} \sim \mathrm{N}(0, \tilde{\sigma}^2)$. Thus, $Y_3 \to \hat{Y} \to \tilde{Y}$ form a Markov chain (Note that the Markovity does not incur any loss among the set of Gaussian test channels). Then, $\mathcal{R}_1$ simplifies to the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< \mathrm{C}\left(\frac{g_{13}^2P(1+g_{32}^2\bar{\alpha}P) + (1+\hat{\sigma}^2+\tilde{\sigma}^2)g_{12}^2P}{(1+\hat{\sigma}^2+\tilde{\sigma}^2)(1+g_{32}^2\bar{\alpha}P)}\right),\\ R_1 &< \mathrm{C}\left(\frac{g_{12}^2P + g_{32}^2\alpha P}{1+g_{32}^2\bar{\alpha}P}\right) - \mathrm{C}\left(\frac{1}{\hat{\sigma}^2+\tilde{\sigma}^2}\right),\\ R_2 &< \mathrm{C}\left(\frac{g_{23}^2P + (1+\hat{\sigma}^2)g_{21}^2P}{1+\hat{\sigma}^2}\right),\\ R_2 &< \mathrm{C}\left(g_{21}^2P + g_{31}^2P\right) - \mathrm{C}(1/\hat{\sigma}^2),\\ R_2 &< \mathrm{C}\left(\frac{(1+g_{31}^2\bar{\alpha}P)g_{23}^2P + (1+\hat{\sigma}^2+\tilde{\sigma}^2)(g_{21}^2+g_{31}^2\bar{\alpha})P}{1+\hat{\sigma}^2+\tilde{\sigma}^2}\right)\\ &\quad - \mathrm{C}\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2(1+\hat{\sigma}^2+\tilde{\sigma}^2)}\right) \end{aligned}$$

for some $\tilde{\sigma}^2, \hat{\sigma}^2 > 0$ and $0 \le \alpha \le 1$, where $\bar{\alpha} = 1-\alpha$ and $\mathrm{C}(x) = \frac{1}{2}\log(1+x)$. On the other hand, by setting $Q = \emptyset$ and $\hat{Y}_3 = Y_3 + \hat{Z}$ with $\hat{Z} \sim \mathrm{N}(0, \sigma^2)$, the noisy network coding inner bound simplifies to the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< \mathrm{C}\left(\frac{g_{13}^2P + (1+\sigma^2)g_{12}^2P}{1+\sigma^2}\right),\\ R_1 &< \mathrm{C}(g_{12}^2P + g_{32}^2P) - \mathrm{C}(1/\sigma^2),\\ R_2 &< \mathrm{C}\left(\frac{g_{23}^2P + (1+\sigma^2)g_{21}^2P}{1+\sigma^2}\right),\\ R_2 &< \mathrm{C}(g_{21}^2P_1 + g_{31}^2P) - \mathrm{C}(1/\sigma^2) \end{aligned}$$

for some $\sigma^2 > 0$.

We further compare with two schemes presented in [6], compress–forward and amplify–forward. The compress–forward inner bound is given by the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< \mathrm{C}\left(\frac{g_{13}^2P + (1+\sigma^2)g_{12}^2P}{1+\sigma^2}\right),\\ R_2 &< \mathrm{C}\left(\frac{g_{23}^2P + (1+\sigma^2)g_{21}^2P}{1+\sigma^2}\right) \end{aligned}$$

for some

$$\begin{aligned} \sigma^2 &\ge \frac{(1+g_{12}^2P)(1+g_{13}^2P) - (g_{12}g_{13}P)^2}{\min\{g_{32}^2, g_{31}^2\}P},\\ \sigma^2 &\ge \frac{(1+g_{21}^2P)(1+g_{23}^2P) - (g_{21}g_{23}P)^2}{\min\{g_{32}^2, g_{31}^2\}P} \end{aligned}$$

and the amplify–forward inner bound is given by the rate pairs (R_1, R_2) such that $R_k < \frac{1}{2}\log\left(\frac{a_k + \sqrt{a_k^2 - b_k^2}}{2}\right)$, $k \in \{1,2\}$ for some $\alpha \le \sqrt{P/(g_{13}^2P + g_{23}^2P + 1)}$, where $a_1 := 1 + \frac{P(g_{12}^2 + \alpha^2 g_{32}^2 g_{13}^2)}{g_{32}^2\alpha^2+1}$, $a_2 := 1 + \frac{P(g_{21}^2 + \alpha^2 g_{31}^2 g_{23}^2)}{g_{31}^2\alpha^2+1}$, $b_1 := \frac{2P\alpha g_{32}g_{13}g_{12}}{g_{32}^2\alpha^2+1}$, and $b_2 := \frac{2P\alpha g_{31}g_{23}g_{21}}{g_{31}^2\alpha^2+1}$.

In Figures 1 and 2, we compare the performance of the proposed scheme (layered NNC) with noisy network coding (NNC), compress–forward (CF), and amplify–forward (AF) for a geometric model. The channel gains are given by $g_{13} = g_{31} = d^{-\gamma/2}$ and $g_{23} = g_{32} = (1-d)^{-\gamma/2}$, where $d \in [0,1]$ is the location of the relay node between nodes 1 and 2 (which are unit distance apart) and $\gamma = 3$ is the path loss exponent. Figure 1 depicts the case where $g_{21} = g_{12} = 1$ (that is, there is a direct link between nodes 1 and 2) and Figure 2 depicts the case $g_{21} = g_{12} = 0$ (that is, there is no direct link between nodes 1 and 2). The first case can be interpreted as the Gaussian two-way relay channel where all nodes transmit in the single frequency band while the second case corresponds to the frequency division between sender–relay and relay–destination (BC) links.

Figure 1 shows that layered noisy network coding outperforms all other schemes where the performance gain increases with channel asymmetry. In Figure 2, noisy network coding provides a substantial gain over compress–forward when the channel is asymmetric and layered noisy network coding provides additional gain. In particular, when the relay node is co-located with node 1 ($d = 0$), layered noisy network coding achieves the capacity.

As a final note, both noisy network coding schemes are the only schemes that are within 1 bit of the cutset bound, independent of the channel gains and power constraint. For the case without direct links between node 1 and 2 the gap can be improved to 1/2 bit (an achievable scheme using lattice codes was also shown to be within 1/2 bit [8] for the case without direct links between source–destination pairs).

ACKNOWLEDGMENTS

The work was supported in part by the DARPA ITMANET program, NSF CAREER grant CCF-0747111, and MKE/IITA IT R&D program 2008-F-004-02.

APPENDIX

Let (T_{j-1}, L_{j-1}) be the index pair chosen at block $j-1$. Define

$$\mathcal{B}_j := \{(t_j, l_j) : (\hat{\mathbf{Y}}_{3j}(l_j|t_j, L_{j-1}, T_{j-1}), \tilde{\mathbf{Y}}_{3j}(t_j|T_{j-1}), \mathbf{X}_{3j}(L_{j-1}|T_{j-1}), \mathbf{U}_j(T_{j-1}), \mathbf{Y}_{3j}) \in \mathcal{T}_{\epsilon'}^{(n)}\}.$$

Then the probability of the event of interest can be bounded as

$$\mathsf{P}\{|\mathcal{B}_j| = 0\} \le \mathsf{P}\{(|\mathcal{B}_j| - \mathsf{E}\,|\mathcal{B}_j|)^2 \ge (\mathsf{E}\,|\mathcal{B}_j|)^2\} \le \frac{\mathrm{Var}(|\mathcal{B}_j|)}{(\mathsf{E}\,|\mathcal{B}_j|)^2}$$

by Chebychev's inequality. Using indicator random variables, $|\mathcal{B}_j| = \sum_{t_j=1}^{2^{n\tilde{R}}} \sum_{l_j=1}^{2^{n\hat{R}}} E(t_j, l_j)$, where $E(t_j, l_j) = 1$ if

$$(\hat{\mathbf{Y}}_{3j}(l_j|t_j, L_{j-1}, T_{j-1}), \tilde{\mathbf{Y}}_{3j}(t_j|T_{j-1}), \mathbf{X}_{3j}(L_{j-1}|T_{j-1}), \mathbf{U}_j(T_{j-1}), \mathbf{Y}_{3j}) \in \mathcal{T}_{\epsilon'}^{(n)}$$

and $E(t_j, l_j) = 0$, otherwise.

Let

$$\begin{aligned}
p_1 &:= \mathsf{P}\{E(1,1) = 1\}, \\
p_2 &:= \mathsf{P}\{E(1,1) = 1, E(2,1) = 1\}, \\
p_3 &:= \mathsf{P}\{E(1,1) = 1, E(1,2) = 1\}, \\
p_4 &:= \mathsf{P}\{E(1,1) = 1, E(2,2) = 1\} = p_1^2.
\end{aligned}$$

Then, $E(|\mathcal{B}_j|) = \sum_{t_j, l_j} \mathsf{P}\{E(t_j, l_j) = 1\} = 2^{n(\hat{R}+\tilde{R})} p_1$. Hence,

$$\begin{aligned}
&E(|\mathcal{B}_j|^2) \\
&= \sum_{t_j, l_j} \mathsf{P}\{E(t_j, l_j) = 1\} \\
&\quad + \sum_{t_j, l_j} \sum_{t_j' \ne t_j} \mathsf{P}\{E(t_j, l_j) = 1, E(t_j', l_j) = 1\} \\
&\quad + \sum_{t_j, l_j} \sum_{l_j' \ne l_j} \mathsf{P}\{E(t_j, l_j) = 1, E(t_j, l_j') = 1\} \\
&\quad + \sum_{t_j, l_j} \sum_{t_j' \ne t_j} \sum_{l_j' \ne l_j} \mathsf{P}\{E(t_j, l_j) = 1, E(t_j', l_j') = 1\} \\
&\le 2^{n(\hat{R}+\tilde{R})} p_1 + 2^{n(\hat{R}+2\tilde{R})} p_2 + 2^{n(2\hat{R}+\tilde{R})} p_3 + 2^{2n(\hat{R}+\tilde{R})} p_1^2.
\end{aligned}$$

Hence, $\mathrm{Var}(|\mathcal{B}_j|) \le 2^{n(\hat{R}+\tilde{R})} p_1 + 2^{n(\hat{R}+2\tilde{R})} p_2 + 2^{n(2\hat{R}+\tilde{R})} p_3$. Now by the joint typicality lemma, we have

$$\begin{aligned}
p_1 &\ge 2^{-n(I(\tilde{Y}_3; X_3, Y_3|U) + I(\hat{Y}_3; Y_3|X_3, U, \tilde{Y}_3) + \delta(\epsilon'))}, \\
p_2 &\le 2^{-n(2I(\tilde{Y}_3; X_3, Y_3|U) + 2I(\hat{Y}_3; Y_3|X_3, U, \tilde{Y}_3) - \delta(\epsilon'))}, \\
p_3 &\le 2^{-n(I(\tilde{Y}_3; X_3, Y_3|U) + 2I(\hat{Y}_3; Y_3|X_3, U, \tilde{Y}_3) - \delta(\epsilon'))}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\mathrm{Var}(|\mathcal{B}_j|)}{\mathsf{E}(|\mathcal{B}_j|)^2} &\le 2^{-n(\hat{R}+\tilde{R}-I(\tilde{Y}_3; X_3, Y_3|U) - I(\hat{Y}_3; Y_3|X_3, U, \tilde{Y}_3) - 3\delta(\epsilon'))} \\
&\quad + 2^{-n(\hat{R}-3\delta(\epsilon'))} + 2^{-n(\tilde{R}-I(\tilde{Y}_3; X_3, Y_3|U) - 3\delta(\epsilon'))},
\end{aligned}$$

which tends to zero as $n \to \infty$, provided that

$$\begin{aligned}
\tilde{R} &> I(\tilde{Y}_3; X_3, Y_3|U) - \delta(\epsilon'), \\
\tilde{R} + \hat{R} &> I(\tilde{Y}_3; X_3, Y_3|U) + I(\hat{Y}_3; Y_3|X_3, U, \tilde{Y}_3) - \delta(\epsilon').
\end{aligned}$$

REFERENCES

[1] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," 2010, submitted to *IEEE Trans. Inf. Theory,* 2010. [Online]. Available: http://arxiv.org/abs/1002.3188

[2] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow," 2007, submitted to *IEEE Trans. Inf. Theory,* 2007. [Online]. Available: http://arxiv.org/abs/0710.3781/

[3] A. F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 789–804, 2006.

[4] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[5] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[6] B. Rankov and A. Wittneben, "Achievable rate region for the two-way relay channel," in *Proc. IEEE International Symposium on Information Theory*, Seattle, WA, July 2006.

[7] S. Katti, I. Maric, A. Goldsmith, D. Katabi, and M. Médard, "Joint relaying and network coding in wireless networks," in *Proc. IEEE International Symposium on Information Theory*, Nice, France, June 2007, pp. 1101–1105.

[8] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel within 1/2 bit," to appear in *IEEE Trans. Inf. Theory*.

[9] A. El Gamal and Y.-H. Kim, "Lecture notes on network information theory," 2010. [Online]. Available: http://arxiv.org/abs/1001.3404