

RESEARCH

Open Access

Network-coded retransmissions in wireless demodulate-and-forward relay channels

Ye Liu* and Chi Wan Sung

Abstract

Consider the case where a source wishes to reliably deliver a number of data packets to the destination with the help of a demodulate-and-forward relay. We discuss two retransmission schemes, each of which can ensure that the packets from the source will be delivered to the destination without error. We first discuss a scheme that combines signals from the source and the relay at the destination using a weighed combiner. The use of network coding which facilitates the retransmission process is then proposed. The generation of network-coded retransmission packets takes the qualities of previously erroneously received packets at the destination into account, and we show that such network coding scheme can make the retransmission process more efficient even if there is only one destination, unlike the existing approaches where the gain of network coding must come from coding opportunities across multiple destinations. Simulation results show that while the first retransmission scheme performs well in the low signal-to-noise ratio (SNR) region, the second scheme has good performance in both low and high SNR regions, thanks to the use of network coding. Moreover, the second scheme induces smaller packet decoding delay than the first scheme.

1 Introduction

1.1 Motivation and contribution

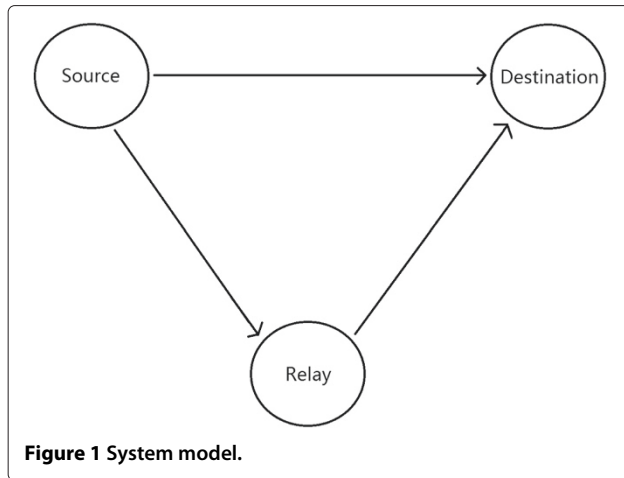
Relay has found itself becoming part of the long term evolution-advanced (LTE-A) due to its ability to extend coverage area and increase users' throughput [1]. Forming a virtual antenna array with the source, works have been done (e.g., [2-4]) so that maximum possible diversity gain can be obtained when the relay works in different modes, namely decode-and-forward (DF), amplify-and-forward, and demodulate-and-forward (DmF). It has been demonstrated that with the help of relays, lower outage error probability can be achieved.

When multiple packets are to be routed to multiple destinations, network coding [5] has been proven to be an efficient solution in many network models. Works have been done to study the advantages of network coding in broadcast channels, including linear network coding [6-9] and non-linear network coding [10]. In DF relay channels, outage analysis was performed in DF relay channels with two destinations in [11], where the use of network coding achieves full diversity with less bandwidth and energy. Retransmission schemes are then designed in [12]

for the same type of network, and network coding opportunities of different lost packets are exploited among destinations, which greatly improves the throughput of the proposed schemes. A network-coded retransmission scheme of multiple sources multiple destinations with the help of a single relay is designed in [13], where the relay performs binary linear network coding on packets from different sources to increase throughput. The use of network coding in DmF multi-user multi-relay channels is also proposed in [14], where a relay is selected to forward network-coded packets by the realized channel gains of the source-to-relay links and relay-to-destination links, and a diversity of two is observed for each source.

In this paper, we discuss the retransmissions of multiple packets in wireless DmF relay channels shown in Figure 1, where the relay is limited to perform demodulation and re-modulation to mitigate the large decoding delay in DF relays. We first design a retransmission scheme based on weighted combining. In this particular scheme, the destination will use a weighted combiner in [3] to combine signals from the source and relay after applying different weights on them. We then demonstrate how a network-coded packet generated by a DmF relay can help the destination to decode the erroneous packets previously received from the source and

*Correspondence: yeliu27-c@my.cityu.edu.hk
Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong



propose a retransmission scheme using both network coding and weighted combiner. We seek for possibilities for the relay to perform network coding based on the signal-to-noise ratios (SNRs) of packets in error at the destination. Rules of forming network-coded packets are very easy to implement. Unlike [6-8] and [12], where the gain of network coding must have come from employing multiple destinations, our approach can make use of network coding to achieve better performance even if there is only one intended destination. The attractive feature of the proposed network-coded retransmission scheme has a high potential to be extended to systems with multiple destinations, which will be a topic for future investigations. Simulation results show that for different relay positions, the second retransmission scheme using network coding requires less retransmissions and has smaller packet decoding delay than the first scheme.

The rest of the paper is organized as follows. In Section 2, the system model will be set up and the weighted combiner will be briefly summarized. In Section 3, retransmission schemes based on the weighted combiner and network coding will be discussed. After that, algorithms for finding network-coded packets will be proposed in Section 4. Simulation results will be given in Section 5. Finally, conclusion and future works will be the content of Section 6.

1.2 Notations

Throughout the paper, we use boldface to represent a column vector (e.g. \mathbf{A} for vector \mathbf{A}) and math calligraphy to represent a set (e.g., \mathcal{B} for set \mathcal{B}). $\|\cdot\|$ denotes the Euclidean norm. $(\cdot)^*$ means the conjugate, $(\cdot)^T$ means the transpose, and $(\cdot)^H$ means conjugate transpose. $\lfloor x \rfloor$ takes

the floor of a real number x . \oplus denotes the bit-wise exclusive OR operation. $Q(\cdot)$ is the Q-function, and $Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-t^2/2) dt$.

2 System model

Consider the relay channel displayed in Figure 1, where the source wishes to transmit its data to the destination via the help of relay. Transmissions are equally slotted in time, and the communication link between any two nodes in the network is modeled as block Rayleigh fading channel, where the real and imaginary parts of the channel coefficients are i.i.d. zero mean Gaussian random variables with variance 0.5 and remain unchanged in a time slot. The received signals at the relay and destination are also corrupted by additive white Gaussian noise with zero mean and σ^2 variance. The relay and the destination have full knowledge of the channel state information of the source-to-relay channel and the source-to-destination channel, respectively, so that coherent detection can be performed. Assume the transmission power of the source and the relay is the same and is normalized to 1. With the realized complex channel coefficient to be h for a communication link, the SNR of a transmitted packet at receiver side is given by $\frac{|h|^2}{\sigma^2}$. We will use h_{SD} , h_{SR} , and h_{RD} to indicate the channel coefficient of the source-to-destination link, the source-to-relay link, and the relay-to-destination link, respectively. The relay is placed somewhere between the source and the destination, and the effect of path loss is taken into consideration.

The source has K source packets of equal number of bits to be transmitted reliably to the destination. All transmissions, either from the source or from the relay, are assigned to different time slots of the same duration. We assume that a fixed-rate ideal channel coding scheme and binary phase-shift keying (BPSK) modulation is applied before sending a packet, where the channel coding scheme is ideal in the sense that decoding error is arbitrarily close to zero if the code rate is less than or equal to the channel capacity. Given a realized SNR Γ at the receiver, the rate $\log(1 + \Gamma)$ bits/s/Hz can be achieved [15]. For a given modulation scheme, the bit error rate (BER) after demodulation is directly related to Γ . Therefore, to correctly receive a packet, the BER after demodulation must be smaller or equal to a threshold value ε_{th} , so that the destination can correctly decode the packet.

The transmission of the K packets can be divided into two phases. In the initial phase, the K source packets are broadcasted to both the relay and the destination. Retransmission phase will start if not all of the K packets are received correctly by the destination. We assume that the control signals, i.e., acknowledgements (ACKs) and negative acknowledgements (NACKs), indicating reception success and failure of the corresponding

packets, respectively, can be reliably received by intended receivers.

When there is no relay, the source can start the retransmission using chase combining-automatic repeat-request (CC-ARQ) scheme [16], where erroneous packets are repeatedly transmitted by applying the same channel coding and modulation scheme. For a packet \mathbf{P}_i that is in error after the initial phase, all available observations at destination are processed by maximal ratio combining (MRC) [17] so that the SNR values of the observations can be added. Once the sum of the SNR values exceeds the *threshold value* T , where

$$T = Q^{-1}(\varepsilon_{\text{th}})^2, \quad (1)$$

the retransmission of \mathbf{P}_i will be halted as \mathbf{P}_i is now reliably received.

If a relay is present, the relay can overhear the transmissions in the initial phase and it may be helpful in the retransmission phase. Limited by its functionality, however, the relay is not able to decode the original data packets from the source; hence, errors will be propagated to the destination if time slots are allocated to the relay in the retransmission phase. Nevertheless, if \mathbf{P}_i is received at the relay with high SNR, then after demodulation and re-modulation, the BER of the re-modulated signal will be small. In that case, the relay is still able to provide useful information on the source packets and help achieve higher retransmission efficiency. A DmF relaying scheme is proposed in [3], where the destination applies a weighted combiner to the incoming signals from the source and the relay, such that a diversity of 2 is achieved. We now describe the relay scheme in [3] for future references.

Consider a symbol vector \mathbf{X} transmitted from the source in the first time slot. Denote the received signals at the relay and the destination as \mathbf{Y}_{SR} and \mathbf{Y}_{SD} , respectively. The relay performs coherent maximal-likelihood demodulation on \mathbf{Y}_{SR} . Denote the re-modulated signal vector at the relay as $\hat{\mathbf{X}}_{\text{R}}$. For BPSK, the i th symbol of $\hat{\mathbf{X}}_{\text{R}}$ satisfies the following relation:

$$\hat{\mathbf{X}}_{\text{R}}(i) = \begin{cases} 1, & \mathbf{Y}_{\text{SR}}(i) \geq 0, \\ -1, & \mathbf{Y}_{\text{SR}}(i) < 0. \end{cases} \quad (2)$$

In the second time slot, the relay transmits $\hat{\mathbf{X}}_{\text{R}}$, and the destination receives \mathbf{Y}_{RD} from the relay. Let Γ_{SD} , Γ_{SR} , and Γ_{RD} be the SNRs of \mathbf{Y}_{SD} , \mathbf{Y}_{SR} , and \mathbf{Y}_{RD} , respectively. The combined symbol vector \mathbf{Y}_{D} can be written as [3]:

$$\mathbf{Y}_{\text{D}} = w_{\text{SD}}\mathbf{Y}_{\text{SD}} + w_{\text{RD}}\mathbf{Y}_{\text{RD}}, \quad (3)$$

where

$$w_{\text{SD}} = h_{\text{SD}}^*, \quad (4)$$

and

$$w_{\text{RD}} = \frac{Q^{-1}[\varepsilon_{\text{eq}}(\Gamma_{\text{SR}}, \Gamma_{\text{RD}})]^2}{\Gamma_{\text{RD}}} h_{\text{RD}}^* \quad (5)$$

are the weights applied to \mathbf{Y}_{SD} and \mathbf{Y}_{RD} , respectively. In (5), $\varepsilon_{\text{eq}}(\Gamma_{\text{SR}}, \Gamma_{\text{RD}})$ is the equivalent BER of a signal at the destination after going through the source-to-relay link, being demodulated and re-modulated at the relay, and going through the relay-to-destination link. The expression of ε_{eq} is given as

$$\varepsilon_{\text{eq}}(\Gamma_{\text{SR}}, \Gamma_{\text{RD}}) = [1 - Q(\sqrt{\Gamma_{\text{SR}}})] Q(\sqrt{\Gamma_{\text{RD}}}) + [1 - Q(\sqrt{\Gamma_{\text{RD}}})] Q(\sqrt{\Gamma_{\text{SR}}}). \quad (6)$$

The BER of \mathbf{Y}_{D} is given as

$$\begin{aligned} \varepsilon_{\text{D}} &= [1 - Q(\sqrt{\Gamma_{\text{SR}}})] \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_{\text{D}}^2}} \\ &\times \exp\left\{-\frac{[x + (w_{\text{SD}}h_{\text{SD}} + w_{\text{RD}}h_{\text{RD}})]^2}{2\sigma_{\text{D}}^2}\right\} dx \\ &+ Q(\sqrt{\Gamma_{\text{SR}}}) \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_{\text{D}}^2}} \\ &\times \exp\left\{-\frac{[x + (w_{\text{SD}}h_{\text{SD}} - w_{\text{RD}}h_{\text{RD}})]^2}{2\sigma_{\text{D}}^2}\right\} dx, \end{aligned} \quad (7)$$

where $\sigma_{\text{D}}^2 = (|w_{\text{SD}}|^2 + |w_{\text{RD}}|^2)\sigma^2$.

It was shown in [3] that the above scheme can achieve the highest possible diversity. In the next section, we will design retransmission schemes based on the above weighted combiner.

3 Proposed retransmission schemes

In this section, we will discuss two retransmission schemes. One is based on the weighted combiner introduced in Section 2, and the other one uses both weighted combiner and network coding.

3.1 Cooperative retransmission

From (7), we can say that when $\varepsilon_{\text{D}} \leq \varepsilon_{\text{th}}$, \mathbf{X} can be correctly decoded. To maximize the efficiency of retransmissions when a relay is available, it is crucial to determine when the relay should retransmit. Also, in order to make sure that a retransmission scheme can always work, the source should delegate the retransmissions to relay only if the relay possesses good enough observations of the source packets. To achieve this robustness, the SNR of a packet received by the relay should exceed T , so that a packet can still be delivered to the destination via the relay in a finite number of retransmissions without knowing the SNR of the packet at the destination. Let the *SNR threshold value at relay* be T_{R} , where packets received by the relay are considered good enough to be forwarded if their SNRs are larger than or equal to T_{R} . It is clear that T_{R} should be set larger than T . After the initial phase, the following threshold-based cooperative retransmission (TCR) scheme will be initiated if there

are packets at the destination that are not decoded correctly. Let $\Gamma_{1,SD}, \Gamma_{2,SD}, \dots, \Gamma_{K,SD}$ be the SNRs of packets 1, 2, ..., K at the destination after the initial phase, respectively. Let $\Gamma_{1,SR}, \Gamma_{2,SR}, \dots, \Gamma_{K,SR}$ be the SNRs of packets 1, 2, ..., K at the relay after the initial phase, respectively. The TCR scheme can be described as follows:

- (1) For the i th packet, $i \in [1, K]$: If $\Gamma_{i,SD} \geq T$, increment i by 1. If $\Gamma_{i,SD} < T$, check the following conditions:
 - (a) $\Gamma_{i,SR} < T_R$,
 - (b) $\Gamma_{i,SR} \geq T_R$.
- (2) If condition (a) is satisfied: The source retransmits \mathbf{X}_i . The relay and the destination combine all different copies of \mathbf{X}_i they received from the source using MRC, so that $\Gamma_{i,SR}$ and $\Gamma_{i,SD}$ are improved. The source stops retransmission until either $\Gamma_{i,SR} \geq T_R$ or $\Gamma_{i,SD} \geq T$. Go back to step 1.
3. If condition (b) is satisfied: The destination finds the required SNR of the receiving signal on the relay-to-destination link, namely Λ_i , such that $\varepsilon_D = \varepsilon_{th}$. The calculation of Λ_i can be done by applying bisection method to (7), after setting $\varepsilon_D = \varepsilon_{th}$, $\Gamma_{SR} = \Gamma_{th,R}$, $w_{SD}h_{SD} = \sigma^2\Gamma_{i,SD}$, $w_{RD}h_{RD} = \frac{Q^{-1}[\varepsilon_{eq}(\Gamma_{i,SR}, \Lambda_i)]^2}{\sigma^2}$. The relay keeps retransmitting $\widehat{\mathbf{X}}_{i,R}$ until the destination has combined enough observations of $\widehat{\mathbf{X}}_{i,R}$ by MRC such that the SNR of the combined copies of $\widehat{\mathbf{X}}_{i,R}$, denoted as $\Gamma_{\widehat{\mathbf{X}}_i}$, is larger than or equal to Λ_i . Go back to step 1.

To implement the TCR scheme, a counting-down method similar to [18] can be adopted. The destination will send ACKs and NACKs sequentially after the source broadcasts all its packets. If the destination sends an ACK for packet i , it will continue sending the control signal (either ACK or NACK) for packet $i + 1$. If the destination sends a NACK for packet i , it will stop and wait for retransmissions of packet i until it can issue an ACK for it. The source and relay can setup their own timers, and the timers are started upon receiving a NACK signal from the destination. The initial value of the timer at the relay can be set according to $\Gamma_{i,SR}$. There will be two situations:

- (1) If $\Gamma_{i,SR} < T_R$, the source's timer will expire first. The source then starts retransmitting the i th packet until either an ACK signal is received from the destination when $\Gamma_i \geq T$ or an ACK signal is received from the relay when $\Gamma_{i,SR} \geq T_R$. In the former case, the retransmission of the i th packet is done. In the latter case, the relay will send an ACK to the source, and the retransmission of the i th packet will be performed by the relay. The retransmission of the i th packet finishes until the destination issues an ACK.

- (2) If $\Gamma_{i,SR} \geq T_R$, the relay's timer will expire first. The relay sends an ACK to the source so that the source will not perform retransmission and then retransmits the i th packet until it receives an ACK from the destination.

Good values of T_R under different channel settings will be obtained via simulations.

3.2 Network-coded cooperative retransmission

In this subsection, we first discuss how network-coded retransmission can be realized with the assumption that the relay has the correct versions of the packets that the source wishes to deliver to the destination. We will then extend the idea to DmF relays, where the relay may not have exact copies of the source's packets.

3.2.1 Network-coded retransmission

Assume that the source wishes to deliver \mathbf{P}_1 and \mathbf{P}_2 to the destination, and the relay has the two packets. After the initial phase, the destination cannot correctly decode the two packets. The two packets' SNRs at the destinations are $\Gamma_{1,SD}$ and $\Gamma_{2,SD}$, respectively, where $\Gamma_{1,SD} + \Gamma_{2,SD} \geq T$. Denote the received signals of \mathbf{P}_1 and \mathbf{P}_2 at destination after the initial phase as $\mathbf{Y}_{1,SD}$ and $\mathbf{Y}_{2,SD}$, respectively. We have

$$\begin{aligned} \mathbf{Y}_{2,SD} &= h_{2,SD}\mathbf{X}_2 + \mathbf{z} \\ \Rightarrow \frac{h_{2,SD}^H}{h_{2,SD}^H h_{2,SD}} \mathbf{Y}_{2,SD} &= \mathbf{X}_2 + \frac{h_{2,SD}^H}{h_{2,SD}^H h_{2,SD}} \mathbf{z}, \end{aligned} \quad (8)$$

where $h_{2,SD}$ is the channel coefficient of the source-to-destination link when \mathbf{P}_2 is sent by the source. We further assume that $\mathbf{P}_1 \oplus \mathbf{P}_2$ is transmitted by the relay and is correctly received by the destination. For ease of further development, we give the following definition of *element-by-element multiplication* between two vectors of equal length.

Definition 1. Let \star denote the element-by-element multiplication between two equal length vectors \mathbf{A} and \mathbf{B} , where $\mathbf{A} \star \mathbf{B} = \mathbf{C}$, $\mathbf{A} = [a_1, a_2, \dots, a_L]$, $\mathbf{B} = [b_1, b_2, \dots, b_L]$, $\mathbf{C} = [c_1, c_2, \dots, c_L]$, such that $c_i = a_i b_i$, $i \in [1, L]$.

After correctly receiving $\mathbf{P}_1 \oplus \mathbf{P}_2$, the destination can perform the following: First, it encodes $\mathbf{P}_1 \oplus \mathbf{P}_2$ into $\mathbf{X}_{1\oplus 2}$ using the same channel coding and modulation schemes as the source's, and then performs element-by-element multiplication between $\mathbf{X}_{1\oplus 2}$ and $\frac{h_{2,SD}^H}{h_{2,SD}^H h_{2,SD}} \mathbf{Y}_{2,SD}$. Since BPSK modulation is applied and $\mathbf{X}_{1\oplus 2}$ is equal to $\mathbf{X}_1 \star \mathbf{X}_2$, we have

$$\begin{aligned}
 & \mathbf{X}_{1\oplus 2} \star \frac{h_{2,SD}^H}{h_{2,SD}^H h_{2,SD}} \mathbf{Y}_{2,SD} \\
 = & \mathbf{X}_1 \star \mathbf{X}_2 \star \mathbf{X}_2 + \mathbf{X}_1 \star \mathbf{X}_2 \star \frac{h_{2,SD}^H}{h_{2,SD}^H h_{2,SD}} \mathbf{z} \\
 = & \mathbf{X}_1 + \mathbf{X}_1 \star \mathbf{X}_2 \star \frac{h_{2,SD}^H}{h_{2,SD}^H h_{2,SD}} \mathbf{z}. \quad (9)
 \end{aligned}$$

We can see from (9) that by such a process, we come up with another noisy observation of \mathbf{P}_1 . The SNR of this observation is equal to $\Gamma_{2,SD}$, since the magnitude of each component of $\mathbf{X}_1 \star \mathbf{X}_2$ is equal to one, and the noise term's power is unchanged. Next, the destination combines $\mathbf{X}_{1\oplus 2} \star \frac{h_{2,SD}^H}{h_{2,SD}^H h_{2,SD}} \mathbf{Y}_{2,SD}$ and $\mathbf{Y}_{1,SD}$ together using the MRC principle, and a noisy observation of \mathbf{P}_1 with SNR $\Gamma_{1,SD} + \Gamma_{2,SD}$ is formed. By the previous channel coding assumption, \mathbf{P}_1 is now correctly received as $\Gamma_{1,SD} + \Gamma_{2,SD} \geq T$. With \mathbf{P}_1 and $\mathbf{P}_1 \oplus \mathbf{P}_2$, \mathbf{P}_2 can be derived by the exclusive OR operation. From this example, it is illustrated that instead of simply repeating the two packets that the destination needs, the relay has an alternative: to transmit the network coded packet, under the condition that the sum SNRs of the two packets at destination is larger than T . The number of retransmitted packets is then reduced from two to one. The idea can also be extended to situations when higher modulation schemes are used, which is discussed in Appendix 1.

In general, if $\Gamma_{1,SD} + \Gamma_{2,SD} + \dots + \Gamma_{m,SD} \geq T$, the relay can send $\mathbf{P}_1 \oplus \mathbf{P}_2$, $\mathbf{P}_1 \oplus \mathbf{P}_3$, ..., $\mathbf{P}_1 \oplus \mathbf{P}_m$ to the destination. The destination will have $m - 1$ more copies of \mathbf{P}_1 with SNRs $\Gamma_{2,SD}$, $\Gamma_{3,SD}$, ..., $\Gamma_{m,SD}$ once it receives all the network-coded packets correctly. Combining all the m copies of \mathbf{P}_1 (the underlying assumption is that the relay gets \mathbf{P}_1 correctly in the initial transmission) with MRC will lead to successful decoding of \mathbf{P}_1 , and then \mathbf{P}_2 , \mathbf{P}_3 , ..., \mathbf{P}_m can be derived. We term such a combination of m packets' SNRs as a *valid m -wise combination*.

3.2.2 Network-coded retransmission in DmF relay networks

In a DmF relay, as decoding is not performed, the re-modulated signal may contain errors. Nevertheless, it is still possible for a DmF relay to forward network-coded packets, and it is also possible for the destination to decode the network-coded packets from the relay correctly, which is demonstrated in the following:

Consider the transmission of two packets \mathbf{P}_1 and \mathbf{P}_2 . After the initial phase, the destination has \mathbf{P}_1 's SNR $\Gamma_{1,SD}$ and \mathbf{P}_2 's SNR $\Gamma_{2,SD}$, where $\Gamma_{1,SD} < T$, $\Gamma_{2,SD} < T$, and $\Gamma_{1,SD} + \Gamma_{2,SD} \geq T$. The relay has observations of \mathbf{P}_1 and \mathbf{P}_2 with SNRs $\Gamma_{1,SR}$ and $\Gamma_{2,SR}$, respectively. After demodulation and re-modulation, the relay has $\widehat{\mathbf{X}}_{1,R}$ and $\widehat{\mathbf{X}}_{2,R}$, and then the network-coded symbol vector $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$ can be

formed. Note that $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$ is essentially a noisy observation of $\mathbf{P}_1 \oplus \mathbf{P}_2$, with BER $P_{\text{athrmbeq}}^m(\Gamma_{1,SR}, \Gamma_{2,SR})$ (see (6)). Let

$$\begin{aligned}
 & \Gamma_{\text{eq}}(\Gamma_{1,SR}, \Gamma_{2,SR}) \\
 = & \{Q^{-1}[\varepsilon_{\text{eq}}(\Gamma_{1,SR}, \Gamma_{2,SR})]\}^2 \\
 = & \{Q^{-1}\{[1 - Q(\sqrt{\Gamma_{1,SR}})] Q(\sqrt{\Gamma_{2,SR}}) + [1 \\
 & - Q(\sqrt{\Gamma_{2,SR}})] Q(\sqrt{\Gamma_{1,SR}})]\}^2. \quad (10)
 \end{aligned}$$

The above equation indicates that now the relay has a noisy observation of $\mathbf{P}_1 \oplus \mathbf{P}_2$ with SNR $\Gamma_{\text{eq}}(\Gamma_{1,SR}, \Gamma_{2,SR})$. The relay retransmits $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$, and the destination gets a noisy observation of it with SNR Γ_{RD} . At this point, we can also say that the destination has a noisy observation of $\mathbf{P}_1 \oplus \mathbf{P}_2$ with SNR $\Gamma_{\text{eq}}(\Gamma_{\text{eq}}(\Gamma_{1,SR}, \Gamma_{2,SR}), \Gamma_{RD})$. If $\Gamma_{\text{eq}}(\Gamma_{\text{eq}}(\Gamma_{1,SR}, \Gamma_{2,SR}), \Gamma_{RD}) \geq T$, then we can say that $\mathbf{P}_1 \oplus \mathbf{P}_2$ is correctly received by the destination.

In order to ensure that the network-coded packet $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$ can be decoded correctly by the destination, the BER of $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$ at the relay must be smaller than or equal to ε_{th} ; otherwise, the BER of the packet received by the destination will never be smaller than or equal to ε_{th} , even if the relay-to-destination is noise-free. Therefore, $\Gamma_{\text{eq}}(\Gamma_{1,SR}, \Gamma_{2,SR})$ in (10) must be no smaller than T . This imposes the following constraint on $\Gamma_{1,SR}$ and $\Gamma_{2,SR}$:

$$\begin{aligned}
 & \Gamma_{1,SR} > T, \\
 & \Gamma_{2,SR} > T, \quad (11)
 \end{aligned}$$

due to the reason that for finite $\Gamma_{1,SR}$ and $\Gamma_{2,SR}$,

$$\Gamma_{\text{eq}}(\Gamma_{1,SR}, \Gamma_{2,SR}) < \min(\Gamma_{1,SR}, \Gamma_{2,SR}). \quad (12)$$

To see why (12) is the case, with some abuse of notations, one realizes that if $\Gamma_{2,SR} = +\infty$, then

$$\begin{aligned}
 \Gamma_{\text{eq}}(\Gamma_{1,SR}, +\infty) &= \{Q^{-1}[P_{\text{eq}}^b(\Gamma_{1,SR}, +\infty)]\}^2 \\
 &= \{Q^{-1}\{[1 - Q(\sqrt{\Gamma_{1,SR}})] Q(\sqrt{+\infty}) \\
 &\quad + [1 - Q(\sqrt{+\infty})] Q(\sqrt{\Gamma_{1,SR}})]\}^2 \\
 &= \{Q^{-1}[Q(\sqrt{\Gamma_{1,SR}})]\}^2 \\
 &= \Gamma_{1,SR}. \quad (13)
 \end{aligned}$$

Also, if we fix $\Gamma_{1,SR}$ to be c and further reduce (10), we have

$$\begin{aligned}
 \Gamma_{\text{eq}}(c, \Gamma_{2,SR}) &= \{Q^{-1}\{[1 - Q(\sqrt{c})] Q(\sqrt{\Gamma_{2,SR}}) + [1 \\
 &\quad - Q(\sqrt{\Gamma_{2,SR}})] Q(\sqrt{c})]\}^2 \\
 &= \{Q^{-1}\{[1 - 2Q(\sqrt{c})] Q(\sqrt{\Gamma_{2,SR}}) \\
 &\quad + Q(\sqrt{c})]\}^2. \quad (14)
 \end{aligned}$$

Since $Q(\sqrt{c}) < 0.5$, $[1 - 2Q(\sqrt{c})] Q(\sqrt{\Gamma_{2,SR}}) + Q(\sqrt{c})$ will be a monotonic decreasing function of $\Gamma_{2,SR}$. Furthermore, because $Q^{-1}(x)$ is a monotonic decreasing function of x , where $0 \leq x \leq 0.5$, $\Gamma_{\text{eq}}(c, \Gamma_{2,SR})$ will be a monotonic

increasing function of $\Gamma_{2,SR}$. We can then infer (12) from (13) and the monotonicity of (14).

Inspired by (10), to ensure that the relay can provide any necessary network-coded packet that can be ultimately received by the destination correctly, it is desirable to set a threshold value *at the relay*, namely $T_{R,NC}$, and make sure that each requested packet by the destination has SNR larger or equal to $T_{R,NC}$ at the relay. The value of $T_{R,NC}$ must be larger than T by (11). After setting $T_{R,NC}$, the required realized SNR of a network-coded packet *at the destination*, namely Λ_{NC} , can be found by

$$\begin{aligned} \Gamma_{eq}(\Lambda_{NC}, T_{R,NC}) &= T \\ \Rightarrow \{Q^{-1}\{[1 - Q(\sqrt{\Lambda_{NC}})] Q(\sqrt{T_{R,NC}}) + [1 \\ &\quad - Q(\sqrt{T_{R,NC}})] Q(\sqrt{\Lambda_{NC}})\}\}^2 = T \\ \Rightarrow \Lambda_{NC} &= \{Q^{-1}\{ \\ &\quad \frac{Q(\sqrt{T}) - Q(\sqrt{T_{R,NC}})}{1 - 2Q(\sqrt{T_{R,NC}})}\}\}^2. \end{aligned} \quad (15)$$

In the situations where $\Gamma_{1,SD} + \Gamma_{2,SD} < T$, sending $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$ will not be able to help the destination decode \mathbf{P}_1 or \mathbf{P}_2 correctly. However, as $Q(\sqrt{\Gamma_{1,SR}}) \leq Q(\sqrt{T_{R,NC}}) < \varepsilon_{th}$, the relay has a good enough observation of \mathbf{P}_1 and thus able to help the destination correctly receiving \mathbf{P}_1 by retransmitting the demodulated and re-modulated version of \mathbf{P}_1 (i.e., $\widehat{\mathbf{X}}_{1,R}$). In general, the relay can help the destination correctly receiving \mathbf{P}_i by retransmitting $\widehat{\mathbf{X}}_{i,R}$, if $\Gamma_{i,SR} \geq T_{R,NC}$. The required SNR of received $\widehat{\mathbf{X}}_{i,R}$ at the destination, namely Λ_i , can be calculated by applying bisection method to (7) as in the TCR scheme. Good $T_{R,NC}$ values under different channel settings will be obtained via simulations.

In general, if $\Gamma_{1,SD} + \Gamma_{2,SD} + \dots + \Gamma_{m,SD} \geq T$ and $\Gamma_{i,SR} \geq T_{R,NC}$, $i \in [1, m]$, the relay can retransmit $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$, $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{2,R}$, ..., $\widehat{\mathbf{X}}_{1,R} \star \widehat{\mathbf{X}}_{m,R}$, each of which must be received by the destination with SNR larger than or equal to Λ_{NC} .

We now propose the threshold-based network-coded cooperative retransmission (TNCCR) scheme, which will be initiated if there are packets that are not correctly received by the destination after the initial phase:

- (1) For \mathbf{P}_i , $i \in [1, K]$ such that $\Gamma_{i,SD} < T$, if $\Gamma_{i,SR} < T_{R,NC}$, then the source retransmits \mathbf{P}_i until $\Gamma_{i,SD} \geq T$ or $\Gamma_{i,SR} \geq T_{R,NC}$. Let $(\Gamma_{a_1}, \Gamma_{a_2}, \dots, \Gamma_{a_k})$ be the k -tuple consisting of all the SNRs at the destination that are less than T , where $k \leq K$.
- (2) Find valid combinations in $(\Gamma_{a_1}, \Gamma_{a_2}, \dots, \Gamma_{a_k})$. Let VC_m be the number of m -wise valid combinations found, where $m \in [2, \omega]$ (i.e., there exists a ω -wise valid combination that is formed by the sum of most SNRs among all valid combinations). Starting from

the smallest m , where $VC_m > 0$, the relay retransmits the network-coded packets corresponding to one of the m -wise valid combinations. Each of the network-coded packets is repeatedly retransmitted until its SNR exceeds Λ_{NC} .

- (3) For \mathbf{P}_i , $i \in [1, K]$ that are neither involved in any valid combinations nor received reliably at the destination, the relay retransmits $\widehat{\mathbf{P}}_i$ until the SNR of $\widehat{\mathbf{P}}_i$ exceeds Λ_i .

The implementation of TNCCR can be done as follows:

- (1) For \mathbf{P}_i that is not correctly received at the destination, do the following: the destination broadcasts a NACK for the packet. Upon receiving the NACK of \mathbf{P}_i , the relay checks whether $\Gamma_{i,SR} < T_{R,NC}$. If it is the case, then the relay sends a NACK, and the source will start retransmitting \mathbf{P}_i . The source stops sending \mathbf{P}_i if it receives an ACK from the relay (when $\Gamma_{i,SR} \geq T_{R,NC}$) or an ACK from the destination (when $\Gamma_{i,SD} \geq T$).
- (2) If there are still packets not correctly decoded at the destination, the destination will look for valid combinations. For the i th valid combination, the destination sends a signal requesting the corresponding network-coded packet(s) with specific indices to the relay. The relay will retransmit the corresponding network-coded packet until the SNR of that packet exceeds Λ_{NC} .
- (3) After receiving all the required network-coded packets correctly, the destination requests the relay for each of the packets that is neither involved in a valid combination nor received correctly.

3.2.3 Delay consideration

We now consider the *average packet decoding delay* [19] of different retransmission schemes, which is an important measure for applications where the sequence of decoding packets does not matter (i.e., multiple-description coding). We first give the definition of packet decoding delay:

Definition 2. The packet decoding delay of a source packet is the number of retransmissions before it is correctly decoded at the receiver.

The following example demonstrates how packet decoding delay is computed.

Example 1. Consider a relay network using TNCCR as its retransmission scheme, where $K = 6$, $\Gamma_1 = \Gamma_2 = 0.5$, $\Gamma_3 = \Gamma_4 = \Gamma_5 = 1/3$, $\Gamma_6 = 0$, and $T = 1$. Assume $\Gamma_{i,SR} \geq T_{R,NC}, \forall i \in [1, 6]$ and the retransmissions are error-free. The relay will retransmit the following packets in sequence: $\mathbf{P}_1 \oplus \mathbf{P}_2$, $\mathbf{P}_3 \oplus \mathbf{P}_4$, $\mathbf{P}_3 \oplus \mathbf{P}_5$, and \mathbf{P}_6 . After the first retransmission, \mathbf{P}_1 and \mathbf{P}_2 can be decoded; therefore,

the packet decoding delay for \mathbf{P}_1 and \mathbf{P}_2 is 1. To decode \mathbf{P}_3 , \mathbf{P}_4 , and \mathbf{P}_5 , the destination has to wait for 3 retransmissions (i.e., the retransmission of $\mathbf{P}_1 \oplus \mathbf{P}_2$, $\mathbf{P}_3 \oplus \mathbf{P}_4$, and $\mathbf{P}_3 \oplus \mathbf{P}_5$), so the packet decoding delays for \mathbf{P}_3 , \mathbf{P}_4 , and \mathbf{P}_5 are all equal to 3. The packet decoding delay for \mathbf{P}_6 is 4.

In general, let the destination request retransmission of $\mathbf{P}'_1, \mathbf{P}'_2, \dots, \mathbf{P}'_n$, where $\mathbf{P}'_i, i \in [1, n]$ can be a source packet (i.e., $\mathbf{P}_r, r \in [1, K]$) or a network-coded packet (i.e., $\mathbf{P}_r \oplus \mathbf{P}_k, r, k \in [1, K]$). Upon reception of $\mathbf{P}'_i, i \in [1, n]$, l_i source packets can be decoded. Let $\mathbf{P}'_i, i \in [1, n]$ requires μ_i retransmissions before being correctly received by the destination. Starting from $i = 1$, the destination will only request \mathbf{P}'_{i+1} if \mathbf{P}'_i is received correctly. The *total packet decoding delay* is given as

$$l_1\mu_1 + l_2(\mu_1 + \mu_2) + \dots + l_n \sum_{i=1}^n \mu_i, \quad (16)$$

and the *average packet decoding delay* is

$$\frac{l_1\mu_1 + l_2(\mu_1 + \mu_2) + \dots + l_n \sum_{i=1}^n \mu_i}{\sum_{i=1}^n l_i}, \quad (17)$$

where $\sum_{i=1}^n l_i$ is the number of source packets that can be decoded after receiving $\mathbf{P}'_1, \mathbf{P}'_2, \dots, \mathbf{P}'_n$.

Example 2. Following Example 1, the average packet decoding delay of retransmitting $\mathbf{P}_1 \oplus \mathbf{P}_2, \mathbf{P}_3 \oplus \mathbf{P}_4, \mathbf{P}_3 \oplus \mathbf{P}_5$, and \mathbf{P}_6 through error-free channels can be calculated as

$$\frac{2 \times 1 + 0 \times 2 + 3 \times 3 + 1 \times 4}{2 + 0 + 3 + 1} = 2.5,$$

suggesting that, on average, the receiver needs to wait for 2.5 retransmissions before decoding a packet.

Next, we show that under some specific conditions, the scheduling of retransmitting network-coded packets in TNCCR is optimum.

Proposition 1. Assume that for each $i \in [1, K]$, where $\Gamma_i < T$, $\Gamma_{i,SR} \geq T_{R,NC}$, and only pairwise and three-wise valid combinations are considered. When the retransmissions are error-free, the scheduling of retransmitting packets in step 2 of TNCCR leads to the minimum average packet decoding delay.

Proof. See Appendix 2. \square

The general scheduling problem that minimizes the average packet decoding delay is however involved and is out of the scope of this paper. In the next section, we will discuss algorithms for finding valid combinations at the destination. The problem is in general hard and requires

high computational complexity. We will show the complexity of the problem and then propose some efficient algorithms.

4 Algorithms searching for valid combinations

After step 1 in TNCCR, we still have k packets yet to be decoded at the destination. Without loss of generality, we assume $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_k$ are still in error at the destination. The problem of finding the maximum number of valid combinations can be stated as follows:

$$\begin{aligned} & \text{given } \mathcal{W} = \{1, 2, \dots, k\}, \Gamma_i < T, i \in [1, k] \\ & \max t, \\ & \text{subject to } t \in \mathbb{Z}^+, \\ & \mathcal{S}_1 \cup \mathcal{S}_2 \cup \dots \cup \mathcal{S}_t \cup \mathcal{B} = \mathcal{W}, \\ & \mathcal{S}_i \cap \mathcal{S}_r = \emptyset, 1 \leq i < r \leq t, \\ & \mathcal{S}_k \cap \mathcal{B} = \emptyset, 1 \leq k \leq t, \\ & \sum_{i \in \mathcal{S}_r} \Gamma_i \geq T, 1 \leq r \leq t, \end{aligned} \quad (18)$$

where \mathcal{B} is the set that contains indices of packets which cannot find themselves involved in any valid combinations. To illustrate an instance of the above problem, suppose $\Gamma_1 = 1.5, \Gamma_2 = 1.2, \Gamma_3 = 0.8, \Gamma_4 = 0.5, \Gamma_5 = 0.2$, and $T = 2$. One solution can be $\mathcal{S}_1 = \{1, 4\}, \mathcal{S}_2 = \{2, 3\}$, and $\mathcal{B} = \{5\}$, so that $t = 2$. This means that instead of retransmitting $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_5$, we can retransmit $\mathbf{P}_1 \oplus \mathbf{P}_4, \mathbf{P}_2 \oplus \mathbf{P}_3$, and \mathbf{P}_5 , such that $t = 2$ transmissions are saved.

The problem in (18) is actually equivalent to the bin-covering problem (or dual bin-packing problem) [20], where the bin-covering problem is defined as packing items of various sizes from a given set into bins so that the maximum number of bins are used, subject to the constraint that each bin must be filled to at least a given threshold. The problem is NP-hard, which means there is no polynomial-time algorithm available to solve it. The upper bound on t in (18) is

$$\text{Upperbound}_{\# \text{Valid_Comb.}} = \left\lfloor \frac{\sum_{i \in \mathcal{W}} \Gamma_i}{T} \right\rfloor. \quad (19)$$

It is easy to see that (19) gives the maximum number of valid possible combinations and sometimes overestimates it. For example, let $\Gamma_1 = \Gamma_2 = 1.8, \Gamma_3 = 0.4, \Gamma_4 = 0.1$, and $T = 2$. In this case, only one valid combination can be found in $\mathcal{W} = \{1, 2, 3, 4\}$, while (19) gives the value 2.

Next, we develop polynomial-time algorithms for two simpler versions of the problem. We will demonstrate later that the number of valid combinations found in the simpler problems is not significantly smaller than the upper bound in (19).

4.1 Pairwise combinations only

We reduce the complexity of (18) by adding one more constraint, namely, we only concentrate on finding combinations of SNRs that involve two elements. The formal statement of the simpler problem is the following:

$$\begin{aligned} & \text{given } \mathcal{W} = \{1, 2, \dots, k\}, \Gamma_i < T, i \in [1, k], \nu = 2, \\ & \max t, \\ & \text{subject to } S_1 \cup S_2 \cup \dots \cup S_t \cup \mathcal{B} = \mathcal{W}, \\ & S_i \cap S_r = \emptyset, 1 \leq i < r \leq t, \\ & S_k \cap \mathcal{B} = \emptyset, 1 \leq k \leq t, \\ & \sum_{i \in S_r} \Gamma_i \geq T, 1 \leq r \leq t, \\ & 1 < |S_i| \leq \nu, \forall i \in [1, t]. \end{aligned} \quad (20)$$

Before we develop algorithms for finding valid combinations, we first present a lemma that can help design efficient algorithms.

Lemma 1. *There is an optimum solution for (20), where $\arg \max_i \{\Gamma_i | i \in \mathcal{W}\}$ is involved in a valid combination, if there exists at least one valid combination.*

Proof. Let $l = \arg \max_i \{\Gamma_i | i \in \mathcal{W}\}$. If there is an optimum solution that does not use l in any valid combination, then swapping l with any element involved in any valid combination found will not affect the optimality of the solution. \square

Inspired by Lemma 1, we may first find a valid pairwise combination involving the largest element in \mathcal{W} . If such a valid combination is found, we can start looking at the second largest element in \mathcal{W} and try to find a valid combination involving it. The details of the algorithm are the following:

Algorithm 1 Find pairwise valid combinations

Input: $\mathcal{W} = \{1, 2, \dots, k\}, \Gamma_1, \Gamma_2, \dots, \Gamma_k, T$.

Output: $S_1, S_2, \dots, S_{t_1}, \mathcal{B}$, where $S_1 \cup S_2 \cup \dots \cup S_{t_1} \cup \mathcal{B} = \mathcal{W}$.

```

1:  $\Phi = (\Gamma_1, \Gamma_2, \dots, \Gamma_k)$ .
2: Sort  $\Phi$  in descending order, then  $\Phi = (\Gamma_{a_1}, \Gamma_{a_2}, \dots, \Gamma_{a_k})$ .
3:  $i = 1, r = N, c = 0$ .
4: while  $i < r$  do
5:   if  $\Gamma_{a_i} + \Gamma_{a_r} \geq T$  then
6:      $c = c + 1, S_c = \{a_i, a_r\}$ 
7:      $i = i + 1, r = r - 1$ 
8:   else
9:      $r = r - 1$ 
10:  end if
11: end while

```

At the end of Algorithm 1, we can represent \mathcal{W} as $\mathcal{W} = S_1 \cup S_2 \cup \dots \cup S_c \cup \mathcal{B}$, where $S_i, i \in [1, t_1]$ is a set that contains indices of packets involved in a pairwise valid combination, and $|S_i| = 2, \mathcal{B}$ is the set that contains indices of packets not involved in any pairwise valid combination.

Theorem 1. *Algorithm 1 maximizes t in (20).*

Proof. Assume that for a given \mathcal{W} , there is at least one valid combination. Also, assume that $\{\Gamma_{a_1}, \Gamma_{a_q}\}$ is found as a valid pairwise combination by Algorithm 1.

From Lemma 1, we know there is an optimum solution, where Γ_{a_1} is involved in a valid combination. Let $\{\Gamma_{a_1}, \Gamma_{a_p}\}$ be a valid combination in the optimum solution. If $a_p = a_q$, then Algorithm 1 has found the valid combination in the optimum solution. Otherwise, if $a_p \neq a_q$, from the mechanism of Algorithm 1, Γ_{a_q} will be the smallest element, such that $\{\Gamma_{a_1}, \Gamma_{a_q}\}$ can form a valid combination, so that $\Gamma_{a_p} \geq \Gamma_{a_q}$. In this case, swapping Γ_{a_p} with Γ_{a_q} will not affect the optimality of the solution.

The analysis above is carried over to other valid combinations found by Algorithm 1 after $\{\Gamma_{a_1}, \Gamma_{a_q}\}$.

Since each valid combination chosen by Algorithm 1 does not affect the optimality of its final solution, we conclude that Algorithm 1 can find an optimum solution of (20) when $\nu = 2$. \square

The complexity of Algorithm 1 is $O(K \log K)$, since the sorting in the second step has complexity of $O(K \log K)$ and other steps has complexity of $O(K)$. It should be noted that when the retransmission is error-free and only pairwise valid combinations are considered, TNCCR with Algorithm 1 will have the least number of retransmissions it can ever achieve. The design of an optimum scheme when the retransmissions are subject to fading in terms of number of retransmissions is, however, involved and out of the scope of this work.

4.2 Pairwise and three-wise valid combinations

Now, we focus on the case where $\nu = 3$ in (20), which means we consider pairwise and three-wise combinations. We first prove that the decision version of the problem is NP-complete by reduction from a NP-complete problem called three-partition [21], and then discuss a heuristic algorithm for it.

The decision version of (20) can be stated as follows:

Instance: k numbers $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ and the corresponding indices set $\mathcal{W} = \{1, 2, \dots, k\}$, where $\Gamma_i < T, i \in [1, k]$.

Question: Can we find at least $\lfloor \frac{N}{\nu} \rfloor$ disjoint subsets of \mathcal{W} such that the sum of the corresponding real numbers of each subset is larger or equal to T , where each subset consists at most ν elements?

For future reference, we call this problem ν -cardinality-constrained bin-covering (ν -CCBC). The three-partition problem is described for ease of reference [21]:

Instance: A finite set \mathcal{A} of $3m$ elements, a bound $B \in \mathbb{Z}^+$, and a 'size' $s(a) \in \mathbb{Z}^+$ for each $a \in \mathcal{A}$, such that each $s(a)$ satisfies $\frac{B}{4} < s(a) < \frac{B}{2}$ and such that $\sum_{a \in \mathcal{A}} s(a) = mB$.

Question: Can \mathcal{A} be partitioned into m disjoint sets S_1, S_2, \dots, S_m , such that for $1 \leq i \leq m$, $\sum_{a \in S_i} s(a) = B$?

Theorem 2. 3-CCBC is NP-complete.

Proof. We wish to solve the three-partition problem instance using the algorithm solving 3-CCBC. Let $\mathcal{W} = 1, 2, \dots, 3m$, and Γ_i be $s(a)$, where $1 \leq i \leq 3m$ and $a \in \mathcal{A}$. Also, let $T = B$. Then, solve the 3-CCBC problem. If the answer is YES, then we find m subsets of \mathcal{W} as $\frac{B}{4} < \Gamma_i < \frac{B}{2}$, which implies that we can find m subsets of \mathcal{A} of the three-partition problem satisfying the respective constraints. Similarly, if the answer is NO, then we cannot find m subsets of \mathcal{A} that satisfy the constraints. The reduction from three-partition to 3-CCBC is obviously in polynomial time, so that 3-CCBC is NP-hard.

As an output of the algorithm of 3-CCBC can be verified in polynomial time, 3-CCBC is in NP. Thus, 3-CCBC is NP-complete. \square

The above theorem implies that achieving the minimum number of retransmissions using TNCCR when the retransmissions are error-free and if pairwise and three-wise valid combinations are allowed would have exponential complexity, which is impractical to implement.

We now focus on solving 3-CCBC by a heuristic. The idea of the algorithm is the following: First, apply Algorithm 1 to find pairwise combinations. Then, let \mathcal{U}' be a set that contains elements in \mathcal{W} that are not involved in any pairwise combinations, search for three-wise combinations in \mathcal{U}' . The searching for three-wise combinations similar to Algorithm 1 is as follows: We first search for a valid combination involving the largest element in \mathcal{U}' . Then, find two other elements Γ_i and Γ_r , such that the sum SNR of these two elements is the smallest, while the sum SNR of Γ_i , Γ_r , and the largest element is larger than or equal to T . The same approach is carried out until no more valid three-wise combinations can be found. The details of the algorithm for $\nu = 3$ is presented in *Algorithm 2*.

In step 7 of Algorithm 2, Algorithm 1 can be used to find x_1 and x_2 , such that $\Gamma_{x_1} + \Gamma_{x_2} \geq T - \Gamma_{b'_1}$ by changing the while-loop condition to 'no valid combination found yet' and T to $T - \Gamma_{b'_1}$. After running Algorithm 2, we have S_1, S_2, \dots, S_{t_1} , where each of them has indices of SNRs that form a valid pairwise combination. $S'_1, S'_2, \dots, S'_{t_2}$, where each of them has indices of SNRs that form a valid three-wise combination. \mathcal{B} contains indices that the

Algorithm 2 Find pairwise and three-wise valid combinations

Input: $\mathcal{W} = \{1, 2, \dots, k\}, \Gamma_1, \Gamma_2, \dots, \Gamma_k, T$.

Output: $S_1, S_2, \dots, S_{t_1}, S'_1, S'_2, \dots, S'_{t_2}, \mathcal{B}$, where $S_1 \cup S_2 \cup \dots \cup S_{t_1} \cup S'_1 \cup S'_2 \cup \dots \cup S'_{t_2} \cup \mathcal{B} = \mathcal{W}$.

- 1: Run Algorithm 1 and get $S_1, S_2, \dots, S_{t_1} \subset \mathcal{W}$.
- 2: $\mathcal{U}' = \mathcal{W} \setminus \bigcup_{i=1}^{t_1} S_i = \{b_1, b_2, \dots, b_{k-2t_1}\}$.
- 3: $t_2 = 0$.
- 4: $\Psi = (\Gamma_{b_1}, \Gamma_{b_2}, \dots, \Gamma_{b_{k-2t_1}}) = (\Gamma_{b_1}, \Gamma_{b_2}, \dots, \Gamma_{b_{k-2t_1-3t_2}})$.
- 5: Sort Ψ in descending order, then $\Psi = (\Gamma_{b'_1}, \Gamma_{b'_2}, \dots, \Gamma_{b'_{k-2t_1-3t_2}})$.
- 6: If there are at least three elements in Ψ , go to the next step; otherwise, go to the last step.
- 7: Find two elements, namely Γ_{x_1} and Γ_{x_2} , in $(\Gamma_{b'_2}, \Gamma_{b'_3}, \dots, \Gamma_{b'_{k-2t_1}})$, such that $\Gamma_{x_1} + \Gamma_{x_2} \geq T - \Gamma_{b'_1}$. If no such $\Gamma_{x_1} + \Gamma_{x_2}$ can be found, go to the last step. Otherwise, go to the next step.
- 8: $t_2 = t_2 + 1$, $S'_{t_2} = \{b'_1, x_1, x_2\}$. Remove $\Gamma_{b'_1}, \Gamma_{x_1}$, and Γ_{x_2} from Ψ . Rename the remaining elements in Ψ so that $\Psi = (\Gamma_{b'_1}, \Gamma_{b'_2}, \dots, \Gamma_{b'_{k-2t_1-3t_2}})$. $\mathcal{U}' = \mathcal{U}' \setminus S'_{t_2}$. Go to step 6.
- 9: $\mathcal{B} = \mathcal{U}'$.

corresponding SNRs are not involved in any valid combinations.

The complexity of the first step in Algorithm 2 is $O(K \log K)$. The complexity of the other steps is $O(K^2)$, since for each $\Gamma_{b'_i}$, it takes a number of steps linear to K to complete step 7; step 7 can be executed for at most $\lfloor \frac{K}{3} \rfloor$ times. So, in total, the complexity of Algorithm 2 is $O(K^2)$.

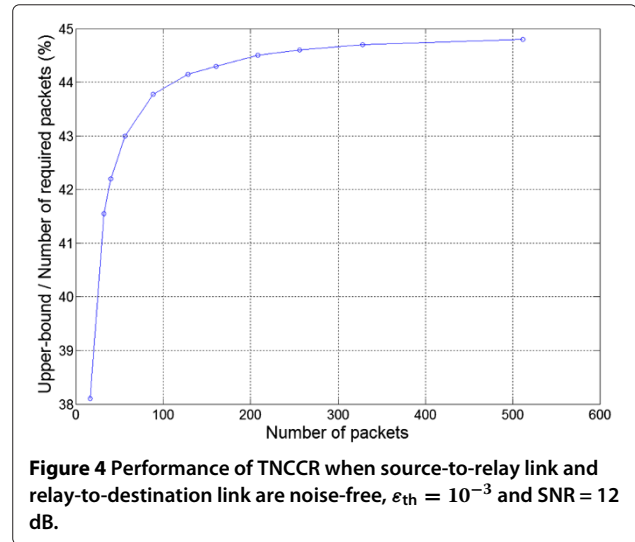
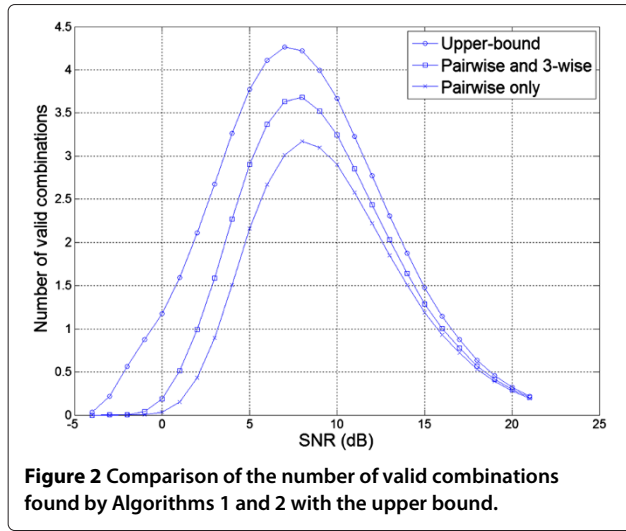
Figure 2 shows the number of valid combinations found by Algorithms 1 and 2, where $K = 16$. The upper bound on the number of valid combinations is also plotted. It can be seen that by executing Algorithm 1 or Algorithm 2, which are of very low complexity, near-upper bound performance can already be achieved.

Figure 3 displays the performance of TNCCR when the source-to-relay link and relay-to-destination link are noise-free. ε_{th} is set to be 10^{-3} . Again, TNCCR with Algorithm 2 attains most of the gains which resulted from network coding.

In Figure 4, the upper bound on reduction of number of retransmissions using TNCCR is plotted when ε_{th} is still 10^{-3} , and the transmission power over noise power is 12 dB. The upper bound on reduction of number of retransmissions can be calculated as $\frac{k - \text{Upperbound}_{\# \text{Valid Comb.}}}{k}$. It can be seen that the gain of TNCCR increases with the number of packets from the source.

5 Simulation results

In this section, we compare the performance of TCR and TNCCR by simulations. For TNCCR, Algorithm 2 is



applied to search for valid combinations. We simulate the retransmission performance in two scenarios:

Scenario 1 (Scn-1): The relay is positioned closer to the source, as displayed in Figure 5a.

Scenario 2 (Scn-2): The relay is positioned at the middle point between the source and the destination, as displayed in Figure 5b.

In the simulations, we take path loss into account. The path loss (in dB) can be calculated as $10\alpha \log_{10}(d)$, where α is the path loss exponent and is set to be 4, and d is the distance between the sender and the receiver. We set the distance between the source and the destination to be 1, and the relative distances of source-to-relay and relay-to-destination in the two scenarios are shown in Figure 5. In Scn-1, the average channel gain power of the source-to-relay channel is 20.915 dB larger than that of the

source-to-destination channel, and the average channel gain power of the relay-to-destination channel is 6.196 dB larger than that of the source-to-destination channel. In Scn-2, the average channel gain power of the source-to-relay channel and the relay-to-destination channel are both 12.04 dB larger than that of the source-to-destination channel.

Figure 6 shows the performance comparisons between TCR and TNCCR in the aforementioned scenarios, where the percentages of reduction on the number of retransmissions using different relay retransmission schemes compared to the CC-ARQ scheme without relay are plotted. ε_{th} is set to be 10^{-3} and 10^{-4} . Good threshold values for TCR and TNCCR in different situations, obtained through extensive simulations, are listed in Table 1. It can be seen that when the SNR is low, TCR and TNCCR have about the same performance. On the other hand, TNCCR has better performance when SNR is relatively high (7 dB or above). The implication of the observations is two-fold:

- (1) At low SNRs, where there are few valid combinations (as implied by Figure 2), there will be few retransmission of network-coded packets from the relay. In this case, TCR and TNCCR are more or less the same.
- (2) At high SNRs, the relay in TNCCR starts retransmitting more network-coded packets as there are more valid combinations. The results suggest that network-coded retransmissions are more efficient than non-network-coded retransmissions in the SNR region concerned.

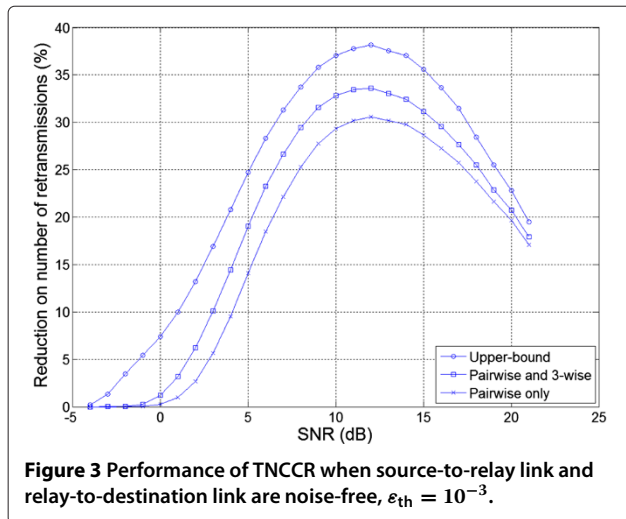


Figure 7 shows the performance of TNCCR scheme when the number of packets from the source varies. It

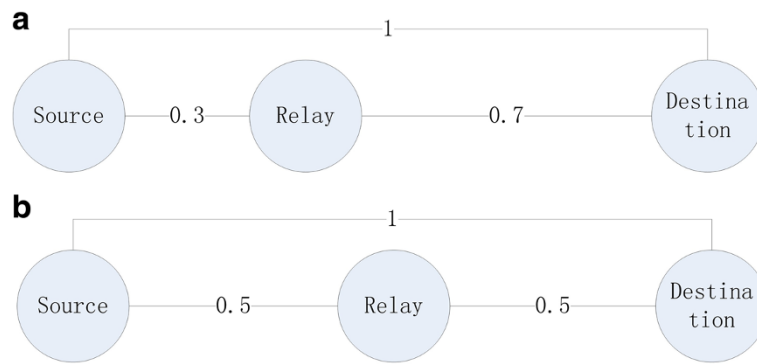


Figure 5 Two different positions of the relay studied in this work. (a) Scn-1, (b) Scn-2.

can be seen that TNCCR can save more retransmissions if more packets from the source are to be transmitted, which is similar to the observation of Figure 4.

Figure 8 shows the average packet decoding delay of TCR and TNCCR schemes in Scn-2, where ε_{th} is set

to be 10^{-3} . The results indicate that TNCCR scheme has less average packet decoding delay than the TCR scheme, which means that the destination experiences smaller delay in the retransmission phase before decoding a source packet using TNCCR.

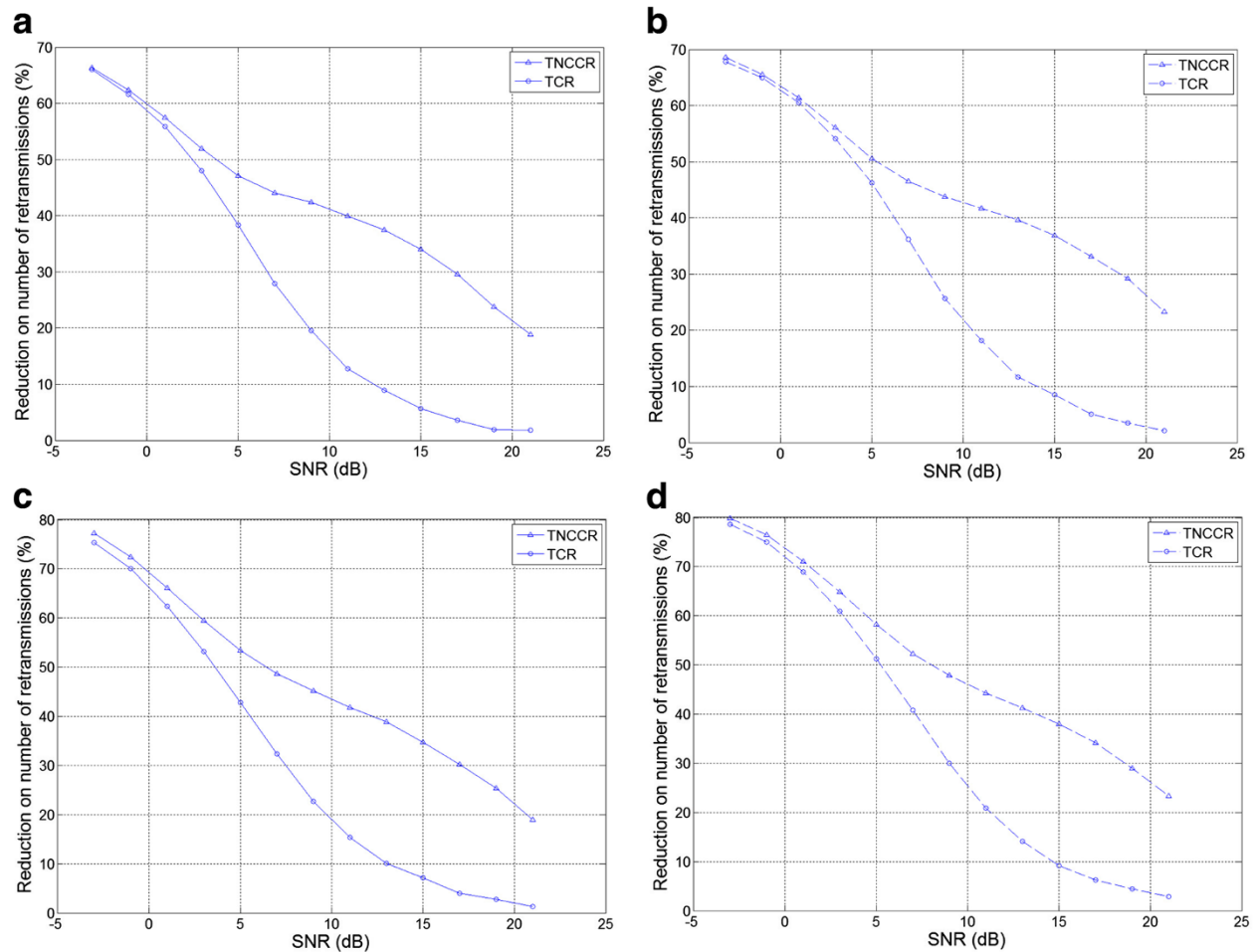


Figure 6 Comparison of MTCR against TCR and TNCCR in different scenarios. (a) $\varepsilon_{th} = 10^{-3}$ in Scn-1, (b) $\varepsilon_{th} = 10^{-4}$ in Scn-1, (c) $\varepsilon_{th} = 10^{-3}$ in Scn-2, and (d) $\varepsilon_{th} = 10^{-4}$ in Scn-2.

Table 1 Threshold values at relay of different retransmission schemes in different scenarios

Threshold value	Scenario	Retransmission scheme	Threshold at relay
$\varepsilon_{th} = 10^{-3}$	Scn-1	TCR	12.10
		TNCCR	11.15
	Scn-2	TCR	10.20
		TNCCR	10.35
$\varepsilon_{th} = 10^{-4}$	Scn-1	TCR	15.48
		TNCCR	15.08
	Scn-2	TCR	14.38
		TNCCR	13.98

6 Conclusions

In this paper, we addressed the retransmission issue in DmF relay channels. We first designed a retransmission scheme based on a weighted combiner. By forwarding high-quality observations, the relay was able to help the system achieve high retransmission efficiency in low SNR region. Then, on top of the weighted combiner, another retransmission scheme based on both weighted combiner and network coding was proposed. The network-coded packets were formed according to the qualities of the previously erroneously received packets at the destination, and they were able to boost the retransmission efficiency at high SNR region.

A possible future direction based on this work may consider a more sophisticated retransmission scheme that is based on incremental-redundancy hybrid automatic repeat request or IR-HARQ, i.e., by considering whether or not the qualities of the packets at the destination can help the relay to decide how parity bits should be generated from which source packets so that system throughput can be increased.

The proposed scheme also has the potential to be applied in wireless broadcast channels. As it has been

demonstrated in 3.2, a network-coded packet generated by considering the SNRs of the received packets at the destination can be used to decode two packets. If the SNRs of the received packets at the destination is not considered (i.e., treat packets in error as erasures), then a network-coded packet can only be used to decode one packet if it is instantly decodable [8]. It is therefore expected

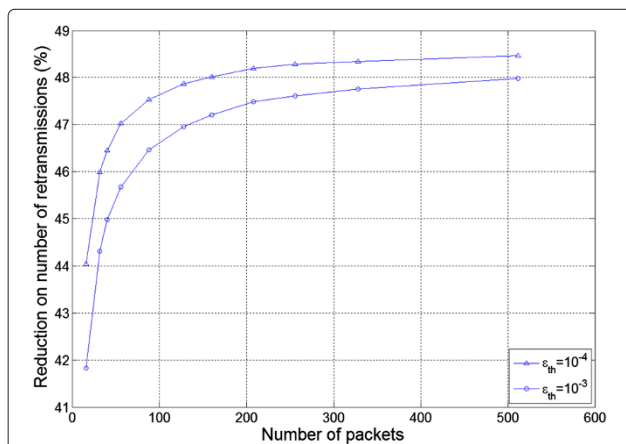


Figure 7 Reduction in the number of retransmissions using TNCCR in Scn-2.

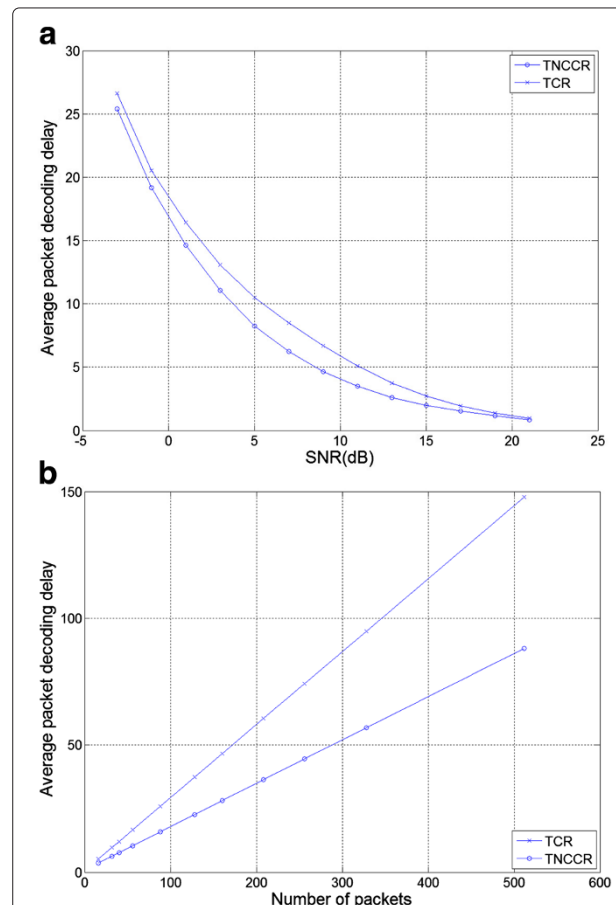


Figure 8 Average packet decoding delay in Scn-2 when $\varepsilon_{th} = 10^{-3}$. (a) Delay as a function of transmission power over noise power. (b) Delay as a function of the number of source packets, where the ratio between transmission power and noise power is 11 dB.

that the proposed network-coded retransmission scheme can result in higher throughput if the network-coded retransmissions are restricted to be instantly decodable. It is also interesting to investigate the possible performance improvement brought by the proposed scheme if a retransmitted packet is not required to be instantly decodable but innovative [9].

Appendices

Appendix 1

Extension to higher modulation schemes

We first demonstrate that network-coded retransmission can be applied when M -PSK modulation is used. We will see that the idea of network-coded retransmission presented in Section 3.2.1 is applicable when the signal constellation of the applied modulation scheme forms a group. From this observation, we can say that network-coded retransmission can also be applied to QAM, as the signals of QAM form a group [22].

M -PSK Consider the following signal constellation of M -PSK, where each point in the constellation is given by $e^{j2\pi\beta/M}$ (where $\beta \in [0, M-1]$ and $j^2 = -1$).

Assume that two signals (namely $X_1 = e^{j2\pi\beta_1/M}$ and $X_2 = e^{j2\pi\beta_2/M}$) are transmitted, where $\beta_1, \beta_2 \in [0, M-1]$. The two signals are corrupted with additive noise (unlike the other sections where fading is present in the channel, here we only assume there is additive noise for simplicity) and the received signals are $Y_1 = X_1 + n_1$ and $Y_2 = X_2 + n_2$, respectively. We now wish to send a network-coded signal that can provide information about X_1 and X_2 .

The M -PSK signal constellation can be related to the following finite Abelian group [23]:

$$U_M \triangleq \{e^{j2\pi\beta/M}, \beta \in [0, M-1]\}, \quad (21)$$

which is closed under the operation of complex multiplication. Let the inverse of X_1 and X_2 be $-X_1 = e^{j2\pi\beta'_1/M}$ and $-X_2 = e^{j2\pi\beta'_2/M}$, respectively, where $-a$ means the inverse of element a in a group and $\beta'_1, \beta'_2 \in [0, M-1]$. Denote \bullet as complex multiplication. If the signal $X_3 = (-X_1) \bullet (-X_2)$ is delivered to the receiver without noise, then at the receiver we can have

$$\begin{aligned} X_3 \bullet Y_1 &= e^{j2\pi\beta'_1/M} \bullet e^{j2\pi\beta'_2/M} \bullet (e^{j2\pi\beta_1/M} + z_1) \\ &= e^{j2\pi\beta'_2/M} + e^{j2\pi\beta'_1/M} \bullet e^{j2\pi\beta'_2/M} \bullet z_1 \\ &= (-X_2) + e^{j2\pi\beta'_1/M} \bullet e^{j2\pi\beta'_2/M} \bullet z_1. \end{aligned} \quad (22)$$

We can see from (22) that an extra observation of X_2 can be obtained from $X_3 \bullet Y_1$. Similarly, an extra observation of X_1 can be obtained from $X_3 \bullet Y_2$. The MRC principle can then be applied to combine multiple noisy observations.

QAM From the above discussion, it is evident that the network-coded retransmission scheme can be applied to a modulation scheme if the signals form a group. In [22], it has been proven that the signal constellation of 2^{2^n} -QAM is given by the group of units in the quotient ring $\mathbb{Z}[i]/2^n + 2^n i$, where n is a positive integer and $\mathbb{Z}[i]$ means Gaussian integers ring. Similar derivations to (22) can be used to show that the network-coded retransmission is applicable to QAM signals.

Appendix 2

Proof of proposition 1

Let the relay retransmit $\mathbf{P}'_1, \mathbf{P}'_2, \dots, \mathbf{P}'_x$ through an error-free channel to the destination, where upon receiving \mathbf{P}'_i , $i \in [1, x]$, l_i source packets can be decoded at the destination. Since we only consider pairwise and three-wise valid combinations, $l_i \in \{0, 1, 2, 3\}$.

Lemma 2. Let $\Gamma_{\alpha,SD}, \Gamma_{\beta,SD}$, and $\Gamma_{\gamma,SD}$ form a three-wise valid combination and, in order to decode source packets, $\mathbf{P}_\alpha, \mathbf{P}_\beta, \mathbf{P}_\gamma, \mathbf{P}'_i$ and \mathbf{P}'_r should be retransmitted, where $i, r \in [1, x]$. Then, a retransmission sequence which retransmits \mathbf{P}'_i followed by \mathbf{P}'_r has smaller average packet decoding delay than a retransmission sequence which retransmits other packets in between \mathbf{P}'_i and \mathbf{P}'_r .

Proof. Let $i, r, a_1, a_2, \dots, a_n \in [1, x]$. Consider the following three retransmission sequences:

- (1) $\mathbf{P}'_i, \mathbf{P}'_{a_1}, \mathbf{P}'_{a_2}, \dots, \mathbf{P}'_{a_n}, \mathbf{P}'_r$
- (2) $\mathbf{P}'_i, \mathbf{P}'_r, \mathbf{P}'_{a_1}, \mathbf{P}'_{a_2}, \dots, \mathbf{P}'_{a_n}$
- (3) $\mathbf{P}'_{a_1}, \mathbf{P}'_{a_2}, \dots, \mathbf{P}'_{a_n}, \mathbf{P}'_i, \mathbf{P}'_r$

Let the total packet decoding delay of sequences (1), (2), and (3) be d_1, d_2 , and d_3 , respectively. We have

$$\begin{aligned} d_1 &= 1 \times 0 + 2l_{a_1} + 3l_{a_2} + \dots + (a_n + 1)l_{a_n} + (a_n + 2)3, \\ d_2 &= 1 \times 0 + 2 \times 3 + 3l_{a_1} + 4l_{a_2} + \dots + (a_n + 2)l_{a_n} \\ &= 2 \times 3 + 2l_{a_1} + 3l_{a_2} + \dots + (a_n + 1)l_{a_n} + l_{a_1} \\ &\quad + l_{a_2} + \dots + l_{a_n}, \\ d_3 &= l_{a_1} + l_{a_2} + \dots + l_{a_n} + (a_n + 1) \times 0 + (a_n + 2)3. \end{aligned}$$

It is obvious that $d_3 \leq d_1$. To see $d_2 \leq d_1$, observe that

$$d_1 - d_2 = 3a_n - (l_{a_1} + l_{a_2} + \dots + l_{a_n}).$$

Since $l_i \in \{0, 1, 2, 3\}, \forall i \in \{a_1, a_2, \dots, a_n\}$, $l_{a_1} + l_{a_2} + \dots + l_{a_n} \leq 3a_n$. Therefore, $d_1 - d_2 \geq 0$.

As the number of source packets that can be decoded after receiving each of the three sequences is the same, sequence (1) gives the largest average packet decoding delay. \square

Lemma 3. Let $l_{a_1} = 2$ and $l_{a_2} = 1$, where $a_1, a_2 \in [1, x]$. Then, retransmitting \mathbf{P}'_{a_1} followed by \mathbf{P}'_{a_2} has smaller average packet decoding delay than retransmitting \mathbf{P}'_{a_2} followed by \mathbf{P}'_{a_1} .

Proof. The average packet decoding delay of retransmitting \mathbf{P}'_{a_1} followed by \mathbf{P}'_{a_2} can be calculated as $\frac{1 \times 2 + 2 \times 1}{2+1} = \frac{4}{3}$. The average packet decoding delay of retransmitting \mathbf{P}'_{a_2} followed by \mathbf{P}'_{a_1} can be calculated as $\frac{1 \times 1 + 2 \times 2}{1+2} = \frac{5}{3}$. The lemma then follows. \square

Lemma 4. Let $l_{a_1} = 2$, $l_{a_2} = 0$, and $l_{a_3} = 3$ so that three source packets can be decoded upon receiving \mathbf{P}'_{a_2} and \mathbf{P}'_{a_3} . Then, retransmitting \mathbf{P}'_{a_1} first gives smaller average packet decoding delay than retransmitting \mathbf{P}'_{a_2} and \mathbf{P}'_{a_3} first.

Proof. The average packet decoding delay of sending \mathbf{P}'_{a_1} first is $\frac{1 \times 2 + 2 \times 0 + 3 \times 3}{2+0+3} = 2.2$. The average packet decoding delay of sending \mathbf{P}'_{a_2} and \mathbf{P}'_{a_3} first is $\frac{1 \times 0 + 2 \times 3 + 3 \times 2}{0+3+2} = 2.4$. The lemma then follows. \square

Lemma 5. Let $l_{a_1} = 0$, $l_{a_2} = 3$, and $l_{a_3} = 1$ so that three source packets can be decoded upon receiving \mathbf{P}'_{a_1} and \mathbf{P}'_{a_2} . Then, retransmitting \mathbf{P}'_{a_1} and \mathbf{P}'_{a_2} first gives smaller average packet decoding delay than retransmitting \mathbf{P}'_{a_3} first.

Proof. The average packet decoding delay of sending \mathbf{P}'_{a_1} and \mathbf{P}'_{a_2} first is $\frac{1 \times 0 + 2 \times 3 + 3 \times 1}{0+3+1} = 2.25$. The average packet decoding delay of sending \mathbf{P}'_{a_3} first is $\frac{1 \times 1 + 2 \times 0 + 3 \times 3}{0+3+1} = 2.5$. The lemma then follows. \square

The above lemmas together suggest the following: To minimize the average packet decoding delay in error-free retransmissions, network-coded packets (each of which can be used to decode two source packets) should be retransmitted first, followed by pairs of network-coded packets (each of which can be used to decode three source packets), followed by source packets. Therefore, TNCCR gives the minimum average packet decoding delay when the retransmissions are error-free and for each $\Gamma_{i,SD} < T$, $\Gamma_{i,SR} \geq T_{R,NC}$, where $i \in [1, K]$. That proves Proposition 1.

Competing interests

Both authors declare that they have no competing interests.

Acknowledgements

This work was partially supported by a grant from the University Grants Committee of the Hong Kong Special Administrative Region, China (Project No.: AoE/E-02/08). Part of the work was presented in Wireless Advanced 2011.

Received: 4 January 2013 Accepted: 21 April 2013
Published: 25 May 2013

References

1. IF Akyildiz, DM Gutierrez-Estevez, EC Reyes, The evolution to 4G cellular systems: Lte-advanced. *Phys. Commun.* **3**(4), 217–244 (2010)
2. J Laneman, D Tse, G Wornell, Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Trans. Inf. Theory.* **50**(12), 3062–3080 (2004). doi:10.1109/TIT.2004.838089
3. TR Wang, A Cano, GB Giannakis, JN Laneman, High-performance cooperative demodulation with decode-and-forward relays. *IEEE Trans. Commun.* **55**(7), 1427–1438 (2007)
4. S Borade, L Zheng, R Gallager, Amplify-and-forward in wireless relay networks: rate, diversity, and network size. *IEEE Trans. Inf. Theory.* **53**(10), 3302–3318 (2007). doi:10.1109/TIT.2007.904774
5. R Ahlswede, N Cai, SYR Li, RW Yeung, Network information flow. *IEEE Trans. Inf. Theory.* **46**(4), 1204–1216 (2000)
6. D Nguyen, T Tran, T Nguyen, B Bose, Wireless broadcast using network coding. *IEEE Trans. Vehicular Technol.* **58**(2), 914–925 (2009). doi:10.1109/TVT.2008.927729
7. L Keller, E Drinea, C Fragouli, in *Fourth Workshop on NetCod*. Online Broadcasting with Network Coding, (2008), pp. 1–6. doi:10.1109/NETCOD.2008.4476183
8. P Sadeghi, R Shams, D Traskov, An optimal adaptive network coding scheme for minimizing decoding delay in broadcast erasure channels. *Eurasip J. Wireless Commun. Netw.* (2010). doi:10.1155/2010/618016
9. HY Kwan, K Shum, CW Sung, Generation of innovative and sparse encoding vectors for broadcast systems with feedback. *IEEE ISI Proc.* 1161–1165 (2011). doi:10.1109/ISIT.2011.6033715
10. QA Li, SH Ting, CK Ho, Nonlinear network code for high throughput broadcasting with retransmissions. *IEEE ISIT*, 2853–2857 (2009). doi:10.1109/ISIT.2009.5205268
11. C Zhi, C Wei, PY Fan, K Ben Letaief, in *IFIP International Conference on NPC*. Relay Aided Wireless Multicast Utilizing Network Coding: Outage Behaviour and Diversity Gain, (2008), pp. 358–364. doi:10.1109/NPC.2008.22
12. PY Fan, C Zhi, C Wei, K Ben Letaief, Reliable relay assisted wireless multicast using network coding. *IEEE J. Selected Areas Commun.* **27**(5), 749–762 (2009)
13. QT Vien, LN Tran, EK Hong, Network coding-based retransmission for relay aided multisource multicast networks. *Eurasip J. Wireless Commun. Netw.* **2011**, 643920 (2011)
14. TW Yune, D Kim, GH Im, Opportunistic network-coded cooperative transmission with demodulate-and-forward protocol in wireless channels. *IEEE Trans. Commun.* **59**(7), 1791–1795 (2011). doi:10.1109/TCOMM.2011.050911.090660
15. D Tse, P Viswanath, *Fundamentals of Wireless Communication*. (Cambridge University Press, New York, 2005)
16. D Chase, Code combining—a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets. *IEEE Trans. Commun.* **33**(5), 385–393 (1985). doi:10.1109/TCOM.1985.1096314
17. JG Proakis, *Digital Communications*, 3rd edn. (McGraw-Hill, New York, 1995)
18. A Bletsas, A Khisti, DP Reed, A Lippman, A simple cooperative diversity method based on network path selection. *IEEE J. Selected Areas Commun.* **24**(3), 659–672 (2006)
19. P Sadeghi, M Yu, Instantly decodable versus random linear network coding: a comparative framework for throughput and decoding delay performance. *CoRR* (2012). doi:arXiv:1208.2387v1
20. SF Assmann, DS Johnson, DJ Kleitman, JYT Leung, On a dual version of the one-dimensional bin packing problem. *J. Algorithms.* **5**(4), 502–525 (1984)
21. MR Garey, DS Johnson, *Computers and intractability: a guide to the theory of NP-completeness*. (W.H. Freeman, San Francisco, 1979)
22. J Rifa, Groups of complex integers used as qam signals. *IEEE Trans. Inf. Theory.* **41**(5), 1512–1517 (1995)
23. FR Kschischang, PG Debuda, S Pasupathy, Block coset codes for m-ary phase-shift keying. *IEEE J. Selected Areas Commun.* **7**(6), 900–913 (1989)

doi:10.1186/1687-1499-2013-136

Cite this article as: Liu and Sung: Network-coded retransmissions in wireless demodulate-and-forward relay channels. *EURASIP Journal on Wireless Communications and Networking* 2013 **2013**:136.