Lossy-Forward Relaying for Lossy Communications: Rate-Distortion and Outage Probability Analyses

Wensheng Lin, *Student Member, IEEE*, Shen Qian, and Tad Matsumoto, *Fellow, IEEE*

***Abstract*—This paper presents an in-depth performance analysis of lossy communications in a single-relay system, where the recovered information is not necessarily lossless in both the relay and the destination. In this system, the relay continues transmitting the sequence with source-relay link errors to the destination even if errors are detected after decoding, i.e., so-called Lossy-Forward (LF) strategy. The problem can be decomposed into two parts as follows: a point-to-point coding problem in the source-relay (S-R) link, and a lossy source coding problem with a LF relay in the source-destination (S-D) and relay-destination (R-D) links. To begin with, we derive the admissible rate region of the lossy source coding problem with a LF relay for a specified distortion requirement. Then, we focus on the analysis of outage probability over block Rayleigh fading channels. Finally, a practical encoding/decoding scheme is proposed for the evaluation of system performance by computer simulations. Due to the suboptimal channel coding and incomplete utilization of joint typicality, the theoretical performance cannot be achieved in the simulation; however, the tendency of curves in simulations matches that in theoretical calculation.**

***Index Terms*—Relaying system, lossy-forward, distributed lossy source coding, rate-distortion, outage probability.**

I. Introduction

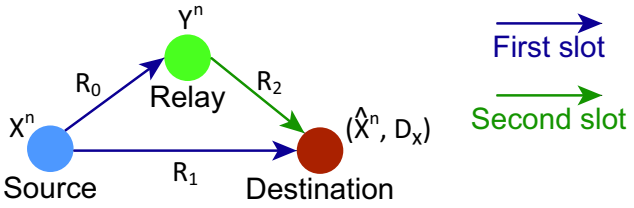


Fig. 1. The simplest system model of a lossy relaying system.

In big data era, transmissions with high fidelity are not always required in wireless cooperative communications networks, such as Internet-of-Things (IoT) networks. Here, we are interested in a basic model of wireless cooperative communications networks as shown in Fig. 1. A source broadcasts the sequence X^n to a destination and a relay, and the destination aims to recover the source sequence with the assistance of the relay. Due to the condition of wireless channels in practical systems, the source-relay (S-R), source-destination (S-D) and relay-destination (R-D) links have to satisfy the rates R_0, R_1 and R_2, respectively. If the capacity constraint on the S-R link is relatively strict, the relay cannot forward the message correctly. Once errors are detected in the decoded data sequence, the traditional Decode-and-Forward (DF) scheme discards the data sequence without forwarding to the destination. However, from the viewpoint of multiterminal source coding, the relay sequence containing intra-link errors has correlation with the source sequence as well. Therefore, the relay still continues to send the error-corrupted sequence Y^n to the destination, which is referred to as Lossy-Forward (LF) [1], [2]. With this method, the destination can refine the final estimate of the source sequence with the side information provided from the relay despite the link rate of relay channel. Compared to DF, LF also reduces the complexity of the relay, because error detection is not needed. More significantly, unlike Amplify-and-Forward (AF), LF does not require amplification of the received analog signal at the relay, which eliminates the well-recognized disadvantage with AF, i.e., noise enhancement and nonlinear distortion. In contrast to the hard decision in the LF relay, soft information relaying (SIR) schemes [3]–[5] encode and forward the soft information to the destination. SIR can both keep the soft information and have distributed coding gains; however, it is difficult to *re-arrange* the signal point at the relay, such as using higher order modulation, for improving the bandwidth efficiency.

From the aspect of whole system, the destination is not able to losslessly reconstruct the source sequence, if the rate triplet (R_0, R_1, R_2), supported by the channel conditions in the S-R, S-D, and R-D links, respectively, does not satisfy the admissible rate region [6]. Actually, lossy reconstructions $\hat{X}^n$ with its distortion level not larger than D_X are also acceptable as exemplified in IoT or sensor application networks. For conciseness, the lossy communications with LF are called lossy LF relaying. With a specified acceptable distortion requirement, we can reduce the power consumption or transmission bandwidth by lossy compression than lossless communication. Thus, the trade-off between the link rates and the expected distortion degree is a very interesting topic in the big data era, especially for numerous electronic devices, of which power is supplied by small battery.

To date, a number of scholars have made efforts to investigate LF. Base on the Slepian-Wolf Theorem [7], Hu and Li [8] for the first time proposed a novel relaying strategy, i.e., LF, to help the destination recover data losslessly. In [9], Cheng *et al.* derived the outage probability for a LF relaying system with three nodes communicating through block Rayleigh fading channels. Qian *et al.* [10] made a comparison of outage probability under spatially and temporally correlated fading among LF, DF and Adaptive Decode-and-Forward (ADF). Then, in [11], Qian *et al.* analyzed the theoretical performance of a LF system with three nodes suffering from independent block Nakagami-m fading. As for the practical techniques related to LF, researchers in [12]–[14] provided diverse coding schemes based on the turbo code [15]. Brulatout *et al.* [16] presented a medium access control (MAC) layer protocol

This work is funded in part by China Scholarship Council (CSC) and in part by JAIST Core-to-Core Program. This work has been also performed in part under JSPS Kakenhi (B)15H04007.

which cooperates with LF techniques in physical layer. In [17], Wolf *et al.* designed an optimal power allocation scheme among a source and two LF relays by taking into account outage probability.

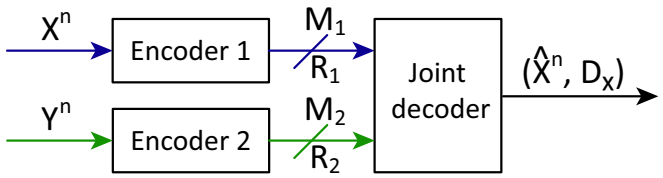


Fig. 2. The multiterminal source coding problem composed of the S-D and R-D links.

Nevertheless, the performance analysis has not been finished yet for the communication systems with lossy reconstructions allowed at the destination, i.e., lossy LF relaying. Notice that it is difficult to directly derive the explicit expression of the distortions resulting from the rate constraints on wireless channels. Shannon provided a way to equivalently determine the distortion corresponding to the channel capacity, i.e., compressing the data sequence by lossy source coding to satisfy the channel capacity which is achievable by lossless channel coding[1]. Since the source coding and channel coding are separately performed, this idea is referred to as the Shannon's lossy source-channel separation theorem [18], [19]. Likewise, to analyze the performance of lossy LF relaying, we start from the fundamentals of multiterminal source coding, which usually requires separate compression at encoders and joint decompression at a common decoder. Since the relay receives data only from the source, the analysis of the S-R link can be easily handled by the Shannon's lossy source-channel separation theorem for point-to-point communication. Regarding the remaining S-D and R-D links, we can first consider a multiterminal source coding problem illustrated in Fig. 2, where the encoder of X and the encoder of Y have to compress the sequences X^n and Y^n into codewords M_1 and M_2 at the rates R_1 and R_2, respectively. Then, we apply the Shannon's lossy source-channel separation theorem and equivalently calculate the distortion due to the channel conditions.

Determining the optimal trade-off between the link rate/rates and the expected distortion belongs to the category of rate-distortion analysis, which identifies the minimum rate requirements to achieve a specified value of distortion. Therefore, we can determine the final distortion inspired by the previous works regarding the rate-distortion function/region of multiterminal source coding [20]–[23]. Ahlswede and Korner derived the rate region for a system that requires the source information to be losslessly recovered with the assistance of a helper in [20]. In the case lossy reconstruction is acceptable, in [21], Wyner and Ziv characterized the rate-distortion function for the lossy source coding problem with uncompressed side information available in decoder. Berger [22] and Tung [23] determined the outer and inner bounds on the achievable rate-distortion region of multiterminal source coding problem with two sources. In [24], Jana and Blahut further extended the Berger-Tung bounds and Wyner-Ziv theorem to a general framework with many sources and one link of uncompressed side information. However, the admissible rate region of the lossy source coding problem with a LF relay is not given by a strict proof yet. We start from the derivation of the admissible rate region for the problem shown in Fig. 2, and then analyze the outage probability over Rayleigh fading channels for the relaying system shown in Fig. 1.

The contributions of this paper are summarized as follows:

- This paper derives the admissible rate region for lossy source coding problem with a LF relay through the proofs of achievability and the converse. In the case of binary sources, we further calculate and present the admissible rate region with a specified distortion requirement.
- Subsequently, based on the derived admissible rate region, we investigate the outage probability when block Rayleigh fading channels is implemented in the relaying system. The numerical results demonstrate the relationship of outage probability to average signal-to-noise ratio (SNR), expected distortion and relay location.
- Moreover, we design a practical encoding/decoding scheme to verify the tendency of theoretical results through computer simulations. Even though there is an obvious gap between the simulation and theoretical results, they show the similar shape of curves. Especially for strict distortion requirements, the performance of the proposed coding scheme is very close to the theoretical limit. We also discuss the reasons that make the simulation results away from the theoretical results.

The outline of the rest of the paper is as follows. Section II describes the problem to be solved under mathematical framework. In Section III, we derive the sufficient and necessary conditions of the admissible rate region for lossy transmissions with the aid of a LF relay. Then, we exploit the derived admissible rate region to analyze the outage probability for block Rayleigh fading channels in Section IV. In addition, Section V provides a design of practical encoding/coding scheme and make a comparison of outage probability between the simulated frame error rate (FER) and theoretical outage probability. Finally, this paper is concluded in Section VI.

II. Problem Statement

The theoretical performance analysis for the system illustrated in Fig. 1 can follow a lossy source coding problem and then Shannon's lossy source-channel separation theorem. In the following, we introduce the source coding problem to be solved and the channel model to be used in the outage probability analysis.

A. Source Coding Problem

Notations. The random variables and their realizations are denoted by uppercase and lowercase letters, respectively. The

[1]By utilizing the duality between source coding and channel coding, the information loss due to channel conditions can be equivalently analyzed by lossy source coding, followed by lossless transmission through wireless channels. Eventually, compression is not performed by the encoder, but fading variation may reduce the rate supported by the channel. However, in theoretical distortion analysis, we can formulate the problem in this way for simplicity and without loss of generality.

finite alphabets of a variable are denoted by calligraphic letters $\mathcal{X}, \mathcal{Y}, \cdots$. The entropy of a random variable X with probability mass function (PMF) $p(x)$ is defined as

$$H(X) = -\sum_{x\in\mathcal{X}} p(x)\log p(x). \tag{1}$$

The mutual information between two random variables X and Y is defined as

$$I(X;Y) = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}. \tag{2}$$

As mentioned above, the transmission in the S-R link is a point-to-point source coding problem, while the transmissions in the S-D and R-D links belong to a multiterminal source coding problem. Since the point-to-point source coding problem is already solved by Shannon in [19], we only focus on the lossy source coding problem with a LF relay depicted in Fig. 2.

By taking values from a finite alphabet $\mathcal{X}$ for each time index t, a common discrete memoryless source X generates independent and identically distributed (i.i.d.) sequence $x^n = \{x(t)\}_{t=1}^n$. The encoder of X encodes the sequence x^n by mapping it into an index as:

$$\varphi_1 : \mathcal{X}^n \mapsto \mathcal{M}_1 = \{1, 2, \cdots, 2^{nR_1}\}. \tag{3}$$

Since the relay sequence $y^n = \{y(t)\}_{t=1}^n$ is an error-corrupted version of x^n, y^n is also an i.i.d. sequence with each bits belonging to a finite alphabet $\mathcal{Y}$. Similar to the encoder of X, the encoder of Y encodes the sequence y^n by assigning an index according to the mapping rule:

$$\varphi_2 : \mathcal{Y}^n \mapsto \mathcal{M}_2 = \{1, 2, \cdots, 2^{nR_2}\}. \tag{4}$$

The joint decoder in the destination node starts decoding after receiving the encoder outputs $\varphi_1(x^n)$ and $\varphi_2(y^n)$. Unlike the distributed compression in encoders, the joint decoder constructs the estimate $\hat{x}^n$ from the index $\varphi_1(x^n)$ with the assistance of the compressed side information $\varphi_2(y^n)$. The recovering progress is implemented by the following mapping as:

$$\psi : \mathcal{M}_1 \times \mathcal{M}_2 \mapsto \mathcal{X}^n. \tag{5}$$

Due to the possible deviation of x from $\hat{x}$, a distortion measure $d_X : \mathcal{X}\times\mathcal{X} \mapsto [0, d_{X,\max}]$ is defined to describe the distortion level between x and its estimate $\hat{x}$. In particular, the Hamming distortion measure is defined for binary sources as

$$d_X(x(t),\hat{x}(t)) = \begin{cases} 1, & \text{if } x(t)\neq\hat{x}(t), \\ 0, & \text{if } x(t)=\hat{x}(t). \end{cases} \tag{6}$$

For the whole sequence, the average distortion between x^n and $\hat{x}^n$ is denoted by

$$d_X(x^n,\hat{x}^n) = \frac{1}{n}\sum_{t=1}^{n} d_X(x(t),\hat{x}(t)). \tag{7}$$

With an acceptable distortion value D_X, the rate region $\mathcal{R}(D_X)$, consisting of all admissible rate pairs (R_1, R_2), is defined as

$$\begin{aligned}\mathcal{R}(D_X) = \{(R_1,R_2) : (R_1,R_2) \text{ is admissible such that } \\ \lim_{n\to\infty}\mathrm{E}[d_X(x^n,\hat{x}^n)] \le D_X+\epsilon, \\ \text{for any } \epsilon>0\}.\end{aligned} \tag{8}$$

For the point-to-point communication in the S-R link, the distortion between X and Y depends on R_0. For the cooperative communications in the S-D and R-D links, the final distortion depends on R_1, R_2 and the correlations between X and Y. Therefore, the final distortion is eventually determined by R_0, R_1 and R_2. After determining the admissible rate region $\mathcal{R}(D_X)$ and the correlations between X and Y, we can obtain the relationship between the final distortion and channel capacities of all three links by utilizing the Shannon's lossy source-channel separation theorem.

Given a set of channel capacities for three links, we can calculate the expected minimum distortion based on the derived admissible rate region. If the expected distortion is larger than a specified distortion requirement, the communications are not reliable and outage event occurs. The channel capacities are random variables in fading channels, and hence the outage event randomly occurs with a probability, which is referred to as the outage probability. With a specified channel model, we can obtain the distributions of channel capacities and further calculate the outage probability, i.e., the probability that the instantaneous channel capacities cannot satisfy the distortion requirement.

B. Channel Model

To make equations more concise, we denote variables for the S-R, S-D and R-D links with subscripts 0, 1 and 2, respectively. The S-R, S-D and R-D links are assumed to suffer from independent block Rayleigh fading, with the complex channel gains as h_0, h_1 and h_2, respectively. Therefore, h_i follows the two dimensional Gaussian distribution.

For the t-th symbol $x_{\mathrm{S}}(t)$ encoded and modulated in the source node, the received signals via the S-R and S-D links are expressed as

$$x_i(t) = \sqrt{G_i}h_i x_{\mathrm{S}}(t) + z_i(t), \text{ for } i\in\{0,1\}, \tag{9}$$

where G_i represents the geometric gains, and z_i denotes the zero-mean additive white Gaussian noise (AWGN) in the corresponding link. Similarly, with the symbol $y_{\mathrm{R}}(t)$ encoded and modulated in the relay node, the destination node receives the signal

$$y_2(t) = \sqrt{G_2}h_2 y_{\mathrm{R}}(t) + z_2(t). \tag{10}$$

Let $E_X = \mathrm{E}[|x_{\mathrm{S}}(t)|^2]$ and $E_Y = \mathrm{E}[|y_{\mathrm{R}}(t)|^2]$ be the transmitting symbol energy, and the variances of z_i be all equal to $N_0/2$ per dimension. Then, the average SNRs are calculated by

$$\overline{\gamma}_i = G_i\cdot\mathrm{E}[|h_i|^2]\cdot\frac{E_X}{N_0}, \text{ for } i\in\{0,1\}, \tag{11}$$

$$\overline{\gamma}_2 = G_2\cdot\mathrm{E}[|h_2|^2]\cdot\frac{E_Y}{N_0}. \tag{12}$$

And the instantaneous SNRs can be expressed as

$$\gamma_i = |h_i|^2 \cdot \overline{\gamma}_i, \text{ for } i \in \{0, 1, 2\}. \tag{13}$$

Then, we can obtain the probability density function (PDF) of instantaneous SNR γ_i as

$$f(\gamma_i) = \frac{1}{\overline{\gamma}_i} \exp(-\frac{\gamma_i}{\overline{\gamma}_i}), \text{ for } i \in \{0, 1, 2\}. \tag{14}$$

For the purpose of simplicity, we assume that the channel state information (CSI) is only available at the receiver sides, and the effect of shadowing is not taken into account.

III. Admissible Rate Region

In this section, we first present the main result of the admissible rate region with General Sources for lossy-LF relaying in *Theorem 1*. Then, we calculate the admissible rate region by rate-distortion coding with binary sources based on *Theorem 1* for the whole system.

A. Admissible Rate Region with General Sources

Theorem 1: Let (X, Y) be a 2-component discrete memoryless source and $d_X(x, \hat{x})$ be a distortion measure. The admissible rate region with acceptable distortion D_X for lossy source coding of X with LF relaying is the set of rate pairs (R_1, R_2) such that

$$R_1 \geq I(X; U|V), \tag{15}$$

$$R_2 \geq I(Y; V), \tag{16}$$

for some conditional PMF $p(u|x)p(v|y)$ and function $\hat{x}(u, v)$ such that $\mathrm{E}[d_X(X, \hat{X})] \leq D_X$, with $U \to X \to Y \to V$ forming a Markov chain.

U and V are auxiliary variables which represent the compressed information of X and Y, respectively. $\hat{X}$ is the lossy recovery of X, which is reconstructed from the compressed information U of the source X and the compressed side information V of the relay information Y. For (16) in the R-D link, the minimum rate R_2 cannot be smaller than the information about Y obtained from the compressed information V, i.e., the mutual information $I(Y; V)$. To better understand (15) in the S-D link, we first consider the case without the assistance of the compressed side information V. Similar to (16), R_1 should be larger than or equal to $I(X; U)$ without the aid of V. Then, by utilizing the compressed side information V in joint decoding, V becomes an already known condition, and R_1 can be further reduced to $I(X; U|V)$.

The proofs of achievability and the converse for *Theorem 1* are presented in Appendix A and Appendix B, respectively.

B. Admissible Rate Region with Binary Sources

In order to draw a precise shape of admissible rate region, we need a specified distribution of source. Since digital signals are quite often assumed in LF, we start to use binary source as an instance in the following. Consider a binary source $X \sim \mathrm{Bern}(0.5)$, it is easy to find that Y, U and V also follow the $\mathrm{Bern}(0.5)$ distribution separately. In order to derive the relationship between R_0 and the distortion occurring in the S-R link, we can equivalently calculate the correlations between X and Y based on the Shannon's lossy source-channel separation theorem. To satisfy the channel capacity by lossy source coding, we have

$$\begin{aligned} R_0 &\geq I(X; Y) \\ &= H(X) - H(X|Y) \\ &= 1 - H_b(p), \end{aligned} \tag{17}$$

where $H_b(\cdot)$ denotes the binary entropy function, and p represents the crossover probability between X and Y. Likewise, for the R-D link with the crossover probability p' between Y and V, we have

$$R_2 \geq I(Y; V) \tag{18}$$

$$= 1 - H_b(p'). \tag{19}$$

From (15), for the S-D link, we have

$$R_1 \geq I(X; U|V) \tag{20}$$

$$= H(U|V) - H(U|X, V)$$

$$= H(U|V) - H(U|X) \tag{21}$$

$$= H_b(p' * p * D_X) - H_b(D_X), \tag{22}$$

where the operation * denotes the binary convolution process, i.e., $a * b = a(1 - b) + b(1 - a)$; (21) and (22) follows since $V \to Y \to X \to U$ forms a Markov chain with the crossover probabilities p', p and D_X, respectively.

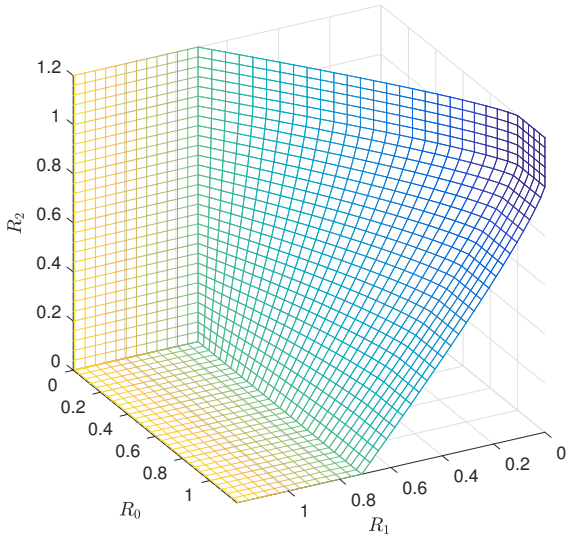
Consequently, we can obtain the admissible rate region with given distortion requirement as

$$\begin{cases} R_0 \geq 1 - H_b(p), \\ R_1 \geq H_b(p' * p * D_X) - H_b(D_X), \\ R_2 \geq 1 - H_b(p'). \end{cases} \tag{23}$$

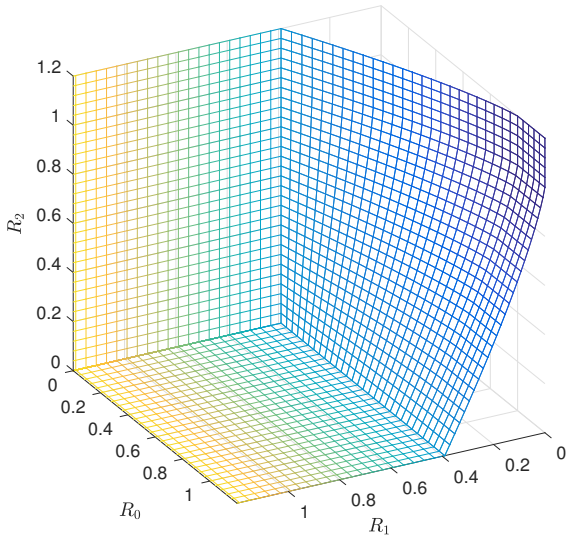
If the acceptable distortion is given, we can illustrate the admissible rate region by rate-distortion coding as in Fig. 3. It is remarkable that arbitrary R_0 and R_2 are admissible if R_1 is not less than $1 - H_b(D_X)$. Obviously, the compressed side information provided by the relay becomes redundant when R_1 is large enough for independent decoding. Hence, the acceptable distortion D_X can be easily satisfied by independent decoding for $R_1 \geq 1 - H_b(D_X)$ according to the lossy source coding theorem for point-to-point communication. Fig. 3(a) and Fig. 3(b) also demonstrate that the admissible rate region extends when the acceptable distortion becomes relatively large. Moreover, the part of surface is not flat for R_0, R_1 and R_2 all being less than 1. Because R_0 or R_2 needs more increase to compensate the decrease of R_1, due to the distortion propagating from the S-R link to the R-D link. Another interesting observation is that the admissible rate region is symmetric with respect to the plane of $R_0 = R_2$. Therefore, the S-R and R-D links have the same importance for system design, such as in determining power allocation and/or relay location.

IV. Outage Probability Analysis

In this section, we provide the derivation of outage probability for the lossy-LF relaying, based on the admissible rate region derived in (23).



(a) $D_X = 0.05$, $1 - H_b(D_X) \approx 0.7136$.



(b) $D_X = 0.15$, $1 - H_b(D_X) \approx 0.3902$.

Fig. 3. The admissible rate region by rate-distortion coding for specified distortion requirement.

A. Outage Event of Lossy LF Relaying

Here, we focus on the transmissions of the S-D and R-D links, which directly determine the occurrence of outage event, i.e., the destination cannot guarantee the reconstruction of X with the distortion smaller than D_X. For the influence of the S-R link, we treat the crossover probability p between X and Y as a parameter determined by R_0. By this means, we can obtain admissible rate region for given R_0 as illustrated in Fig. 4, where the rate pair (R_1, R_2) is achievable if (18) and (20) are satisfied. To facilitate the outage calculation provided later in this paper, the inadmissible rate region is divided to two sub-regions, α and β, as indicated by

$$\begin{cases} \alpha \triangleq \{0 \le R_1 \le I(X;U|Y), 0 \le R_2\}, \\ \beta \triangleq \{I(X;U|Y) \le R_1 \le I(X;U|V), \\ \qquad 0 \le R_2 \le H(Y)\}. \end{cases} \tag{24}$$

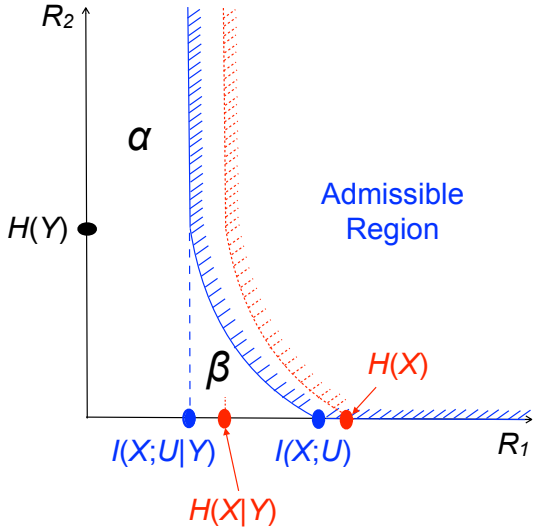


Fig. 4. The admissible rate region for X and Y; the blue solid line indicates the admissible rate region with acceptable distortion D_X; the red dashed line indicates the admissible rate region without distortion.

To conveniently calculate $I(X;U|Y)$, we can utilize the result in (22) by letting $V = Y$ and $p' = 0$. Consequently, we have

$$\begin{cases} \alpha \triangleq \{0 \le R_1 \le H_b(p * D_X) - H_b(D_X), 0 \le R_2\}, \\ \beta \triangleq \{H_b(p * D_X) - H_b(D_X) \\ \qquad \le R_1 \le H_b(p' * p * D_X) - H_b(D_X), \\ \qquad 0 \le R_2 \le 1\}. \end{cases} \tag{25}$$

Intuitively, the rate region defined in (22) indicates that:

1) For $H(Y) \le R_2$, Y can successfully decoded with $Y = V$, i.e., $p' = 0$. The transmission with distortion D_X can be supported as long as $R_1 \ge H_b(0 * p * D_X) - H_b(D_X) = H_b(p * D_X) - H_b(D_X)$, which reduces to the Wyner-Ziv theorem.
2) Even with $0 < R_2 < H(Y)$, Y can be partially recovered at the destination as V. V containing errors serves as the compressed side information for recovering X as long as $R_1 \ge H_b(p' * p * D_X) - H_b(D_X)$.
3) In the case $R_2 = 0$ ($p' = 0.5$), i.e., the R-D link is broken down, the conditions in (22) become to $R_1 \ge H_b(0.5 * p * D_X) - H_b(D_X) = 1 - H_b(D_X)$, which reduces to the classical rate-distortion function.

Based on the discussion above, (22) can be rewritten explicitly as

$$R_1 \ge \begin{cases} H_b(p * D_X) - H_b(D_X), & \text{for } H(Y) \le R_2, \\ H_b(p' * p * D_X) - H_b(D_X), & \text{for } 0 < R_2 < H(Y), \\ 1 - H_b(D_X), & \text{for } R_2 = 0. \end{cases} \tag{26}$$

With the help of the compressed side information V, the outage event occurs when the rate pair (R_1, R_2) falls inside the inadmissible regions, i.e., region α and β in Fig. 4. The

outage probability $\mathrm{P_{out}}$ can be defined by taking average over all the transmissions, which results in

$$\begin{aligned}
\mathrm{P_{out}} =& \Pr\{(R_1,R_2)\in\alpha\cup\beta\}\\
=& \Pr\{p=0,(R_1,R_2)\in\alpha\cup\beta\}\\
&+\Pr\{p\in(0,0.5],(R_1,R_2)\in\alpha\cup\beta\}\\
=& \Pr\{p=0,(R_1,R_2)\in\alpha\}\\
&+\Pr\{p=0,(R_1,R_2)\in\beta\}\\
&+\Pr\{p\in(0,0.5],(R_1,R_2)\in\alpha\}\\
&+\Pr\{p\in(0,0.5],(R_1,R_2)\in\beta\}\\
=& \Pr\{0\le R_1\le H_b(p*D_X)-H_b(D_X),0\le R_2,\\
&\quad p=0\}\\
&+\Pr\{H_b(p*D_X)-H_b(D_X)\\
&\quad\le R_1\le H_b(p'*p*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,p=0\}\\
&+\Pr\{0\le R_1\le H_b(p*D_X)-H_b(D_X),0\le R_2,\\
&\quad 0<p\le 0.5\}\\
&+\Pr\{H_b(p*D_X)-H_b(D_X)\\
&\quad\le R_1\le H_b(p'*p*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,0<p\le 0.5\}\\
=& P_{1,\alpha}+P_{1,\beta}+P_{2,\alpha}+P_{2,\beta},
\end{aligned}\tag{27}$$

where $P_{1,\alpha}$, $P_{1,\beta}$, $P_{2,\alpha}$ and $P_{2,\beta}$ are defined for conciseness. The first subscript 1 and 2 represent the events $p=0$ and $p\in(0,0.5]$, while the second subscript α and β represent that the rate pair (R_1,R_2) falls inside the region α and β, respectively.

B. Outage Derivation

For calculating the outage probability, first we establish the relationship between γ_i and R_i for $i\in\{0,1,2\}$. Since orthogonal transmissions are assumed in the system, from the Shannon's lossy source-channel separation theorem, the relationship between the instantaneous channel SNR γ_i and its corresponding rate constraint R_i are given by

$$R_i=\Theta_i(\gamma_i)=\begin{cases}\dfrac{C(\gamma_0)}{r_X}=\dfrac{E^n}{2r_X}\log_2\left(1+\dfrac{2\gamma_0}{E^n}\right), i=0,\\ \dfrac{C(\gamma_1)}{r_X}=\dfrac{E^n}{2r_X}\log_2\left(1+\dfrac{2\gamma_1}{E^n}\right), i=1,\\ \dfrac{C(\gamma_2)}{r_Y}=\dfrac{E^n}{2r_Y}\log_2\left(1+\dfrac{2\gamma_2}{E^n}\right), i=2,\end{cases}\tag{28}$$

where r_X and r_Y represent the channel coding rates for X^n and Y^n, respectively; $C(\cdot)$ is the Shannon capacity using Gaussian codebook, and E^n is the signaling dimensionality.

By combining the results with (17), the crossover probability p between X and Y can be expressed with the function of γ_0 as

$$p=H_b^{-1}\left[1-\Theta_0(\gamma_0)\right],\tag{29}$$

with $H_b^{-1}(\cdot)$ denoting the inverse function of $H_b(\cdot)$.

With the assumption that each link suffers from statistically independent block Rayleigh fading, each term of the outage probability expression in (27) can be further expressed as

$$\begin{aligned}
P_{1,\alpha} =& \Pr\{0\le R_1\le H_b(0*D_X)-H_b(D_X),0\le R_2,\\
&\quad p=0\}\\
=& \Pr\{0\le R_1\le 0,0\le R_2,p=0\}\\
=& \Pr\left\{\Theta_1^{-1}(0)\le\gamma_1\le\Theta_1^{-1}(0),\Theta_2^{-1}(0)\le\gamma_2,\right.\\
&\quad\left.\Theta_0^{-1}(1)\le\gamma_0\right\}\\
=& \int_{\Theta_2^{-1}(0)}^{\infty}d\gamma_2\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}(0)}d\gamma_1\\
&\cdot\int_{\Theta_0^{-1}(1)}^{\infty}f(\gamma_0)f(\gamma_1)f(\gamma_2)d\gamma_0,\\
=& 0,
\end{aligned}\tag{30}$$

$$\begin{aligned}
P_{1,\beta} =& \Pr\{H_b(0*D_X)-H_b(D_X)\\
&\quad\le R_1\le H_b(p'*0*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,p=0\}\\
=& \Pr\{0\le R_1\le H_b(p'*D_X)-H_b(D_X),\\
&\quad 0\le R_2\le 1,p=0\}\\
=& \Pr\left\{\Theta_1^{-1}(0)\le\gamma_1\right.\\
&\quad\le\Theta_1^{-1}[H_b(\xi(\gamma_2,D_X))-H_b(D_X)],\\
&\quad\left.\Theta_2^{-1}(0)\le\gamma_2\le\Theta_2^{-1}(1),\Theta_0^{-1}(1)\le\gamma_0\right\}\\
=& \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}d\gamma_2\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[H_b(\xi(\gamma_2,D_X))-H_b(D_X)]}d\gamma_1\\
&\cdot\int_{\Theta_0^{-1}(1)}^{\infty}f(\gamma_0)f(\gamma_1)f(\gamma_2)d\gamma_0\\
=& \frac{1}{\overline{\gamma}_2}\exp\left(-\frac{\Theta_0^{-1}(1)}{\overline{\gamma}_0}\right)\int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}\exp\left(-\frac{\gamma_2}{\overline{\gamma}_2}\right)\cdot\left[1\right.\\
&\left.-\exp\left(-\frac{\Theta_1^{-1}\{H_b[\xi(\gamma_2,D_X)]-H_b(D_X)\}}{\overline{\gamma}_1}\right)\right]d\gamma_2,
\end{aligned}\tag{31}$$

$$\begin{aligned}
P_{2,\alpha} =& \Pr\{0\le R_1\le H_b(p*D_X)-H_b(D_X),0\le R_2,\\
&\quad 0<p\le 0.5\}\\
=& \Pr\left\{\Theta_1^{-1}(0)\le\gamma_1\right.\\
&\quad\le\Theta_1^{-1}\{H_b\left[\xi(\gamma_0,D_X)\right]-H_b(D_X)\},\\
&\quad\left.\Theta_2^{-1}(0)\le\gamma_2,\Theta_0^{-1}(0)\le\gamma_0<\Theta_0^{-1}(1)\right\}\\
=& \int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)}d\gamma_0\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}\{H_b[\xi(\gamma_0,D_X)]-H_b(D_X)\}}d\gamma_1\\
&\cdot\int_{\Theta_2^{-1}(0)}^{\infty}f(\gamma_2)f(\gamma_1)f(\gamma_0)d\gamma_2\\
=& \frac{1}{\overline{\gamma}_0}\exp\left(-\frac{\Theta_2^{-1}(0)}{\overline{\gamma}_2}\right)\int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)}\exp\left(-\frac{\gamma_0}{\overline{\gamma}_0}\right)\cdot\left[1\right.\\
&\left.-\exp\left(-\frac{\Theta_1^{-1}\left\{H_b\left[\xi(\gamma_0,D_X)\right]-H_b(D_X)\right\}}{\overline{\gamma}_1}\right)\right]d\gamma_0,
\end{aligned}\tag{32}$$

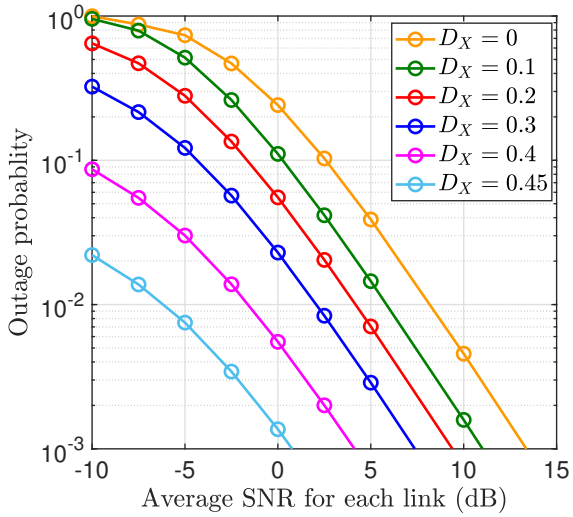


Fig. 5. Outage probability with different acceptable distortions.

and

$$
\begin{aligned}
&P_{2,\beta} \\
&= \Pr\{H_b(p * D_X) - H_b(D_X) \\
&\qquad \le R_1 \le H_b(p' * p * D_X) - H_b(D_X), \\
&\qquad 0 \le R_2 \le 1, 0 < p \le 0.5\} \\
&= \Pr\big\{\Theta_1^{-1}\{H_b[\xi(\gamma_0, D_X)] - H_b(D_X)\} \\
&\qquad \le \gamma_1 \le \Theta_1^{-1}\{H_b[\mu(\gamma_2, \gamma_0) * D_X] - H_b(D_X)\}, \\
&\qquad \Theta_2^{-1}(0) \le \gamma_2 \le \Theta_2^{-1}(1), \\
&\qquad \Theta_0^{-1}(0) \le \gamma_0 < \Theta_0^{-1}(1)\big\} \\
&= \int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)} d\gamma_0 \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)} d\gamma_2 \\
&\quad \cdot \int_{\Theta_1^{-1}\{H_b[\xi(\gamma_0,D_X)]-H_b(D_X)\}}^{\Theta_1^{-1}\{H_b[\mu(\gamma_2,\gamma_0)*D_X]-H_b(D_X)\}} f(\gamma_1) f(\gamma_2) f(\gamma_0) d\gamma_1 \\
&= \frac{1}{\overline{\gamma}_0 \overline{\gamma}_2} \int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)} d\gamma_0 \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)} \exp\left(-\frac{\gamma_0}{\overline{\gamma}_0} - \frac{\gamma_2}{\overline{\gamma}_2}\right) \\
&\quad \cdot \left[\exp\left(-\frac{\Theta_1^{-1}\{H_b[\xi(\gamma_0, D_X)] - H_b(D_X)\}}{\overline{\gamma}_1}\right)\right. \\
&\quad \left. - \exp\left(-\frac{\Theta_1^{-1}\{H_b[\mu(\gamma_2, \gamma_0) * D_X] - H_b(D_X)\}}{\overline{\gamma}_1}\right)\right] d\gamma_2,
\end{aligned}
\tag{33}
$$

where $\Theta_i^{-1}(\cdot)$ denoting the inverse function of $\Theta_i(\cdot)$, $\xi(\gamma_i, \tilde{p}) = H_b^{-1}[1 - \Theta_i(\gamma_i)] * \tilde{p}$ and $\mu(\gamma_i, \gamma_j) = H_b^{-1}[1 - \Theta_i(\gamma_i)] * H_b^{-1}[1 - \Theta_j(\gamma_j)]$. Since there is not an explicit expression for the inverse of binary entropy function, it is hard to further calculate the integral and obtain a precise closed form. Instead, we utilize computer to calculate the numerical results for analyzing the outage probability.

C. Numerical Results

The performance of outage probabilities for specified acceptable distortion D_X is presented in Fig. 5, where average SNR is set at the same value for all three links. Clearly, the lossy LF relaying achieves lower outage probability with larger acceptable distortion D_X. It should be noticed that the outage probability equals to zero when $D_X = 0.5$. This is because that $D_X = 0.5$ indicates any distortion can be accepted at the destination, and therefore, there will be no more outage in this case.

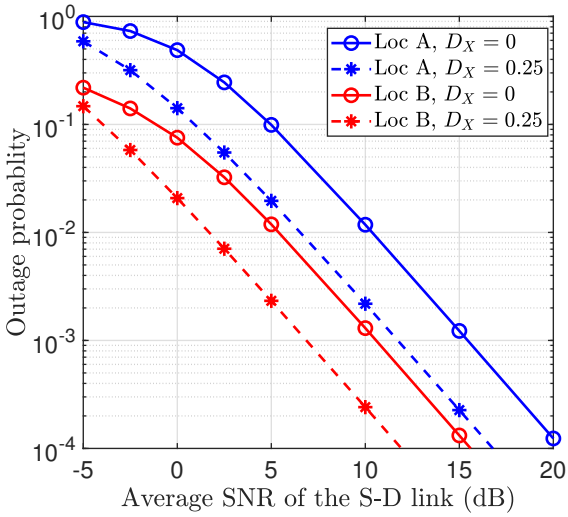


Fig. 6. Outage performances of the lossy LF relaying for different relay locations.

The outage curves of the lossy LF relaying are shown in Fig. 6 for two different relay location scenarios. With s_i for $i \in \{0, 1, 2\}$ denoting the distance of its corresponding link, we set $s_0 = s_1 = s_2$ in location scenario A (Loc A), while $s_0 = 0.25 s_1$ and $s_2 = 0.75 s_1$ in location scenario B (Loc B). In either the Loc A or Loc B, lower outage probability can be achieved by allowing distortion at the destination. Moreover, since the distances of the S-R and R-D links in Loc B are both smaller than that in Loc A, outage events occur with lower probability in Loc B than in Loc A.

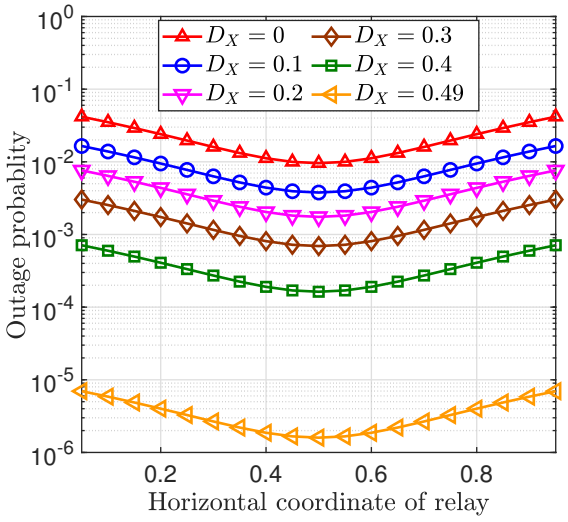


Fig. 7. The optimal relay positions of the lossy LF relaying, where $\overline{\gamma}_1 = 5$ dB.

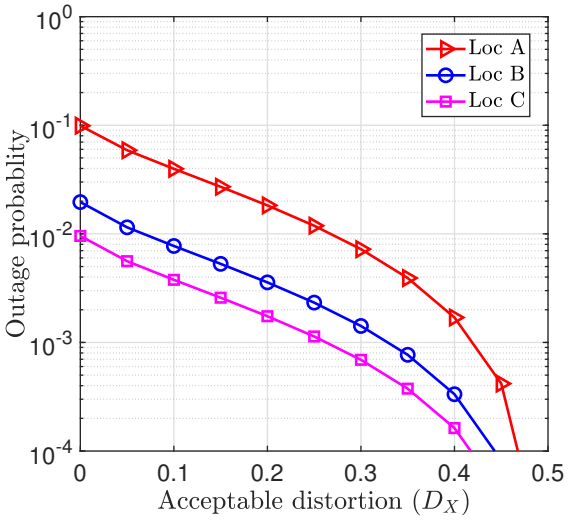


Fig. 8. Outage probability versus acceptable distortion where $\overline{\gamma}_1 = 5$ (dB).

Fig. 7 shows the impact of the relay location on the outage probability, with $\overline{\gamma}_1 = 5$ dB. The relay is located on the line between the source and the destination. It is found that the lowest outage probability can be achieved when the relay is located at the midpoint regardless the acceptable distortion. It is also observed that the outage curves are symmetric with respect to the midpoint of the S-D link. This is because in the lossy LF relaying, the errors due to the S-R link can be corrected at the destination, and therefore, the midpoint ($s_0 = s_2$) is the optimal point where the contributions of the S-R and R-D links are balanced. This phenomenon indicates that the S-R and R-D links are of the same significance for system design, which perfectly matches with the finding in Fig. 3.

Fig. 8 shows the outage probability versus the acceptable distortion D_X, with different relay location scenarios are considered. We set $s_0 = s_1 = s_2$ in Loc A, $s_0 = 0.25s_1$ and $s_2 = 0.75s_1$ in Loc B, and $s_0 = s_2 = 0.5s_1$ in Loc C. It is observed that when the relay at the same location, outage probability decreases as the acceptable distortion increases. It can also be seen from the figure that, the outage performance in Loc B is superior than that obtained in Loc A. This is because the quality of the S-R link in Loc B is better than that in Loc A, resulting in lower probability of the S-R link transmission failure. From intuitive discussion for Fig. 7, we can understand the fact that the lossy LF relaying shows the best outage performance in Loc C, since the relay is at the midpoint. Another interesting finding is that, the outage probability decreases almost linearly with D_X when the value of acceptable distortion is small (roughly less than 0.3); however, the outage probability decreases significantly when D_X is larger than 0.3. This observation, which results from the exact calculations of outage probability with diverse D_X, can explain the reason why the gap between the curve with $D_X = 0.4$ and that with $D_X = 0.49$ suddenly becomes large in Fig. 7.

V. Performance Evaluation

A. Simulation Design

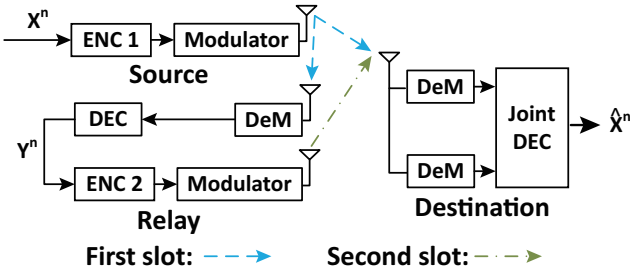
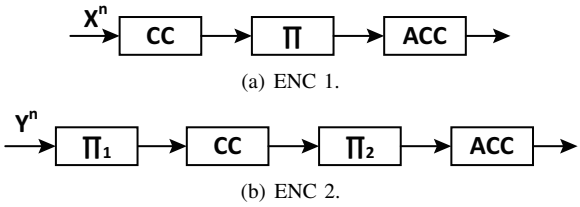


Fig. 9. The system model for simulation.

Here, we start to evaluate the system performance for a practical wireless communications network. As illustrated in Fig. 9, there are three nodes in the system containing source, relay and destination. In the first slot, the source node encodes sequence X^n by ENC 1 and broadcasts the modulated signal through Rayleigh channels. Then, the relay node decodes the received signal by DEC after demodulation (DeM) and makes hard decision into Y^n, while the destination node just stores the received signal. In the second slot, the relay node encodes Y^n by ENC 2 and subsequently sends the modulated signal to the destination node. As soon as the destination node receives the signal from the relay node, it starts to jointly decodes the received signals and finally outputs the estimate $\hat{X}^n$.



(a) ENC 1.

(b) ENC 2.

Fig. 10. The structure of encoders.

The structure of encoders is shown in Fig. 10[2]. In the source node, X^n is encoded by a convolutional code (CC) for the first step. For the sake of utilizing the principle of turbo code in decoding, CC is concatenated with an interleaver Π and an accumulator (ACC) [14]. To obtain the iteration gains between X^n and Y^n in joint decoding, Y^n is interleaved by Π_1 at the beginning of encoding. Then, the interleaved sequence is encoded by the same means as the process in the source node.

Fig. 11 depicts the structure of the joint decoder in the destination node. To begin with, the demodulated signal in each link is separately decoded by the decoder of ACC (ACC^{-1}) and the decoder of CC (CC^{-1}). In the local iteration, the extrinsic information is exchanged between ACC^{-1} and CC^{-1} via an interleaver Π and a deinterleaver Π^{-1}. After CC^{-1} outputs the *a posteriori* log-likelihood ratio (LLR^p)

[2]The purpose of the simulation is to compare the performance tendency with theoretical performance, and also with other forwarding schemes. Therefore, we choose relatively simple component codes to reduce simulation time. To approach the theoretical limit by utilizing stronger coding scheme is left as the future work.

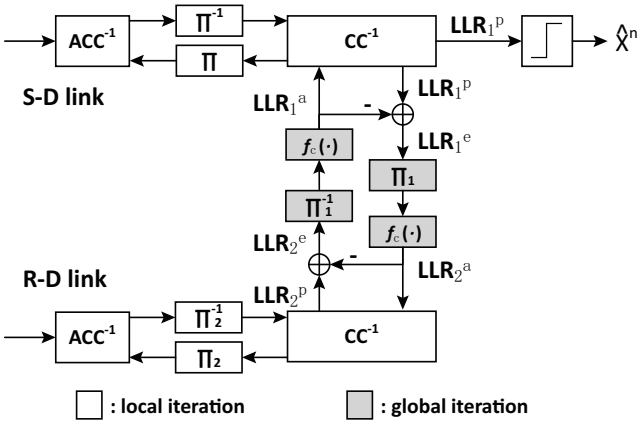


Fig. 11. The structure of joint decoder.

TABLE I
PARAMETER SETTINGS

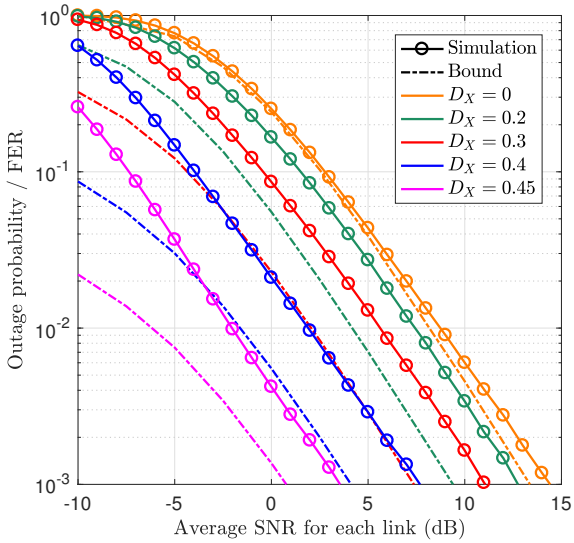
Parameter	Value
Frame length	10^4 bits
Number of Frames	10^6
Rate of CC	$1/2$
Generator polynomial of CC	$G = ([3, 2]3)_8$
Type of interleaver	random interleaver
Modulation method	BPSK
Decoding algorithm for CC	BCJR algorithm [29]
Maximum iteration time	30

at the end of local iteration, the joint decoder calculates the extrinsic LLR (LLR^e) by subtracting the *a priori* LLR (LLR^a) from LLR^p. When exchanging LLR between X^n and Y^n, we take the error probability of Y^n into consideration based on the correlation model [25]. The error probability of Y^n is first estimated by the algorithm proposed in [26], and then the *a priori* LLR is updated by the LLR updating function $f_c(\cdot)$ [27] with the extrinsic LLR as input. By this means, the relay information provides less extrinsic information if more errors exist in Y^n, and hence X^n is insulated from the errors in Y^n in joint decoding[3]. Due to the interleaving process on Y^n before CC, LLR_1^e should be interleaved by Π_1 and LLR_2^e should be deinterleaved by Π_1^{-1} when exchanging the extrinsic information in the global iteration. Finally, the estimate $\hat{X}^n$ is made by hard decision from LLR_1^p, if the maximum iteration time is exceeded or no more gains of the mutual information on LLR_1^p can be obtained in iterations.

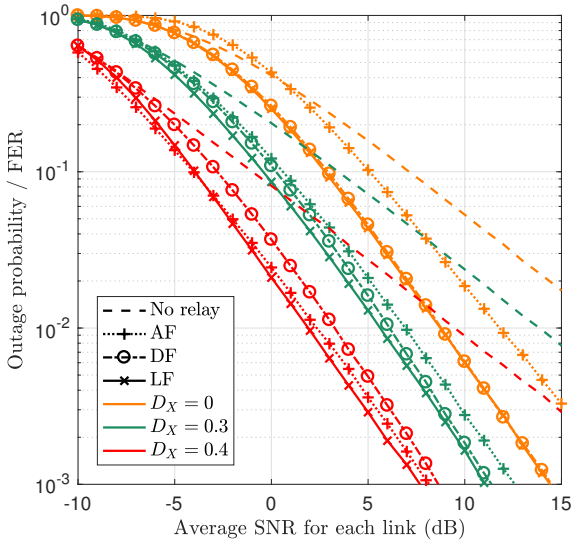
B. Simulation Results

The simulation result with parameter settings listed in Table I is shown in Fig. 12, which compares the theoretical outage probability and FER in simulation. For simplicity, average

[3]It should be emphasized that although the final distortion will not be worse if the relay always forwards the sequence to the destination, the performance gain becomes very small when the relay sequence contains too many errors. However, the relay still needs to consume the same power for forwarding, resulting in lower power efficiency. The trade-off between outage and energy efficiency may be handled by a thresholding technique [28]; however, it is out of the scope of this paper.



(a) Comparison between theoretical and simulation results.



(b) Comparison among different forwarding schemes.

Fig. 12. Outage probability in simulation.

SNR is set at the same value for the S-D, S-R and R-D links. In Fig. 12(a), it is clear that the simulation result has the same tendency and similar slope as the theoretical bound, even though there is an obvious gap between them. Moreover, the gap between the simulation and theoretical results becomes larger as D_X increases. This phenomenon indicates that the practical scheme used in simulation is more efficient when the distortion requirement is more strict. There are two major factors which result in the loss of system performance. First, notice that with relatively simple channel coding scheme, it is hard to achieve the Shannon limit, and hence there is also a gap between the FER in simulation and the theoretical outage probability of the network, as a whole. Another significant factor is that, the practical coding scheme in simulation cannot utilize the joint typicality as efficiently as the random binning coding scheme in the proof of achievability for *Theorem 1*.

Fig. 12(b) compares the practical performance for diverse

relaying schemes, including AF, DF, LF, and the case without relay. Obviously, the curves with a relay have the same decay of the performance curve independently of the relaying scheme, while the slope of the curve without a relay is less steep compared to the curves with a relay. This observation demonstrates the diversity gains achieved by introducing a relay, although there are some gaps between different relaying schemes. It is found that AF has a worse performance than DF and LF for relatively small distortion requirement, because DF and LF can eliminate the errors from the S-R link, while AF amplifies the signals along with the noise. In addition, by encoding again at the relay, the data received from the S-D and R-D links are equivalent to distributed turbo codes, and hence DF and LF have coding gains while AF cannot. Nonetheless, when the acceptable distortion becomes very large, e.g., $D_X = 0.4$, AF has a better performance than DF. The reason is that large D_X requires even lower SNR, which makes more errors exist in the decoding result at the relay; therefore, the DF relay discards the data sequences and stop forwarding more frequently. However, if the average SNR is not too small (larger than -4 dB), the system with a LF relay still has a lower outage probability than that with a AF relay, due to the utilization of correlations in the error-corrupted sequences. Only when the average SNR is very small and the acceptable distortion is very large, can AF achieve lower outage probability than LF. This is because with LF, hard decision will eliminate the information rather than the noise when the SNR is very small. In low average SNR region, the instantaneous SNR is frequently small, and hence the relay sequence contains too many errors resulting from the hard decision by LF. On the other hand, AF still keeps relatively large volume of soft information of the source sequence, if large distortion is acceptable.

VI. Conclusion

We have analyzed the performance of lossy LF relaying, where distortion is allowed in the destination with the assistance of a LF relay. To begin with, we divided the system into two sub problems, i.e., lossy point-to-point communication for the S-R link, and the multiterminal source coding problem for the S-D and R-D links. The sub problem for the S-R link can be easily solved by the Shannon's lossy source coding theorem and lossy source-channel separation theorem. Then, for the multiterminal source coding problem in the S-D and R-D links, we derive the admissible rate region through the proofs of achievability and the converse. We further determine the relationship between final distortion and the rate constraints due to the channel condition, by applying the Shannon's lossy source-channel separation theorem to the admissible rate region. Moreover, we analyze the outage probability for specified distortion requirements over block Rayleigh fading channels. Finally, we design a simulation system to evaluate the practical performance of FER. Comparing to the theoretical outage probability, we find that the tendency of simulation result matches with theoretical analysis. Especially for the case with strict distortion requirement, the FER in simulation is very close to the theoretical outage probability.

Appendix A
Proof of Achievability for Theorem 1

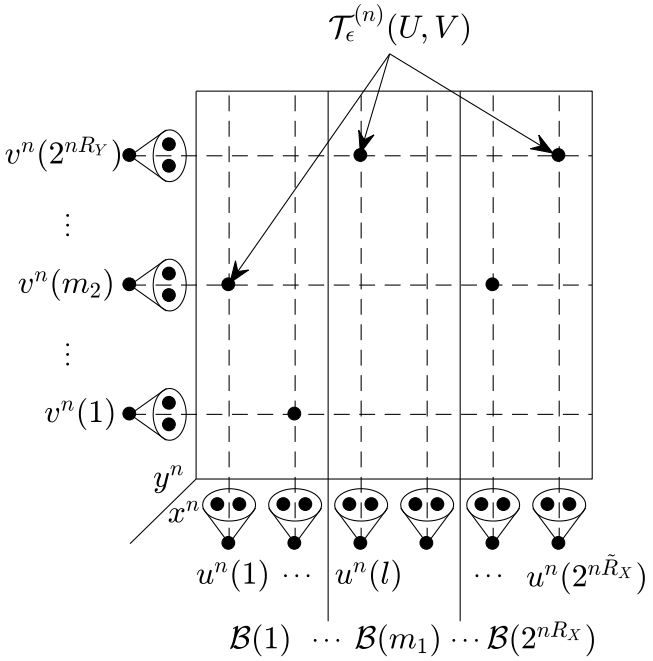


Fig. 13. The coding scheme for the proof of achievability.

For two correlated sequences, if one of the sequence is given, the possible alternatives of another sequence will also be determined, so that two sequences follow the joint PDF of two random variables. This property is referred to as joint typicality, which can be utilized to save coding rate for correlated sources. We use a random binning and joint typicality encoding scheme shown in Fig. 13, and analyze its expected distortion for the proof of achievability. In the following, we assume that $\epsilon'' < \epsilon' < \epsilon$.

Codebook generation. Fix a conditional PMF $p(u|x)p(v|y)$ and a function $\hat{x}(u,v)$ such that $\mathrm{E}[d_X(X,\hat{X})] \le D_X/(1+\epsilon)$. Let $\tilde{R}_1 \ge R_1$. Randomly and independently generate $2^{n\tilde{R}_1}$ sequences $u^n(l) \sim \prod_{t=1}^n p_U(u), l \in \mathcal{L} = \{1,2,\cdots,2^{n\tilde{R}_1}\}$. Partition the set of indices l into equal-size bins $\mathcal{B}(M_1) = \{(M_1-1)2^{n(\tilde{R}_1-R_1)}+1,\cdots,M_1 2^{n(\tilde{R}_1-R_1)}\}$. Note that this process is equivalent to vector quantization of a source [30]. Then, randomly and independently generate 2^{nR_2} sequences $v^n(m_2) \sim \prod_{t=1}^n p_V(v), m_2 \in \mathcal{M}_2 = \{1,2,\cdots,2^{nR_2}\}$. This codebook structure is utilized in the encoders and the decoder.

Encoding. Upon observing x^n, the encoder of X finds an index $l \in \mathcal{L}$ such that $(u^n(l),x^n) \in \mathcal{T}_{\epsilon''}^{(n)}$. If there is more than one such index l, the encoder of X selects the smallest one among them. If there is no such index l, the encoder of X sets $l=1$. Then, the encoder of Y finds an index m_2 such that $(v^n(m_2),y^n) \in \mathcal{T}_{\epsilon''}^{(n)}$. If there is more than one such index m_2, the encoder of Y selects the smallest one among them. If there is no such index m_2, the encoder of Y sets $m_2=1$. The encoder of X and the encoder of Y send the indices m_1 and m_2 such that $l \in \mathcal{B}(M_1)$, respectively.

Decoding. The joint decoder finds the unique index $\hat{l} \in \mathcal{B}(M_1)$ such that $(u^n(\hat{l}), v^n(m_2)) \in \mathcal{T}_\epsilon^{(n)}$. If there is such a unique index $\hat{l}$, the reconstruction is computed bit by bit as $\hat{x}_t(u_t(\hat{l}), v_t(m_2))$; otherwise, $\hat{x}^n$ is set to an arbitrary sequence in $\mathcal{X}^n$.

We start to analyze the expected distortion of this random binning scheme. Let L denote the index for the chosen sequence U^n, M_1 be the corresponding bin index, and $\hat{L}$ be the decoded index. Moreover, let M_2 denote the index for the chosen sequence V^n. Define the "error" event

$$\mathcal{E} = \{(U^n(\hat{L}), V^n(M_2), X^n, Y^n) \notin \mathcal{T}_\epsilon^{(n)}\}, \tag{34}$$

and consider the following events:

$$\mathcal{E}_1 = \{(U^n(l), X^n) \notin \mathcal{T}_{\epsilon''}^{(n)} \text{ for all } l \in \mathcal{L}\}, \tag{35}$$

$$\mathcal{E}_2 = \{(V^n(m_2), Y^n) \notin \mathcal{T}_{\epsilon''}^{(n)} \text{ for all } m_2 \in \mathcal{M}_2\}, \tag{36}$$

$$\mathcal{E}_3 = \{(U^n(L), X^n, Y^n) \notin \mathcal{T}_{\epsilon'}^{(n)}\}, \tag{37}$$

$$\mathcal{E}_4 = \{(U^n(L), X^n, V^n(M_2), Y^n) \notin \mathcal{T}_\epsilon^{(n)}\}, \tag{38}$$

$$\mathcal{E}_5 = \{(U^n(\tilde{l}), V^n(M_2)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{l} \in \mathcal{B}(M_1), \tilde{l} \neq L\}. \tag{39}$$

$\mathcal{E}_1$ and $\mathcal{E}_2$ represent that encoding error events happen in the encoder of X and the encoder of Y, respectively. $\mathcal{E}_4$ represents the failure of joint typicality decoding, while $\mathcal{E}_3$ is the sub event of $\mathcal{E}_4$. $\mathcal{E}_5$ occurs when there are more than one decoding result, i.e., a decoding error event also happens. Notice that the "error" event occurs only if $(U^n(L), X^n, V^n(M_2), Y^n) \notin \mathcal{T}_\epsilon^{(n)}$ or $\tilde{l} \neq L$. By the union of the events bound, we have

$$\mathrm{P}(\mathcal{E}) \leq \mathrm{P}(\mathcal{E}_1) + \mathrm{P}(\mathcal{E}_2) + \mathrm{P}(\mathcal{E}_1^c \cap \mathcal{E}_3) + \mathrm{P}(\mathcal{E}_3^c \cap \mathcal{E}_4) + \mathrm{P}(\mathcal{E}_5). \tag{40}$$

We bound each term as follows. First, by the covering lemma, $\mathrm{P}(\mathcal{E}_1)$ and $\mathrm{P}(\mathcal{E}_2)$ both tend to zero as $n \to \infty$ if

$$\tilde{R}_1 > I(X;U) + \delta(\epsilon''), \tag{41}$$

$$R_2 > I(Y;V) + \delta(\epsilon''). \tag{42}$$

Notice that $\mathcal{E}_1^c = \{(U^n(L), X^n) \in \mathcal{T}_{\epsilon''}^{(n)}\}$ and $Y^n | \{U^n(L) = u^n, X^n = x^n\} \sim \prod_{t=1}^n p_{Y|X}(y_t|x_t)$. By the conditional typicality lemma, $\mathrm{P}(\mathcal{E}_1^c \cap \mathcal{E}_3)$ tends to zero as $n \to \infty$.

To bound $\mathrm{P}(\mathcal{E}_3^c \cap \mathcal{E}_4)$, let $(u^n, x^n, y^n) \in \mathcal{T}_{\epsilon'}^{(n)}(U, X, Y)$, and consider

$$\begin{aligned} &\mathrm{P}\{V^n(M_2) = v^n | U^n(L) = u^n, X^n = x^n, Y^n = y^n\} &(43)\\ &= \mathrm{P}\{V^n(M_2) = v^n | Y^n = y^n\} \\ &= p(v^n|y^n). &(44)\end{aligned}$$

First, notice that by the covering lemma, $\mathrm{P}\{V^n(M_2) \in \mathcal{T}_{\epsilon'}^{(n)}(V|y^n) | Y^n = y^n\}$ converges to 1 as $n \to \infty$, i.e., $p(v^n|y^n)$ satisfies the first condition of the Markov lemma [6]. Then, similar to the Lemma 12.3 in [6] for the proof of the Berger-Tung inner bound, $p(v^n|y^n)$ also satisfies the second condition of the Markov lemma. Hence, according to the Markov lemma, we have

$$\begin{aligned} &\lim_{n\to\infty} \mathrm{P}\Big\{(u^n, x^n, y^n, V^n(M_2)) \in \mathcal{T}_\epsilon^{(n)} | U^n(L) = u^n, \\ &\qquad\qquad X^n = x^n, Y^n = y^n\} \\ &= 1, &(45)\end{aligned}$$

if $(u^n, x^n, y^n) \in \mathcal{T}_{\epsilon'}^{(n)}(U, X, Y)$ and $\epsilon' < \epsilon$ is sufficiently small. Hence, $\mathrm{P}(\mathcal{E}_3^c \cap \mathcal{E}_4)$ tends to zero as $n \to \infty$.

To bound $\mathrm{P}(\mathcal{E}_5)$, similar to the Lemma 11.1 in [6] for the achievability proof of the Wyner-Ziv Theorem, we have

$$\mathrm{P}(\mathcal{E}_5) \leq \mathrm{P}\{(U^n(\tilde{l}), V^n(M_2)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{l} \in \mathcal{B}(1)\}. \tag{46}$$

By the mutual packing lemma $\mathrm{P}(\mathcal{E}_5)$ tends to zero as $n \to \infty$, if

$$\tilde{R}_1 - R_1 < I(U;V) - \delta(\epsilon). \tag{47}$$

Further combining the inequalities (41) and (47), we have

$$\begin{aligned} R_1 &> \tilde{R}_1 - I(U;V) + \delta(\epsilon) \\ &> I(X;U) + \delta(\epsilon'') - I(U;V) + \delta(\epsilon) \\ &= I(X,V;U) - I(U;V) + \delta'(\epsilon) &(48)\\ &= I(X;U|V) + \delta'(\epsilon), &(49)\end{aligned}$$

where (48) follows since $U \to X \to Y \to V$ forms a Markov chain and by defining $\delta'(\epsilon) = \delta(\epsilon'') + \delta(\epsilon)$. Hence, we have shown that $\mathrm{P}(\mathcal{E})$ tends to zero as $n \to \infty$ if inequalities (42) and (49) are satisfied. Notice that $(U^n(L), V^n(M_2), X^n, Y^n) \in \mathcal{T}_\epsilon^{(n)}$ if there is no "error". Therefore, by the law of total expectation and the typical average lemma, the asymptotic distortion, averaged over the random codebook and encoding, is upper bounded as

$$\begin{aligned} &\lim_{n\to\infty} \sup \mathrm{E}[d_X(X^n, \hat{X}^n)] \\ &\leq \lim_{n\to\infty} \sup [d_{X,\max} \cdot \mathrm{P}(\mathcal{E}) \\ &\qquad + (1+\epsilon) \cdot \mathrm{E}[d_X(X, \hat{X})] \cdot \mathrm{P}(\mathcal{E}^c)] &(50)\\ &\leq D_X, &(51)\end{aligned}$$

if the inequalities in (42) and (49) are satisfied. Finally, by the continuity of mutual information and taking $\epsilon \to 0$, we complete the proof of achievability for *Theorem 1*.

Appendix B
Proof of the Converse for Theorem 1

First, consider

$$\begin{aligned} nR_1 &\geq H(M_1) \\ &\geq H(M_1|M_2) &(52)\\ &\geq I(X^n; M_1|M_2) \\ &= \sum_{t=1}^n I(X_t; M_1|M_2, X^{t-1}) &(53)\\ &\geq \sum_{t=1}^n I(X_t; M_1|M_2, X^{t-1}, Y^{t-1}) &(54)\\ &= \sum_{t=1}^n I(X_t; M_1, X^{t-1}, Y^{t-1}|M_2, X^{t-1}, Y^{t-1}), &(55)\end{aligned}$$

where (52) and (54) hold since the condition reduces entropy; (53) holds according to the chain rule for mutual information. By identifying $U_t = (M_1, X^{t-1}, Y^{t-1})$ and

$V_t = (M_2, X^{t-1}, Y^{t-1})$, noting that $U_t \to X_t \to Y_t$ and $X_t \to Y_t \to V_t$ form Markov chains, we have

$$nR_1 \geq \sum_{t=1}^{n} I(X_t; U_t|V_t). \tag{56}$$

Then, consider

$$\begin{aligned} nR_2 &\geq H(M_2) \\ &\geq I(Y^n; M_2) \\ &= \sum_{t=1}^{n} I(Y_t; M_2|Y^{t-1}) \\ &= \sum_{t=1}^{n} I(Y_t; M_2, Y^{t-1}) \qquad (57) \\ &= \sum_{t=1}^{n} I(Y_t; M_2, X^{t-1}, Y^{t-1}) \qquad (58) \\ &= \sum_{t=1}^{n} I(Y_t; V_t), \qquad (59) \end{aligned}$$

where (57) and (58) hold by that Y_t is independent of (X^{t-1}, Y^{t-1}) since X and Y are memoryless sources, and (59) holds by the same identifying of V_t as in the derivation of constraint on R_1. This completes the proof of the converse for *Theorem 1*.

REFERENCES

[1] X. Zhou, M. Cheng, X. He, and T. Matsumoto, "Exact and approximated outage probability analyses for decode-and-forward relaying system allowing intra-link errors," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 7062–7071, Dec. 2014.

[2] P.-S. Lu, X. Zhou, K. Anwar, and T. Matsumoto, "Joint adaptive network–channel coding for energy-efficient multiple-access relaying," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2298–2305, Jun. 2014.

[3] M. A. Karim, J. Yuan, Z. Chen, and J. Li, "Soft information relaying in fading channels," *IEEE Wireless Commun. Lett.*, vol. 1, no. 3, pp. 233–236, Jun. 2012.

[4] M. H. Azmi, J. Li, J. Yuan, and R. Malaney, "LDPC codes for soft decode-and-forward in half-duplex relay channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1402–1413, Aug. 2013.

[5] J. Li, M. A. Karim, J. Yuan, Z. Chen, Z. Lin, and B. Vucetic, "Novel soft information forwarding protocols in two-way relay channels," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2374–2381, Jun. 2013.

[6] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge Univ. Press, 2011.

[7] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[8] R. Hu and J. Li, "Exploiting Slepian-Wolf codes in wireless user cooperation," in *IEEE 6th Workshop on Signal Process. Advances in Wireless Commun.*, New York, USA, Jun. 2005, pp. 275–279.

[9] M. Cheng, K. Anwar, and T. Matsumoto, "Outage probability of a relay strategy allowing intra-link errors utilizing Slepian-Wolf theorem," *EURASIP J. on Advances in Signal Process.*, vol. 2013, no. 1, p. 34, Dec. 2013.

[10] S. Qian, J. He, M. Juntti, and T. Matsumoto, "Fading correlations for wireless cooperative communications: Diversity and coding gains," *IEEE Access*, Apr. 2017.

[11] S. Qian, X. Zhou, X. He, J. He, M. Juntti, and T. Matsumoto, "Performance analysis for lossy-forward relaying over Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10 035–10 043, Nov. 2017.

[12] B. Zhao and M. C. Valenti, "Distributed turbo coded diversity for relay channel," *Electron. Lett.*, vol. 39, no. 10, pp. 786–787, May 2003.

[13] H. H. Sneessens, J. Louveaux, and L. Vandendorpe, "Turbo-coded decode-and-forward strategy resilient to relay errors," in *IEEE Int. Conf. on Acoustics, Speech and Signal Process.*, Las Vegas, NV, USA, Apr. 2008, pp. 3213–3216.

[14] K. Anwar and T. Matsumoto, "Accumulator-assisted distributed turbo codes for relay systems exploiting source-relay correlation," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1114–1117, Jul. 2012.

[15] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[16] M. Brulatout, H. Khalife, V. Conan, S. Szott, M. Natkaniec, K. Kosek-Szott, and L. Prasnal, "A cooperative MAC protocol for lossy forwarding networks," in *Wireless Days (WD)*. Toulouse, France: IEEE, Mar. 2016, pp. 1–3.

[17] A. Wolf, M. Matthé, A. Festag, and G. Fettweis, "Outage based power allocation for a lossy forwarding two-relaying system," in *IEEE 20th Int. Workshop on Comput. Aided Modelling and Design of Communication Links and Networks*, Guildford, UK, Sep. 2015, pp. 267–272.

[18] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[19] ——, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec*, vol. 4, no. 142-163, p. 1, Mar. 1959.

[20] R. Ahlswede and J. Korner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 21, no. 6, pp. 629–637, Nov. 1975.

[21] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[22] T. Berger, "Multiterminal source coding," in *The Inform. Theory Approach to Commun., G. Longo, Ed.* New York: Springer-Verlag, 1978, pp. 171–231.

[23] S. Y. Tung, "Multiterminal source coding," Ph.D. dissertation, Sch. of Elect. Eng., Cornell Univ., Ithaca, New York, 1978.

[24] S. Jana and R. Blahut, "Canonical description for multiterminal source coding," in *IEEE Int. Symp. on Inform. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 697–701.

[25] J. Garcia-Frias and Y. Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 555–559, Apr. 2005.

[26] X. He, X. Zhou, K. Anwar, and T. Matsumoto, "Estimation of observation error probability in wireless sensor networks," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1073–1076, Jun. 2013.

[27] X. Zhou, X. He, K. Anwar, and T. Matsumoto, "GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems," *IEICE Trans. on Commun.*, vol. 95, no. 12, pp. 3654–3662, Dec. 2012.

[28] A. Irawan, K. Anwar, and T. Matsumoto, "Lossy forwarding HARQ for parallel relay networks," *Wireless Personal Commun.*, vol. 95, no. 2, pp. 915–930, Jul. 2017.

[29] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[30] A. Gersho and R. M. Gray, *Vector quantization and signal compression*. Springer Science & Business Media, 2012, vol. 159.