Impact of Network Coding on System Delay for Multi-source Multi-destination Scenarios

Zhiguo Ding, *Member, IEEE*, Zheng Ma, *Member, IEEE* and Kin K. Leung,

Fellow, IEEE

Abstract

Existing work has shown that random coding across multi-cast sessions can reduce the system delay significantly, however, such a scheme requires the strong assumption that each source has the priori information of other sources' messages. Actually the broadcasting nature of radio propagation can provide an opportunity to realize collaboration across sessions without causing much system overhead. In this paper, we propose the application of network coding to multi-source multi-destination (MSMD) scenarios and provide formal analysis for the improvement of system delay. In particular, two types of analytical results have been developed, one based on the outage probability and the other based on the use of practical convolutional codes. Monte-Carlo simulation results have also been provided to demonstrate the delay performance of the proposed network coded protocol.

I. INTRODUCTION

In wireless communication systems, it has been a challenging task to achieve robust transmission and small delay simultaneously, particularly for certain quality-of-service demanding services. Many traditional techniques may be efficient to improve the reception reliability, but have deteriorating effects on the system delay. For example, cooperative diversity has been recognized as a low cost and efficient method to combat the wireless unreliability caused by multipath fading, but the fact that relay transmission consumes extra bandwidth resource implies that some cooperative protocols could enlarge the system delay, i.e. the protocols in [1]. Originally developed to increase the capacity for wireline networks, network coding has recently received a lot of attentions and it has been shown to yield positive gain for delay performance in wireless communications [2], [3]. Traffic pattern is important for an efficient application of network coding to wireless communications. In [4]–[7], efficient network coded protocols have been developed for the two-way relaying channels, and in [8], [9] several cooperative multiple access transmission protocols

based on network coding have been proposed. In [10] the scenario with two source-destination pairs and one relay has been studied, where the achievable rates have been obtained. For such two-way relaying and multiple access channels, it has been demonstrated that the use of network coding combined with cooperative diversity can not only just increase reception reliability, but also improve system throughput and delay performance.

Interference channel, also known as multi-source multi-destination scenarios (MSMD), is one of the fundamental building blocks of wireless communications. Different to other traffic patterns, such as multiple access and two-way relaying channels, MSMD is severely interference limited. A traditional way for such scenarios is to serve multiple source-destination pairs one at each time, which is intended to avoid interference but not efficient and robust. In case that deep channel fading happens for certain pairs, a large number of retransmission are required and hence large system delay becomes unavoidable. [11] is one of the first tries to apply network coding to MSMD scenarios and the key idea of [11] is to encourage source nodes collaborating with each other. By applying random coding across the multiple sessions, each source transmits a mixture of all source messages, which is analogue to network coding. Comparing with non-cooperative schemes, random coding across multiple sessions may bring more interference, however, it is shown in [11] that such a coded scheme is much more reliable, which is due to the reason that each source transmission can serve all destination simultaneously. However, each source needs the priori information of other sources' information, so all source messages can be mixed together. This could be a strong assumption since source nodes are not co-located and extra system overhead is required for information exchanging between source nodes.

In this paper, our aim is to study the impact of network coding on the system delay in multi-source multi-destination scenarios. Different to [11], we are interested in how to achieve the collaboration among multiple source-destination pairs without causing too much system overhead. In specific, the two-hop transmission strategy is focused and the use of intermediate relays is introduced into MSMD scenarios. Instead of asking one source transmitting each time, all source nodes will broadcast their messages simultaneously. At relays, mixtures of all source messages are observed because of the broadcasting nature of radio propagation. Rather than to ask relays to separate the mixture, the idea of network coding is used and relays are allowed to forward the mixtures. In such a way, random coding across multiple sessions is realized without causing any system overhead. To further improve the performance of the proposed transmission protocol, the opportunistic use of relays is also used to exploit multi-user diversity. Various decoding methods can be utilized at destinations to solve the mixture, where the criteria of zero forcing are used due to its simplicity. Two types of analytical results are developed for the overall system delay. One is based on the outage probability, which can be tightly bounded by the error probability of

maximum likelihood for infinite length of data blocks and high SNR. In addition to such a theoretical upper bound, we also provide analytical results based on the use of practical convolutional codes. Monte-Carlo simulation results have also provided to demonstrate the performance of the proposed network coded transmission protocol with comparisons to comparable schemes.

This paper is organized as follows. Transmission strategies for one hop MSMD scenarios are first discussed in Section II to highlight the importance of the collaboration among multiple pairs. Then in Section III we will focus on the two-hop MSMD scenarios, where a network coding assisted transmission protocol is proposed. Analytical results for system delay will be provided in Section IV, and Monte-Carlo simulation results are provided in Section V. Finally, concluding remarks are given in Section VI.

II. A Brief Study on Transmission Strategies for One Hop Scenarios

Consider that there are M pairs of sources and destinations, and there is no intermediate node in the context of one hop scenarios. Time division duplexing is considered in this paper for its simplicity. The baseline transmission strategy is the traditional non-cooperative scheme, where one source transmits its head-of-line packet continuously until the packet is correctly received by its corresponding destination. At each time slot, only one source is transmitting and the other sources will take their turn to transmit their head-of-line packets in a round-robin way.

Recently proposed in [11], an alternative strategy is to apply random linear coding across the multicast sessions, where each source will transmit a linear combination of the M head-of-line packets. As a result, each destination has to decode all M packets in order to extract its own message from the mixture. Although more demanding requirements have been imposed on the receive capability of the destinations, it was shown in [11] that more robust performance and less retransmissions can be obtained benefited by such source cooperation.

However, the strategy of coding across sessions relies on the assumption that each source has the priori information of the packets the other sources transmit, which is not realistic in practice. A straightforward realization of the cross-coding strategy is to adopt the direct transmission strategy for the first few time slots, during which each source broadcasts its message and listen to the transmissions of other nodes when it is idle. After each source has obtained the knowledge of other sources' messages, the strategy of coding across sessions can then be applied. Fig. 1 shows the performance of these three one-hop transmission strategies in terms of the averaged number of required retransmissions. The number of sources is $M = 2$. The outage probability is used to determine whether a message can be decoded correctly as in [1] where the target data rate is $R = 1$ bits/s/Hz. As can be seen from the figure, the practical protocol of coding across sessions can offer positive performance gain compared with the non-cooperative scheme, however, it can only realize a portion of the performance achieved by the transmission strategy with the ideal
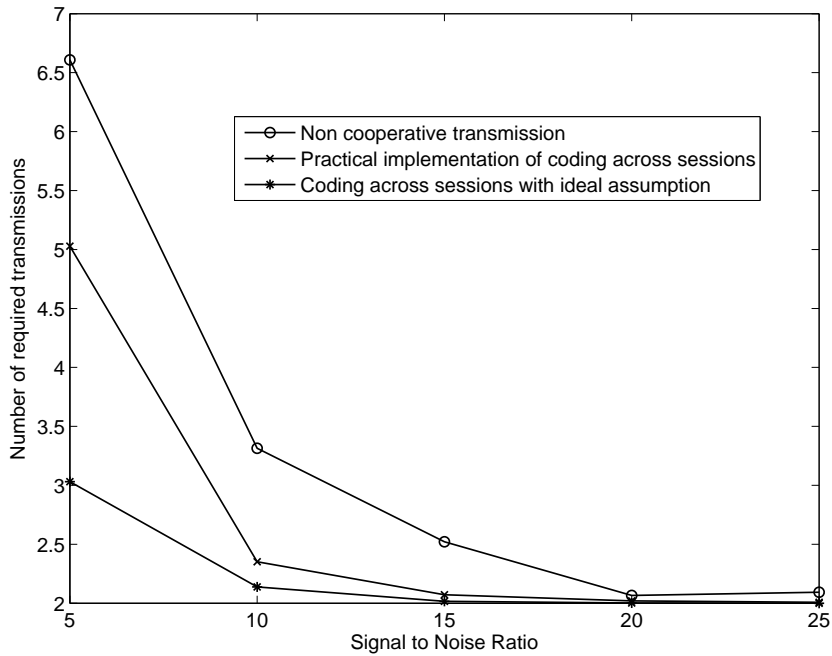
Fig. 1. Averaged number of required retransmissions vs SNR. The number of the source nodes is $M = 2$.

assumption. Such a performance loss is mainly due to the fact that extra bandwidth resource has to be consumed before the sources can benefit from the random coding across the pairs.

The message delivered in Fig. 1 triggers an important question about how to efficiently implement the idea of coding across multicast sessions. In the context of one hop scenario, it seems inevitable to consume extra bandwidth resource as the price of source cooperation. However, in the context of two-hop scenarios, the broadcasting nature of radio propagation enables us to exploit the benefit of random coding across sessions without suffering the loss of bandwidth efficiency, as shown in the next section.

III. Two-Hop Transmissions: Cooperative and Non-Cooperative Strategies

Consider a two-hop communication scenario with M source-destination pairs and L intermediate relays. Each source aims to send its head-of-line packet to its corresponding receiver. Assume that there is no direct link between the sources and destinations as in [12]. Such an assumption can be justified by the fact that a node two hops away from a transmitter is most likely in the strong interference range of another transmitter. Furthermore, a node two hops away from a transmitter along the route can transmit by using overlapping medium access control [13] to increase the overall system throughput, and hence it is not able to receive signals due to the half duplexing constraint. The time division duplexing mode is used, which means the channel responses are reciprocal between the transmitter and receiver. Note that it is straightforward to extend the proposed network coded transmission protocol to the scenario where the destination hears the transmitter two hops away.

A traditional way for such a scenario is to apply the non-cooperative strategy, where time division multiple access is used and there is no cooperation among multiple source-destination pairs. Each source will first try to deliver its message to an intermediate node which decodes the message and then forwards it to the corresponding destination. Despite its simplicity, the non-cooperative strategy suffers loss of bandwidth efficiency and reception reliability as shown in the end of this section. As proposed in [11], coding across multicast sessions can increase transmission robustness and reduce the latency. Different to one hop scenarios, random coding across multicast sessions can be accomplished easily due to the broadcasting characteristics of radio propagation. Furthermore, in the context of multicast communication scenarios, the use of network coding brings the advantage that each relay transmission can help all destinations simultaneously, whereas one relay transmission can only help one source-destination pair each time for non-cooperative strategies.

A. Protocol Description and Signal Model

The cooperative coding strategy for two-hop scenarios can be described as following. At the first time slot, all sources broadcast their head-of-line packets simultaneously. Hence at this time slot, each relay receives the superposition of the M messages

$$y_{R_n} = \sum_{m=1}^{M} h_{mR_n} s_m + n_{R_n}, \quad n \in \{1, \ldots, L\}. \tag{1}$$

where s_m is the message transmitted from the mth source, n_{R_n} is the additive Gaussian noise at the relay and h_{mR_n} is the coefficient for the channel between the mth source and the relay R_n. To simplify notation, consider that each source transmits a symbol each time slot rather than a packet. In this paper, all wireless channels are assumed to be independent identical Rayleigh fading. For Rician or other types of fading, the developed signal model is still valid and similar analytical results can be obtained by taking the characteristics of particular fading into account. After this first transmission, all relays received a mixture of the M transmitted messages with different combination coefficients. For the next hop transmission, each relay will broadcast its received mixture to all destinations, which is analogous to the strategy of random coding across multicast sessions [11]. And thanks for the broadcasting nature of radio propagation, cooperation among multiple source-destination pairs does not consume extra bandwidth resource and the next hop transmission can benefit such random coding without any loss of bandwidth efficiency.

Due to the dynamic nature of radio propagation, the channel quality of one relay connecting to the destinations and sources varies, which is crucial to the system performance. Relays are scheduled to transmit in a way that a relay with better connection to the sources and destinations should be used earlier, where the details for relay selection will be discussed at the end of this section. The amplify-forward strategy is used here for relay transmission. During the next n time slots, the scheduled relays

will take their turns to forward the mixture to the destination, and the m-th destination receives

$$y_{D_{l,m}} = h_{R_l D_m} \hat{y_{R_l}} + n_{l+1}, \quad l = 1, \ldots, n, \tag{2}$$

where $\hat{y_{R_l}} = y_{R_l}/\beta$, $\beta_l = \sqrt{M+1/\rho}$ and ρ is denoted as signal-to-noise ratio (SNR) and $h_{R_l D_m}$ is channel coefficient from l-th relay to m-th destination. Note that the averaged transmission power constraint is applied to each relay. So after n time slots, the observations at the destination can be expressed as

$$\begin{bmatrix} y_{D_{1,m}} \\ \vdots \\ y_{D_{n,m}} \end{bmatrix} = \begin{bmatrix} \frac{h_{R_1 D_m}}{\sqrt{M}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{h_{R_n D_m}}{\sqrt{M}} \end{bmatrix} \begin{bmatrix} h_{1R_1} & \cdots & h_{MR_1} \\ \vdots & \ddots & \vdots \\ h_{1R_n} & \cdots & h_{MR_n} \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_M \end{bmatrix} + \begin{bmatrix} n_{1,m} + \frac{h_{R_1 D} n_{R_1}}{\sqrt{M}} \\ \vdots \\ n_{n,m} + \frac{h_{R_n D} n_{R_n}}{\sqrt{M}} \end{bmatrix}, \tag{3}$$

which can be denoted as

$$\mathbf{y}_{m,n} = \mathbf{D}_{m,n} \mathbf{H}_n \mathbf{s} + \mathbf{n}_{m,n} \tag{4}$$

The relays will keep forwarding their received mixture until all destinations have correctly received their corresponding source messages, where the criterion for successful transmission will be discussed later. It is possible that all relays have been scheduled to transmit, but at least one destination can not decode its source messages, e.g., $n \geq L$. In such a case, the relays will be reused, which means that there are some repeated rows in the matrices, $\mathbf{D}_{m,n}$ and $\mathbf{H}_n$.

B. Detection at Destinations

There are many choices of the criteria to determinate whether one destination can receive the source message correctly. One option is to ask each destination to decode all source messages and we can apply the capacity region of multiple access channel (MAC) since the signal model in (3) is exactly a MAC model. As shown in [14], [15], the error event based on the capacity region of MAC is the union of the events

$$E_{\mathcal{A}} \triangleq \left\{ \mathcal{I}(\mathbf{s}_{\mathcal{A}}; \mathbf{y} | \mathbf{s}_{\mathcal{A}^c}, \mathbf{H}_n = H, \mathbf{D}_{m,n} = D) \leq \sum_{i \in \mathcal{A}} R_i \right\}. \tag{5}$$

which denotes the error event that the information of the users in the subset $\mathcal{A}$ can not be decoded correctly. While the error probability based on MAC can tell us the optimal performance from the information theoretic aspect, complicated successive decoding is required at the receivers and hence it is difficult to realize it in practice.

Alternatively, we will use the principle of zero forcing [16], [17], a linear receiver which is not only easy to implement, but also helpful to simplify the development of analytical results. Applying the principle of zero forcing detection, a simplified signal model can be obtained as

$$\begin{aligned} (\mathbf{H}_n^H \mathbf{H}_n)^{-1} \mathbf{H}_n^H \mathbf{D}_{m,n}^{-1} \mathbf{y}_{m,n} &= \mathbf{s} + (\mathbf{H}_n^H \mathbf{H}_n)^{-1} \mathbf{H}_n^H \mathbf{D}_{m,n}^{-1} \mathbf{n}_{m,n} \tag{6} \\ &= \mathbf{s} + \tilde{\mathbf{n}}_{m,n} \end{aligned}$$

The signal-to-noise ratio (SNR) for the i-th source message at the m-th destination after n time slots can be expressed as

$$\rho_{i,m}^n = \frac{P}{E\{\tilde{n}_{i,m,n}\}},$$

where $\tilde{n}_{i,m,n}$ is the i-th element of the vector $\tilde{\mathbf{n}}_{m,n}$. The noise power $E\{\tilde{n}_{i,m,n}\}$ can be obtained from the noise covariance matrix as

$$\begin{aligned}
\mathbf{C}_{m,n} &= E\{\tilde{\mathbf{n}}_{m,n}\tilde{\mathbf{n}}_{m,n}^H\} \\
&= (\mathbf{H}_n^H\mathbf{H}_n)^{-1}\mathbf{H}_n^H\tilde{\mathbf{C}}_{m,n}\mathbf{H}_n(\mathbf{H}_n^H\mathbf{H}_n)^{-1}
\end{aligned} \tag{7}$$

where $\tilde{\mathbf{C}}_{m,n} = \mathbf{D}_{m,n}^{-1}\bar{\mathbf{C}}_{m,n}(\mathbf{D}_{m,n}^{-1})^H$ and $\bar{\mathbf{C}}_{m,n} = \text{diag}\{1 + |h_{R_1 D_m}|^2/M, \cdots, 1 + |h_{R_n D_m}|^2/M\}$. So we can have $\tilde{\mathbf{C}}_{m,n} = \text{diag}\{1 + M/|h_{R_1 D_m}|^2, \cdots, 1 + M/|h_{R_n D_m}|^2\}$. Based on these SNRs, $\rho_{i,m}^n$, the performance of the addressed protocol, in terms of delay and stability, can be obtained. Note that successful detection requires the knowledge of the channel coefficients at the destinations, which can be obtained through a dedicated control channel.

In summary, the proposed transmission protocol requires the assumptions as following. It is assumed that there is no direct link between the sources and destinations. Time division duplexing has been adopted, so the channels between the transmitter and receiver are reciprocal. Furthermore, we assume that each relay has the access to its local channel information, and each destination has the access to the channel coefficients to accomplish the zero-forcing detection as shown in (6). So compared to the scheme proposed in [11], the proposed transmission protocol requires extra system overhead to ensure the nodes to have access to the necessary channel information. However, the proposed scheme does not require the strong assumption that each source needs to know what is transmitted by the other sources, which can reduce the system overhead significantly compared to the scheme in [11]. For the scenarios where channels are changing rapidly, the proposed scheme may require more system overhead than the scheme in [11], but for the scenarios with quasi-static fading channels, the proposed scheme is more spectrally efficient.

C. Distributed scheduling strategy of relay forwarding

As discussed in [18], relay selection can be accomplished in a distributed and effective way. In particular, each relay calculates its backoff time inversely proportional to its channel quality, which means the relay with the best channel quality will seize the control of the channel. An important question is what is the desirable criterion of the channel quality, which is focused in the following.

Different to those scenarios with single S-D pair, the design of relay scheduling/selection for the addressed multicast sessions is more complicated. Due to the use of the random coding across the sessions, one relay has to serve more than one S-D. Assume that each relay has the access to its local channel

state information (CSI), $[h_{1R_m}, \cdots, h_{MR_m}, h_{R_mD_1}, \cdots, h_{R_mD_M}]$. Among possible choices of the criteria for relay quality, the one used in this paper is to order the relays in a descending order according to its worst link, e.g.,

$$|h_{R_1,min}|^2 \geq \cdots \geq |h_{R_L,min}|^2,$$

where $|h_{R_m,min}|^2 = \min\left(|h_{1R_m}|^2, \cdots, |h_{MR_m}|^2, |h_{R_mD_1}|^2, \cdots, |h_{R_mD_M}|^2\right)$. Consider that there is only one S-D pair, recall that a well known relay selection criterion for amplify forward protocols is the harmonic mean of the incoming and outgoing channels of each relay,

$$\frac{|h_{1R_m}|^2|h_{R_mD_1}|^2}{|h_{1R_m}|^2 + |h_{R_mD_1}|^2}.$$

For such a special case, the proposed criterion is to select relays according to the minimum of their incoming and outgoing channels, $\min\left(|h_{1R_m}|^2, |h_{R_mD_1}|^2\right)$, which is the same as the harmonic mean. Note that the source-relay channels have been also included into the criterion for relay selection, which is to avoid the use of a relay which has a poor connection to the source nodes. An extreme example is that there is a relay which has no connection to all of the source nodes. Even if this relay has strong connections to the destination nodes, the relay cannot help the source nodes and hence we cannot use this relay.

IV. PERFORMANCE ANALYSIS FOR THE TWO-HOP TRANSMISSIONS

In this paper, we are interested in the impact of the proposed transmission protocol on the system delay and stability. Define $\bar{N}$ as the averaged number of retransmissions required to deliver the M head-of-line packets to their associated destinations. So the averaged delay and the maximum stable arrival rate for the addressed two-hop scenario can be obtained as

$$\bar{T} = \bar{N}T_p, \quad \lambda = \frac{M}{\bar{N}},$$

where T_p is the time duration of one time slot. Obviously the key step to study the delay and stability is to find the averaged number of required time slots. Define $P(N = n)$ as the probability for the event that n retransmissions ensure all messages are corrected decoded by all destinations, but $n - 1$ transmissions can not. Then the averaged number of the minimum retransmissions can be obtained as

$$\bar{N} = \sum_{n=1}^{\infty} nP(N = n). \tag{8}$$

The probability, $P(N = n)$, can be obtained in different ways dependent on the definition of the error probability. In this paper, we will use two types of error probability, outage probability and maximum-likelihood (ML) decoding error for convolutional codes. Conditioned on high SNR and infinity length of coding, the outage probability can be tightly bounded the ML probability of detection error. So in the following, the outage probability is first to use in order to get some closed-form expression of $P(N = n)$. Later the error probability based on practical length of coding will be studied.

A. Analytical results based on outage probability

An outage event occurs when the mutual information supported by the instantaneous receive SNR is less than the target data rate. Hence based on the simplified model in (6), the probability $P(N = n)$ can be defined as[1]

$$P(N = n + 1) = P(\rho_{min}^{n-1} \leq \phi, \rho_{min}^{n} \geq \phi) \tag{9}$$

where $\phi = 2^R - 1$, R is the target data rate, and ρ_{min}^{n} is the minimum SNR among the M subchannels, e.g., $\rho_{min}^{n} = \min(\rho_{m,m}^{n})$, $\forall m \in \{1, \ldots, M\}$. Note that a symmetric network is considered here, where the target rates for all sessions are the same. The use of relay scheduling means that the elements in $\mathbf{H}_n$ and $\mathbf{D}_{m,n}$ are no longer complex Gaussian distributed. To obtain tractable analytical expressions, we first construct an auxiliary signal model as following

$$\bar{\mathbf{y}}_n = \mathbf{s} + \bar{\mathbf{n}}_n, \tag{10}$$

which has the new noise covariance matrix as

$$\begin{aligned}
\bar{\mathbf{C}}_n &= (\mathbf{H}_n^H \mathbf{H}_n)^{-1} \mathbf{H}_n^H (1 + M) \mathbf{I}_M \mathbf{H}_n (\mathbf{H}_n^H \mathbf{H}_n)^{-1} \\
&= (1 + M)(\mathbf{H}_n^H \mathbf{H}_n)^{-1}.
\end{aligned} \tag{11}$$

Conditioned on the assumption that there are an infinity number of relays, the use of the strategy of relay scheduling can ensure that no relay is scheduled twice and those used relays can have good enough outgoing channels, $\frac{1}{|h_{R_m D_m}|^2} \leq 1$. So with infinity relays, we can have the following inequality

$$\bar{\mathbf{C}}_n(i, i) \geq \mathbf{C}_{m,n}(i, i), \quad \forall i, m \in \{1, \ldots, M\}, \tag{12}$$

which means that the use of such a simplified signal model could result in more transmissions than the original model. Hence in the following, we will focus on this simplified signal model to develop tractable analytical results. Define $\tilde{N}$ as the total transmission number required to deliver source messages to all destinations by using the simplified model in (10), and we can have the following inequality

$$\bar{N} = \sum_{n=1}^{\infty} nP(N = n) \leq \sum_{n=1}^{\infty} nP(\tilde{N} = n). \tag{13}$$

To obtain tractable expressions, the probability $P(\tilde{N} = n + 1)$ is first expressed as

$$P(\tilde{N} = n + 1) = P(\max\{N_1, \cdots, N_M\} = n + 1) \tag{14}$$

[1]For the proposed transmission protocol, all M sources transmit simultaneously during the first time slot, and then the selected relays will forward their mixtures to the destinations. So the transmission delay consists of two parts, one due to the source transmission and the other due to relay transmissions. The former only consumes one time slot. So for notation simplicity, the expression $P(N = n + 1)$ will be used in the following, where n denotes the number of the relay retransmission.

where N_m denotes the number of transmissions required for reliable communication between the mth source-destination pair by using the simplified model in (10). Define $\gamma_{m,n}$ as the effective channel gain at the m-th subchannel of the model in (10) after n time slots[2], and hence the corresponding SNR can be written as $\frac{\gamma_{m,n}\rho}{M+1}$. To further simplify the development, it is assumed that the elements in the matrix $\mathbf{H}_n$ are i.i.d complex Gaussian distributed, which has no effect to the inequality in (12). Provided that $\mathbf{H}_n$ is a Gaussian random matrix, the density function of the effective channel gain at m-th subchannel after n transmissions $\gamma_{m,n}$ can be expressed as [16], [17]

$$f_{\gamma_{m,n}}(\gamma) = \frac{e^{-\gamma}}{(n-M)!}(\gamma)^{n-M}$$

which is Chi-square distributed with $2(n-M+1)$ degree of freedom. And the probability $P(N_m = n)$ can be expressed as

$$P(N_m = n+1) = P(\gamma_{m,n-1} \leq \epsilon, \gamma_{m,n} \geq \epsilon), \tag{15}$$

where $\epsilon = \frac{(2^R-1)(M+1)}{\rho}$. While the density of $\gamma_{m,n-1}$ and $\gamma_{m,n}$ can be easily found, obviously the two variables are not independent to each other and it is not clear how to get their joint density function which is needed to obtain the probability $P(N_m = n)$. Define $\Delta_{m,n} = \gamma_{m,n} - \gamma_{m,n-1}$. An intuition is that $\Delta_{m,n}$ should be exponentially distributed and independent to $\gamma_{m,n-1}$ **if** $\gamma_{m,n-1}$ is the sum of $(n-M)$ i.i.d complex Gaussian variables and $\gamma_{m,n}$ is the addition of $\gamma_{m,n-1}$ with another i.i.d Gaussian variable. Although the relationship between the two variable is not as explicit as expected, the intuition for the density of $\Delta_{m,n}$ is still valid, as shown in the following lemma.

Lemma *1:* Consider a $N \times M$ complex Gaussian matrix $\mathbf{H}_N$. Define

$$\gamma_{m,N} = \frac{1}{[(\mathbf{H}_N^H \mathbf{H}_N)^{-1}]_{m,m}},$$

where $[\mathbf{A}]_{m,m}$ denotes the ith element on the diagonal of $\mathbf{A}$. It can be proved that the difference between $\gamma_{m,N}$ and $\gamma_{m,N+1}$, denoted as $\Delta_{m,N}$, is independent to $\gamma_{m,N}$, and its cumulative density function is

$$f_{\Delta_{m,N}}(x) = 1 - e^{-x}.$$

Proof: See Appendix. ∎

By using Lemma 1, the probability $P(N_m = n+1)$ can be found as

$$
\begin{aligned}
P(N_m = n+1) &= P(\gamma_{m,n-1} \leq \epsilon, \gamma_{m,n} \geq \epsilon) \\
&= \int_0^\epsilon e^{-(\epsilon-z)} \frac{z^{n-M-1}}{(n-M-1)!} e^{-z} dz = \frac{e^{-\epsilon}}{(n-M)!} \epsilon^{n-M}.
\end{aligned}
\tag{16}
$$

[2]To distinguish the original and simplified models, γ is used to denote the SNR based on the simplified model and ρ is used to denote the SNR based on the original model.
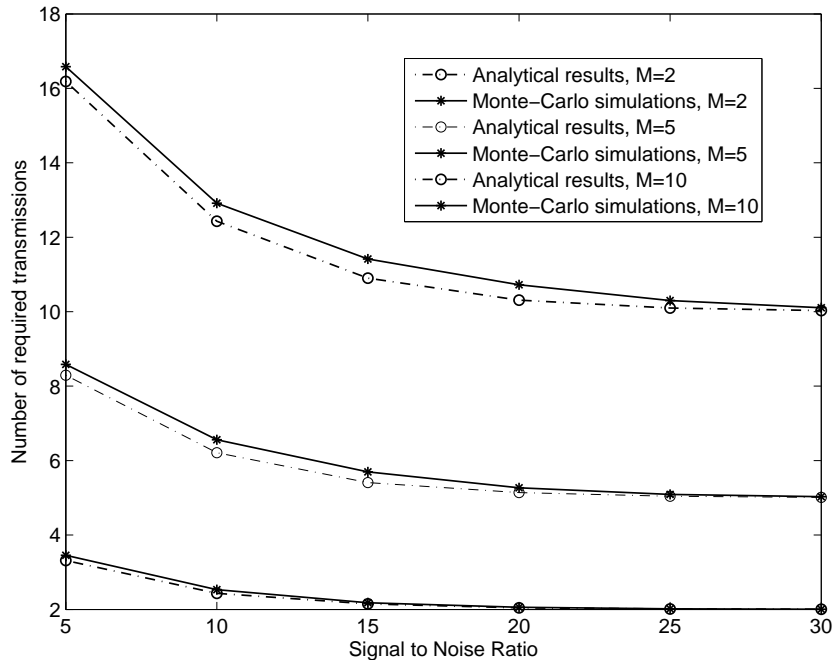
Fig. 2. Value of $\sum_{n=1}^{\infty} nP(\tilde{N} = n)$ obtained by the Monte-Carlo simulations and analytical expressions vs SNR. The targeted data rate is $R = 1$ bit/s/Hz.

Consider that the numbers of required time slots for M source-destination pairs, $N_1, \cdots, N_M$, can be sorted in descending order as

$$N_{(1)} \leq N_{(2)} \cdots \leq N_{(M)}$$

By using order statistics [19], we can find the probability of $N_{(M)}$, the largest number of transmissions

$$P(\tilde{N} = n + 1) = P(N_{(M)} = n + 1) = \left(\sum_{k=M}^{n} P(N_i = k) \right)^M - \left(\sum_{k=M}^{n-1} P(N_i = k) \right)^M. \qquad (17)$$

And the expectation of the largest number of retransmissions can be upper bounded as

$$\bar{N}_{(M)} \leq \sum_{n=1}^{\infty} nP(\tilde{N} = n) = (M+1)\left(P(N_i = M)\right)^M \qquad (18)$$

$$+ \sum_{n=M+2}^{\infty} n \left(\left(\sum_{k=M}^{n-1} P(N_i = k) \right)^M - \left(\sum_{k=M}^{n-2} P(N_i = k) \right)^M \right)$$

Recall that the use of order statistics in (17) requires that all variables $N_1, \cdots, N_M$ are independent. It is important to note that the multiple sub-channels for linear receivers are not strictly independent, however, the approximation of independence is used here to make the analytical results tractable as in [20]. In Fig. 2, the analytical results with the independence approximation and the results obtained by Monte-Carlo simulations are shown to be close to each other.

To obtain insights of the performance achieved by the proposed protocol, we use the exponential expansion and the assumption of medium SNR to obtain some approximations. The probability $P(N_i =$

$n + 1$) can be simplified as

$$P(N_i = n + 1) \approx \frac{1 - \epsilon}{(n - M)!} \epsilon^{n-M}.$$

And based on this approximation, we can obtain

$$(P(N_i = M))^M \approx (1 - \epsilon)^M \approx 1 - M\epsilon \tag{19}$$

Similarly we can have

$$\sum_{n=M+2}^{\infty} n \left(\left(\sum_{k=M}^{n-1} P(N_i = k) \right)^M - \left(\sum_{k=M}^{n-2} P(N_i = k) \right)^M \right) \tag{20}$$

$$\approx \sum_{n=M+2}^{\infty} n \left(\left(\sum_{k=M}^{n-1} \frac{(1 - \epsilon)\epsilon^{k-M}}{(k - M)!} \right)^M - \left(\sum_{k=M}^{n-2} \frac{(1 - \epsilon)\epsilon^{k-M}}{(k - M)!} \right)^M \right)$$

$$\approx (M + 2) \left(\left(1 - \epsilon^2\right)^M - (1 - \epsilon)^M \right) \approx (M + 2)M\epsilon \tag{21}$$

And finally for the expected number of required time slots can be upper bounded

$$\bar{N} \leq \bar{N}_{(M)} \approx (M + 1)(1 - M\epsilon) + (M + 2)M\epsilon = M(1 + \epsilon) + 1, \tag{22}$$

conditioned on the assumption that there is infinity number of relays. As can be observed from (22), the expected number of the required time slots scales linearly with the number of the packages, which is consistent to the results provided in [11]. However, the scheme proposed in [11] requires that each transmitter has the non-casual priori information of the messages sent by other transmitters.

B. Two comparable schemes

1) Non-cooperative direct transmission scheme: For non-cooperative direct transmission, as shown in [11], the averaged number of one-hop transmissions for the i-th time slot can be expressed as

$$P(N_i = n) = P\left(\rho(n - 1)|h_m|^2 \leq \phi \quad \& \quad \rho n|h_m|^2 \geq \phi\right) \tag{23}$$

$$= e^{-\frac{2^R - 1}{n\rho}} - e^{-\frac{2^R - 1}{(n-1)\rho}}$$

In total, the expected number of required time slots for the two-hop multisessions can be obtained as following

$$\bar{N}_D = 2M\bar{N}_i = 2M \sum_{n=1}^{\infty} n \left(e^{-\frac{2^R - 1}{n\rho}} - e^{-\frac{2^R - 1}{(n-1)\rho}} \right) \tag{24}$$

Conditioned on $\epsilon = \frac{2^R - 1}{\rho} \leq 1$, we can have the following approximation

$$\bar{N}_D \approx 2M(1 - \epsilon) + 2M \sum_{n=2}^{\infty} n \left(\frac{\epsilon}{n - 1} - \frac{\epsilon}{n} \right) \tag{25}$$

$$\approx 2M(1 - \epsilon) + 2\epsilon M \sum_{k=1}^{\infty} \frac{1}{k}$$

There is no closed-form expression of the sum $\sum_{k=1}^{\infty} \frac{1}{k}$ which can be expressed as

$$\sum_{k=1}^{\infty} \frac{1}{k} = \zeta(1)$$

where $\zeta(\cdot)$ is [21] (Eq(0.233.1)). As discussed in [21], $z = 1$ is the only singular point of the Zeta function, $\zeta(1) \to \infty$. As a result, the expected number of retransmissions for non cooperative transmission becomes also infinity.

$$\bar{N}_D \quad \approx \quad 2M(1 - \epsilon) + \epsilon M \zeta(1) \to \infty. \tag{26}$$

Comparing (22) with (26), we can observe that the introduce of coding cross session has improved the robustness of transmissions. Interestingly (26) contradicts with the intuition that the number of required retransmissions for direct transmission should scale with the number of packages. This is due to the fact that channel coefficients have been assumed to be constant. So with non-zero probability, it can be expected that one channel goes through deep fading, $|h_m| \to 0$, which could cause an extreme large number of retransmissions.

2) Best-relay transmission scheme: For the addressed M sessions, the best-relay transmission scheme consists of M stages. At each stage, only one pair of source and destination are involved. In specific, at the first time slot for each stage, the source broadcasts its message to all relays, and during the following time slots, only one relay with the best channel condition keeps transmitting until the message is received by the corresponding destination. The criterion of relay selection is based on the harmonic mean of the incoming and outgoing channels of each relay. And the rest S-D pairs will take their turn to transmit. For the m-th pair, the probability of the event that $n + 1$ time slots can ensure correct reception is

$$P(N_{m,BR} = n + 1) = P(\rho\mu_{m,n-1} \leq \phi, \rho\mu_{m,n} \geq \phi),$$

where $\rho\mu_{m,n}$ is the received SNR after $(n + 1)$ transmissions, $\mu_{m,n} = n\alpha$ and α is the harmonic mean of the incoming and outgoing channels of the chosen relay. And the expected total number of required time slots for multisessions can be written as

$$\bar{N}_{BR} = M \sum_{n=1}^{\infty} (n + 1)P(N_{m,BR} = n + 1).$$

Due to the difficult to obtain the probability $P(N_{m,BR} = n + 1)$, the performance of the best-relay transmission scheme will be only evaluated by using simulation results.

C. Performance analysis based on convolutional coding

In this subsection, we will attach convolutional codes to the proposed scheme and derive its code bound. Assuming that data sequence from each source has been convolutional encoded by a convolutional

code. Still the relays forward the messages to destinations by employing amplify-forward strategy until all destinations have decoded correctly. Let $\mathcal{O}_0^n$, $\mathcal{O}_u^n$ and $\mathcal{O}_d^n$ denote, respectively, the events "decoded sequence contains no errors", "decoded sequence contains undetected errors" and "decoded sequence contains detected errors" when there are n transmissions. Clearly, $P(\mathcal{O}_0^n) + P(\mathcal{O}_u^n) + P(\mathcal{O}_d^n) = 1$. Further assuming $\mathcal{O}_u^n$ is actually negligible, hence $\mathcal{O}_d^n$ can be approximated by

$$P(\mathcal{O}_d^n) = 1 - P(\mathcal{O}_0^n). \tag{27}$$

Notice that the joint probability [22] $P(\mathcal{O}_d^1, \mathcal{O}_d^2, \cdots, \mathcal{O}_d^n)$ can be upper bounded as

$$P(\mathcal{O}_d^1, \mathcal{O}_d^2, \cdots, \mathcal{O}_d^n) \leq P(\mathcal{O}_d^n). \tag{28}$$

Now the error probability that the m-th user's message can be decoded in N transmissions can be expressed as

$$P(N = n + 1) = P(\mathcal{O}_d^1, \cdots, \mathcal{O}_d^{n-1}, \bar{\mathcal{O}}_d^n) \tag{29}$$

Where $\bar{\mathcal{O}}_d^n$ is complementary event of $\mathcal{O}_d^n$.

By considering (27) and (28), we have $P(N = n + 1)$ bounded by

$$P(N = n + 1) = P(\mathcal{O}_d^{n-1}) - P(\mathcal{O}_d^n) \tag{30}$$

For a frame length K convolutional coded data, $P(\mathcal{O}_d^n)$ can be bounded by

$$P(\mathcal{O}_d^n) \geq 1 - (1 - P(E^n))^K \tag{31}$$

where $P(E^n)$ is the probability of a decoding error event of Viterbi decoding after the nth transmission. According to the Viterbi decoding convolutional codes bounds, the mth sub-sessions error probability $P(E^n)$ can be upper bounded as

$$P(E^n) < \sum_{d=d_{free}}^{\infty} \beta_d Q(\sqrt{2\rho_{m,n} R_c d}) \tag{32}$$

where β_d, d_{free} and R_c denote, respectively, distance spectra, free distance and the rate of employed convolutional code. the Q function is defined as $Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\frac{x^2}{2}} dx$. Substituting (31) into (30) and using approximation $(1 - x)^n \approx 1 - nx$ when x is sufficient small, we may rewrite $P(N = n + 1)$ approximately as

$$P(N = n + 1) = (1 - P(E^n))^K \approx K \left(P(E^{n-1}) - P(E^n) \right) \tag{33}$$

Now the error probability can be expressed as

$$\begin{aligned} P(N = n + 1) = \ & K \sum_{d=d_{free}}^{\infty} \beta_d Q \left(\sqrt{2\rho_{m,n-1} R_c d} \right) \\ & - K \sum_{d=d_{free}}^{\infty} \beta_d Q \left(\sqrt{2\rho_{m,n} R_c d} \right) \end{aligned} \tag{34}$$

Note that the Q function can be upper bounded as

$$Q(\alpha) \leq \frac{1}{2} e^{-\frac{x^2}{2}} \tag{35}$$

The (34) is bounded by

$$P(N = n + 1) \leq \sum_{d=d_{free}}^{\infty} \frac{K\beta_d e^{-\rho_{m,n-1} R_c d}}{2} - \sum_{d=d_{free}}^{\infty} \frac{K\beta_d e^{-\rho_{m,n} R_c d}}{2} \tag{36}$$

Since $\rho_{m,n}$ are Chi-square distributed with $2(n - M + 1)$ degree of freedom, $P(N = n + 1)$ can be formulated by its expectation

$$P(N = n+1) \leq \int_0^{\infty} \left(\frac{\gamma^{(n-M-1)} e^{-\gamma}}{(n-M-1)!} \sum_{d=d_{free}}^{\infty} \frac{K\beta_d e^{-\rho\gamma R_c d}}{2} \right) \mathrm{d}\gamma - \int_0^{\infty} \left(\frac{\gamma^{(n-M)} e^{-\gamma}}{(n-M)!} \sum_{d=d_{free}}^{\infty} \frac{K\beta_d e^{-\rho\gamma R_c d}}{2} \right) \mathrm{d}\gamma \tag{37}$$

By interchanging the integrating and sum operation, we have

$$P(N = n + 1) \leq \sum_{d=d_{free}}^{\infty} \int_0^{\infty} \left(\frac{K\beta_d \gamma^{(n-M-1)} e^{-(\rho R_c d+1)\gamma}}{2(n-M-1)!} \right) \mathrm{d}\gamma - \sum_{d=d_{free}}^{\infty} \int_0^{\infty} \left(\frac{K\beta_d \gamma^{(n-M)} e^{-(\rho R_c d+1)\gamma}}{2(n-M)!} \right) \mathrm{d}\gamma \tag{38}$$

The (38) can be integrated and simplified as

$$P(N = n + 1) \leq \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{\rho R_c d}{(\rho R_c d+1)^{n-M+1}} \right) \approx \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{1}{(\rho R_c d+1)^{n-M}} \right) \tag{39}$$

By combining (18) and (39), the expectation of largest number of re-transmission under convolutional coded case can be written as

$$\bar{N}_{(M)} = (M + 1) \left(1 - \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{1}{\rho R_c d+1} \right) \right)^M \tag{40}$$

$$+ \sum_{n=M+2}^{\infty} n \left(\left(\sum_{k=M}^{n-1} \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{\rho R_c d}{(\rho R_c d+1)^{k-M+1}} \right) \right)^M - \left(\sum_{k=M}^{n-2} \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{\rho R_c d}{(\rho R_c d+1)^{k-M+1}} \right) \right)^M \right)$$

Let $\epsilon_{c0} = \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{1}{\rho R_c d+1} \right)$, $\epsilon_{c1} = \sum_{k=M}^{n-1} \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{\rho R_c d}{(\rho R_c d+1)^{k-M+1}} \right)$ and $\epsilon_{c2} = \sum_{d=d_{free}}^{\infty} \frac{K\beta_d}{2} \left(\frac{\rho R_c d}{(\rho R_c d+1)^{n-M}} \right)$, at the high SNR region, (40) can be approximated as

$$\bar{N}_{(M)} \approx (M + 1)(1 - M\epsilon_{c0}) + \sum_{n=M+2}^{\infty} nM\epsilon_{c1}^{M-1}\epsilon_{c2} \tag{41}$$

Notice when SNR goes to very large value, $\epsilon_{c0}, \epsilon_{c1}$ and ϵ_{c2} would be extremely small, which indicates the average number of transmissions of proposed protocol should converge to $M + 1$ eventually.

V. NUMERICAL RESULTS

In this section, the performance of the proposed two-hop transmission protocol is evaluated with the comparison to the two comparable schemes, the two-hop non cooperative direct transmission scheme and the two-hop best-relay scheme. The elements of the channel and noise matrices are zero-mean, circular complex Gaussian random variables, where the variances of the channel and noise are set according to the signal-to-noise ratio. A symmetric system is considered here where all pairs of sources and destinations have the same targeted data rate and use the same convolutional coding.

First the expected number of required transmissions based on the *outage probability* is studied for the three transmission schemes. The targeted data rate is set as $R = 1$bits/s/Hz. In Fig. 3, the performance of the proposed cross-coding scheme is compared with the direct transmission. The number of relays is not fixed, which is to show the impact of the relay number on the system performance. As discussed in the earlier section, the use of this non cooperative scheme could result in an infinite number of transmissions in case of deep fading. During the simulation, we provide a threshold for the number of retransmissions for the non cooperative scheme. If a destination can not receive its source message reliably after this threshold value, transmission will be stopped and the required transmission will be set as the threshold. During the simulation, such a threshold is set as 100. As can be seen from Fig. 3, the proposed scheme can achieve significantly performance gain over the non cooperative scheme particularly in low SNR range, where the second factor in (25) becomes dominant due to small value of ρ. Notice that the number of retransmissions for direct transmission seems converting to a fixed value, which is contradicted to the analytical results provided in Section IV.B.2. This is mainly due to the fact that the number of retransmissions is caped by the fixed-value threshold, and with an infinite number of simulation runs, it can be expected that the number of retransmission for the direct transmission approaches infinity.

In Fig. 4, the best relay scheme is shown as the comparable scheme. At low SNR, the proposed transmission scheme suffers some performance loss, which is due to the fact that multi-user diversity can not be exploited as effectively as the best relay scheme. For the proposed scheme, a used relay has to serve more than one pairs of sources and destinations at the same time, which imposes more demanding requirement for the quality of relay. This phenomenon is analogs to the so-called channel hardening effect that the more antennas are equipped by terminals, it is more difficult to exploit multi user diversity. For the best relay scheme, relay selection is based on a single S-D pair, which means that the best relay scheme only needs a small number of relays to achieve acceptable performance. However, at high SNR, the proposed scheme can always achieve better performance than the best-relaying scheme which is due to the fact that less bandwidth resource has been consumed during the second hop transmission for the proposed scheme. In particular, for the proposed scheme, one single relay transmission can help
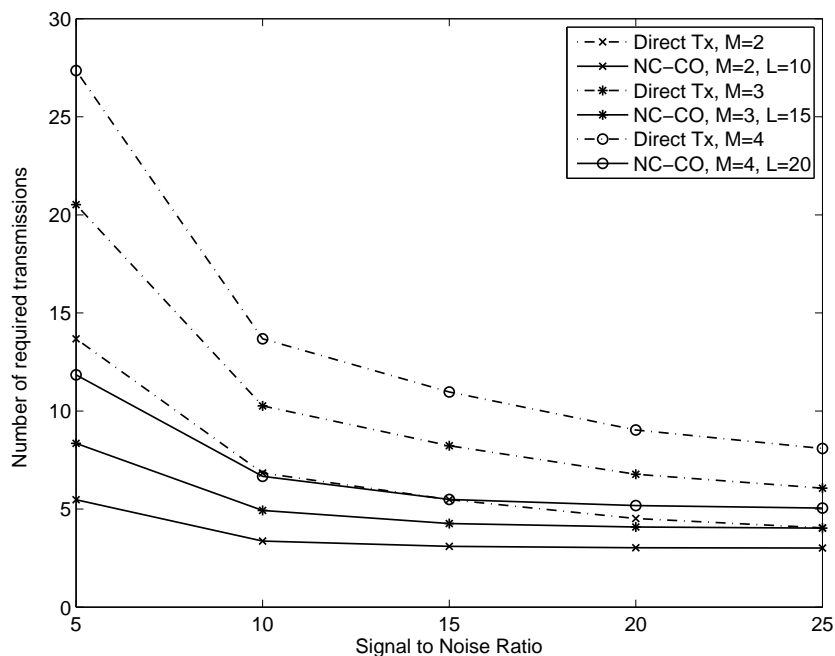
Fig. 3. Averaged number of required retransmissions vs SNR. The targeted data rate is $R = 1$ bit/s/Hz.

all destination nodes simultaneously, whereas only one source-destination pair can be benefited by single relay transmission.

In the *convolutional coded* case, a $(5, 7)_{oct}$ systematic recursive convolutional code is considered with 100 information bits in each coded frame. Fig. 5 shows the performance comparison between convolutional coded cooperative transmission and direct transmission. Slightly different from the ideal random codes case, the practical coded transmission system may suffer the more transmissions at low SNR (e.g. less than 10dB) when M goes to large value due to the poor cooperative gain in low SNR region. But at medium to high SNR region, our protocol outperforms the direct transmission case significantly in terms of the number of transmission. Especially, the number of transmission would converge to $M + 1$ while the increasing of SNR. Comparing Fig. 5 with the previous two figures, we can observe that the number of transmissions based on convolutional coding is much larger than that based on outage probability, particularly at low SNR. One reason is that the outage probability can only closely bound the error probability of maximum likelihood if SNR is large enough, as discussed in [14], [23]. Furthermore, it is well known that simple convolutional coding can not realize the Shannon capacity, particularly at low SNR. Provided the use of more sophisticated error control codes, such as turbo or LDPC codes, the performance gap between the cooperative and non cooperative schemes can be reduced.
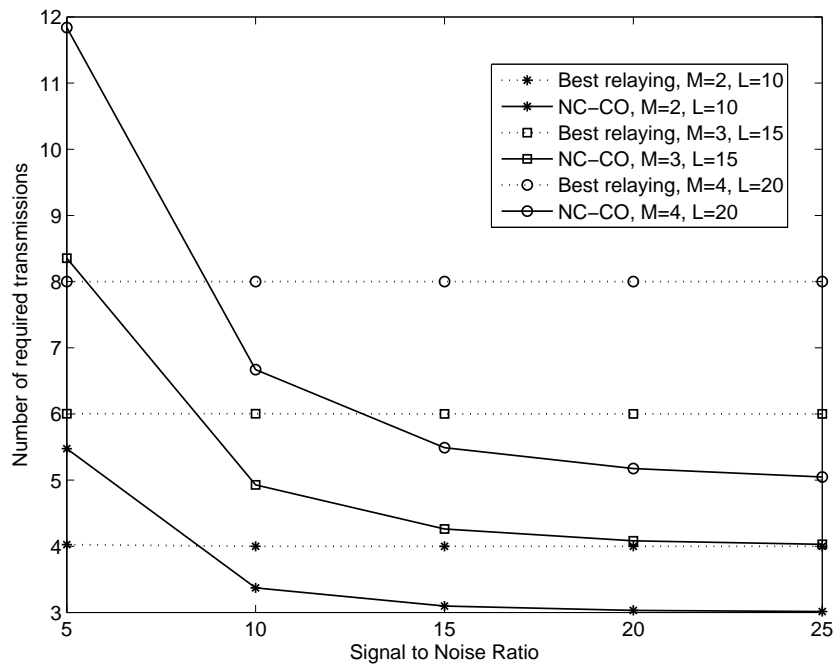
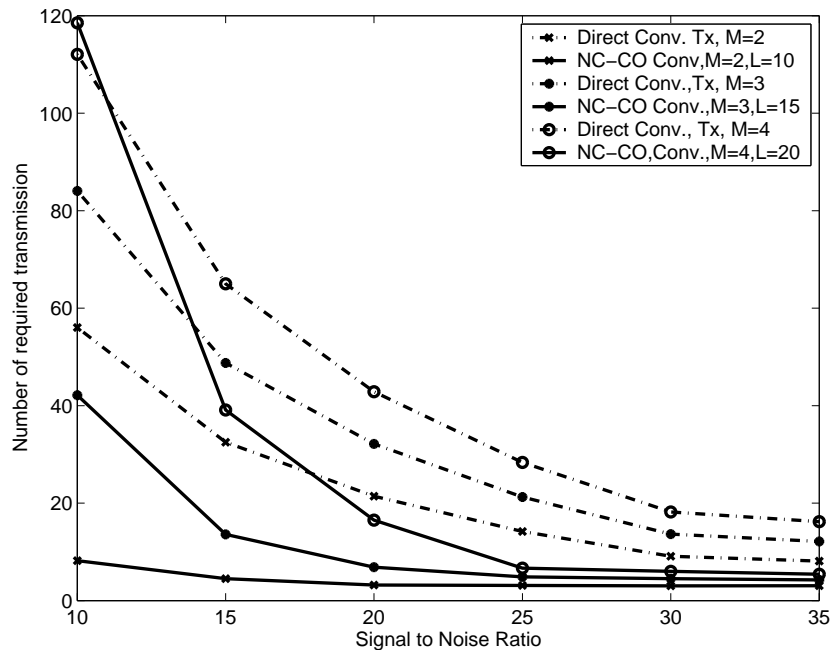Fig. 4.   Averaged number of required retransmissions vs SNR. The targeted data rate is $R = 1$ bit/s/Hz.

Fig. 5.   Averaged number of required retransmissions vs SNR under $(5, 7)_{oct}$ systematic convolutional coed case. The targeted data rate is $R = 1$ bit/s/Hz.

VI. CONCLUSION

The work in [11] has shown that random coding across multi-cast sessions can reduce the system delay significantly, however, such a scheme requires the strong assumption that each source has the priori information of other sources' messages. Actually the broadcasting nature of radio propagation can

provide an opportunity to realize collaboration across sessions without causing much system overhead. In this paper, we have proposed the application of network coding to multi-source multi-destination (MSMD) scenarios and provided formal analysis for the improvement of system delay. In particular, two types of analytical results have been developed, one based on the outage probability and the other based on the use of practical convolutional codes. Monte-Carlo simulation results have also been provided to demonstrate the delay performance of the proposed network coded protocol.

APPENDIX

Proof: Without loss any generality, the SNR for the first stream will be focused, $\gamma_{1,N}$, which can be recalled as

$$\gamma_{1,N} = \frac{1}{[(\mathbf{H}_N^H \mathbf{H}_N)^{-1}]_{1,1}}.$$

In the following, the index of the stream, 1, will be omitted for simplicity. By asking one more relay to forward a new mixture, the channel matrix after $N+1$ transmissions is $\mathbf{H}_{N+1} = \begin{bmatrix} \mathbf{H}_N^H & \mathbf{h}_{N+1} \end{bmatrix}^H$, where $\mathbf{h}$ is the $M \times 1$ vector containing the channels between the new used relay and M sources. Alternatively, the channel matrix can be written as $\mathbf{H}_N = \begin{bmatrix} \mathbf{g}_N & \tilde{\mathbf{H}}_N \end{bmatrix}$ where $\mathbf{g}_N$ denotes the first column of $\mathbf{H}_N$. Hence the SNR after N transmissions can be expressed as

$$\begin{aligned}
\gamma_N &= \frac{1}{[(\mathbf{H}_N^H \mathbf{H}_N)^{-1}]_{1,1}} = \frac{\det(\mathbf{H}_N^H \mathbf{H}_N)}{\det(\tilde{\mathbf{H}}_N^H \tilde{\mathbf{H}}_N)} \\
&= \mathbf{g}_N^H (\mathbf{I}_N - \mathbf{P}_N) \mathbf{g}_N,
\end{aligned} \tag{42}$$

where $\mathbf{P} = \tilde{\mathbf{H}}_N (\tilde{\mathbf{H}}_N^H \tilde{\mathbf{H}}_N)^{-1} \tilde{\mathbf{H}}_N^H$ and the last equality follows the property of the determinant of block matrices. The difference between the two SNRs, γ_{N+1} and γ_N, can be expressed as

$$\gamma_{N+1} - \gamma_N = \mathbf{g}_{N+1}^H [\mathbf{I}_{N+1} - \mathbf{P}_{N+1}] \mathbf{g}_{N+1} - \mathbf{g}_N^H [\mathbf{I}_N - \mathbf{P}_N] \mathbf{g}_N \tag{43}$$

Note that the relationship between the two channel vectors is $\mathbf{g}_{N+1} = \begin{bmatrix} \mathbf{g}_N^T & g(N+1) \end{bmatrix}^T$, where $g(N+1)$ denotes the $(N+1)$-th element of the vector $\mathbf{g}_{N+1}$. Hence the difference between the two SNR can be expressed as

$$\begin{aligned}
\gamma_{N+1} - \gamma_N &= \mathbf{g}_{N+1}^H [\mathbf{I}_{N+1} - \mathbf{P}_{N+1}] \mathbf{g}_{N+1} - \mathbf{g}_{N+1}^H [\mathbf{I}_{N+1} - \tilde{\mathbf{P}}_N] \mathbf{g}_{N+1} \\
&= \mathbf{g}_{N+1}^H [\tilde{\mathbf{P}}_N - \mathbf{P}_{N+1}] \mathbf{g}_{N+1}
\end{aligned} \tag{44}$$

where $\tilde{\mathbf{P}}_N = \begin{bmatrix} \mathbf{P}_N & \mathbf{0}_{N,1} \\ \mathbf{0}_{1,N} & 1 \end{bmatrix}$. It is interesting to observe that both the two matrices, $\mathbf{P}_{N+1}$ and $\tilde{\mathbf{P}}_N$, are idempotent since $\mathbf{P}_{N+1}^2 = \mathbf{P}_{N+1}$ and $\tilde{\mathbf{P}}_N^2 = \tilde{\mathbf{P}}_N$. Recall that an idempotent matrix $\mathbf{A}$ has an important property,

$$\mathbf{A}\mathbf{v} = \begin{cases} \mathbf{v}, & \text{if} \quad \mathbf{v} \quad \text{in the range of} \quad \mathbf{A} \\ 0, & \text{if} \quad \mathbf{v} \quad \text{in the null space of} \quad \mathbf{A} \end{cases}, \tag{45}$$

where $\mathbf{v}$ is a column vector. Alternatively, it can be shown that an idempotent matrix has the eigenvalues which is either one or zero.

The rank of $\tilde{\mathbf{P}}_N$ is $M+1$, and the rank of $\mathbf{P}_{N+1}$ is M since the channel matrices are assumed to be full column rank. Denote $\boldsymbol{\Psi}_N$ as the null space of the matrix $\tilde{\mathbf{P}}_N$ and $\boldsymbol{\Phi}_N$ as its range. The two spaces, $\boldsymbol{\Psi}_{N+1}$ and $\boldsymbol{\Phi}_{N+1}$, are defined in a similar way. Consider a vector $\mathbf{v}$ from the null space of $\mathbf{P}_N$, i.e. $\mathbf{P}_N\mathbf{v} = 0$. Obviously the vector $\tilde{\mathbf{v}} = \begin{bmatrix} \mathbf{v}^T & 0 \end{bmatrix}^T$ is in the null space of $\tilde{\mathbf{P}}_N$ and $\mathbf{P}_{N+1}$, $\mathbf{v} \in \boldsymbol{\Psi}_N$ and $\mathbf{v} \in \boldsymbol{\Psi}_{N+1}$. The rank of $\mathbf{P}_N$ is M and hence the dimension of its null space is $N - M$. Hence we can find $N - M$ common vectors shared by the two null spaces $\boldsymbol{\Psi}_N$ and $\boldsymbol{\Psi}_{N+1}$. As a result, the range of the two matrices share M vectors, denoted as $\mathbf{u}_i$ for $1 \leq i \leq M$. According to (45), we can have

$$\mathbf{P}_{N+1}\mathbf{u}_i = \tilde{\mathbf{P}}_N\mathbf{u}_i = \mathbf{u}_i$$

which implies that the M vectors in the shared range of the two matrices are in the null space of the $\tilde{\mathbf{P}}_N - \mathbf{P}_{N+1}$. Combined with the $N - M$ vectors shared by the null space of the two matrices, we can find N independent eigenvectors which are correspondent to the eigenvalue 0. The only left vector, denoted as $\mathbf{w}$, which is shared by the null space of $\mathbf{P}_{N+1}$ and the range of $\tilde{\mathbf{P}}_N$. And hence we can have

$$[\tilde{\mathbf{P}}_N - \mathbf{P}_{N+1}]\mathbf{w} = \mathbf{w}$$

which provides us the only non-zero eigenvalue. In summary, it can be proved that $[\tilde{\mathbf{P}}_N - \mathbf{P}_{N+1}] = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$ where $\boldsymbol{\Lambda}$ has only one non-zero eigenvalue which is equal to 1. Since $\mathbf{U}$ is an unitary matrix, the variable $\mathbf{g}_{N+1}$ has the same density function as $\tilde{\mathbf{g}}_{N+1} = \mathbf{U}\mathbf{g}_{N+1}$. By using such a property, we can observe that the difference between the two SNRs can now be expressed as

$$\gamma_{N+1} - \gamma_N = |\tilde{\mathbf{g}}_{N+1}(n)|^2 \tag{46}$$

where n is the index of the non-zero eigenvalue of the matrix. Since $\tilde{\mathbf{g}}_{N+1}$ has the same distribution as $\mathbf{g}_{N+1}$, $\gamma_{N+1} - \gamma_N$ is essentially distributed.

To find the density function of the number transmissions, the joint PDF of $x = \gamma_{N+1} - \gamma_N$ and γ_N is needed. Define $f(\gamma_N, x)$ as their joint density function. So the probability of γ_{N+1} can be expressed as

$$P(\gamma_{N+1} \leq \epsilon) = \int_0^\epsilon \int_0^{\epsilon - \gamma_N} f(\gamma_N, x) dx d\gamma_N$$

Define the conditional density function as $f_{x|\gamma_N}(x)$ and we have $f(\gamma_N, x) = f_{x|\gamma_N}(x)f_{\gamma_N}(\gamma_n)$. Now the probability of γ_{N+1} can be expressed as

$$\begin{aligned} P(\gamma_{N+1} \leq \epsilon) &= \int_0^\epsilon f_{\gamma_N}(\gamma_N) F_{x|\gamma_N}(\epsilon - \gamma_N) d\gamma_N \\ &= \int_0^\epsilon f_{\gamma_N}(\epsilon - z) F_{x|\gamma_N}(z) dz \end{aligned} \tag{47}$$

Recall γ_{N+1} is a Chi-square with $2(N-M+1)$ degree distributed variable, and hence the CDF of γ_{N+1} is the same as the density function of a variable which is the sum of one exponentially distributed variable and one independently Chi-square with $2(N-M+2)$ degree distributed variable. So we can have the following equality

$$\int_0^\epsilon f_{\gamma_N}(\epsilon - z)\left(F_{x|\gamma_N}(z) - F_x(z)\right)dz = 0 \tag{48}$$

which is holds for all $\epsilon \geq 0$. Given the fact that $f_{\gamma_N}(t) \geq 0$, we can conclude that $F_{x|\gamma_N}(z) = F_x(z)$ for all z, which implies the independence between x and γ_N.

$\blacksquare$

REFERENCES

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[2] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Information Theory*, vol. 46, pp. 1204–1217, Jul. 2000.

[3] A. Eryilmaz, A. Ozdaglar, and M. Medard, "On delay performance gains from network coding," in *Proc. 2006 40th Annual Conference on Information Sciences and Systems*, Mar. 2006, pp. 864–870.

[4] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "Xors in the air: Practical wireless network coding," in *Proc. ACM SIGCOMM*, Sept. 2006, pp. 243–254.

[5] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *Proc. ACM SIGCOMM*, Sept. 2007, pp. 397–408.

[6] S. Zhang, S. Liew, and P. Lam, "Physical layer network coding," in *Proc. 12th Annual International Conference on Mobile Computing and Networking (ACM MobiCom 2006)*, Sept. 2006, pp. 63 – 68.

[7] Z. Ding, K. K. Leung, D. L. Goeckel, and D. Towsley, "On the study of network coding with diversity," *IEEE Trans. on Wireless Communications*, vol. 8, pp. 1247 – 1259, Mar. 2009.

[8] Y. Chen, S. Kishore, and J. Li, "Wireless diversity through network coding," in *Proceeding of IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2006, pp. 1681 – 1686.

[9] Z. Ding, T. Ratnarajah, and K. K. Leung, "On the study of network coded af transmission protocol for wireless multiple access channels," *IEEE Trans. on Wireless Communications*, vol. 8, pp. 118–123, Jan. 2009.

[10] I. Maric, A. Goldsmith, and M. Medard, "Information-theoretic relaying for multicast in wireless networks," in *Proc. Military Communications Conference. (MILCOM 2007)*, Oct. 2007, pp. 1 –7.

[11] R. Cogill, B. Shrader, and A. Ephremides, "Stability analysis of random linear coding across multicast sessions," in *Proc. IEEE International Symposium on Information Theory (ISIT-08)*, Toronto, Canda, Jul. 2008, pp. 31–35.

[12] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels : performance limits and space-time signal design," *IEEE Journal on Select. Areas in Comm.*, vol. 22, pp. 1099–1109, Aug. 2004.

[13] S. Boppana and J. M. Shea, "Overlapped carrier-sense multiple access (ocsma) in wireless ad hoc networks," *IEEE Trans. on Mobile Computing*, vol. 8, pp. 369–383, Mar. 2009.

[14] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Information Theory*, vol. 50, pp. 1859–1874, Sept. 2004.

[15] Z. Ding, T. Ratnarajah, and C. Cowan, "On the diversity-multiplexing tradeoff for wireless cooperative multiple access systems," *IEEE Trans. Signal Processing*, pp. 4627–4638, Sept. 2007.

[16] D. A. Gore, R. W. Heath, and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Communications Letters*, vol. 6, pp. 491–493, Nov. 2002.

[17] M. Rupp, C. Mecklenbrauker, and G. Gritsch, "High diversity with simple space time block codes and linear receivers," *Proc. GLOBECOM*, vol. 2, pp. 302–306, Dec. 2003.

[18] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE Journal on Select. Areas in Comm.*, vol. 24, pp. 659–672, Mar. 2006.

[19] H. A. David and H. N. Nagaraja, *Order Statistics*.   John Wiley,Third Edition, 2003.

[20] A. Hedayat and A. Nosratinia, "Outage and diversity of linear receivers in flat fading MIMO channel," *IEEE Trans. on Signal Processing*, vol. 55, pp. 5868–5873, Dec. 2007.

[21] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed.   Academic Press, 2000.

[22] S.Kallel, "Analysis of a type II hybrid ARQ scheme with code combining," *IEEE Trans. Communications*, vol. 38, no. 8, pp. 1133–1137, Aug. 1998.

[23] L. Zheng and D. N. C. Tse, "Diversity and multiplexing : a fundamental tradeoff in multiple antenna channels," *IEEE Trans. Information Theory*, vol. 49, pp. 1073–1096, May 2003.