Design and Analysis of Relay-aided Broadcast using Binary Network Codes

Lu Lu, Ming Xiao, and Lars K. Rasmussen
Communication Theory Laboratory
KTH - Royal Institute of Technology - ACCESS Linnaeus Center, Sweden
Email: {lu.lu, ming.xiao, lars.rasmussen}@ee.kth.se

Abstract— **We consider a base-station broadcasting a set of order-insensitive packets to a user population over packet-erasure channels. To improve efficiency we propose relay-aided transmission using instantaneously-decodable binary network coding. The proposed coding schemes have the benefits of minimal decoding delay and low complexity. We further analyze the performance of the resulting broadcast schemes, and show that significant improvements in transmission efficiency are obtained as compared to previously proposed ARQ and network-coding-based schemes.**

Index Terms— **Broadcast, Instantaneously Decodable, Binary Network Coding, ARQ**

I. Introduction

In a wireless digital broadcast system, a base station (BS) forwards information packets to a user population over wireless channels. To meet reliability requirements error control coding is typically introduced at various layers in the protocol stack. At higher layers the corresponding BS-to-user links are typically modelled as packet-erasure channels where a received packet is either error-free or dropped as being in error.

Automatic repeat-request (ARQ) protocols [1] have traditionally been used for error-control in such broadcast systems. Though the protocol is straightforward, ARQ becomes inefficient in systems with many users. To increase the transmission efficiency, the use of network coding for retransmissions has been proposed [2]–[5]. In these schemes, packets lost by different users are jointly encoded with a suitable network code, leading to a reduction in the total number of transmitted packets required for retransmission.

The use of a relay can further improve the efficiency, as promised by fundamental results on the broadcast relay channel [6]–[8]. Due to practical constraints, it is considered a challenge to provide full duplex operation at the relay [9]. Likewise, if the BS and the relay transmit simultaneously multiple-access interference at the user nodes becomes another significant practical challenge [10]. It is therefore relevant to consider time-division transmissions where the relay and the BS transmit in different time slots, thus alleviating both problems.

In this work, we investigate relay-aided broadcast with network coding. We focus on order-insensitive packet delivery applications, and consider a class of instantaneously-decodable network coding (IDNC) schemes [2]–[5], [11] for retransmission. In such schemes, each network-coded retransmission packet contains at most one missing information packet for each intended receiver. As compared to random linear network coding [12], IDNC schemes enjoy some benefits but also suffer some drawbacks. For example, an IDNC scheme may not be throughput-optimal. This is in contrast to random linear network coding with sufficiently large alphabet size and packet length, where each coded packet is innovative with high probability. However, to ensure successful decoding in the random case, e.g., by matrix inversion, the users need to wait for a sufficient amount of coded packets to arrive, leading to a potential large decoding delay. Moreover, the matrix inversion at the user side is computational complex [2]. Conversely, when IDNC is used the decoding delay at the user side is minimal [2], [3], since a network-coded packet can be decoded immediately by the users. In addition IDNC can be easily implemented over GF(2), where only binary XOR operations is required in the encoding and decoding processes. It follows that the complexity is significantly reduced compared to codes operating over a larger field.

As our main contribution, we propose a new efficient instantaneously-decodable network coding scheme for the relay-aided broadcast system [13]. We further analyze the performance of the resulting broadcast scheme, and show that significant improvements in transmission efficiency is obtained as compared to previously proposed ARQ and network-coding-based schemes, e.g., [1], [11].

The remainder of the paper is organized as follows. In Section II, the system model is defined, and in Section III we describe retransmission schemes based on ARQ and IDNC, respectively. We then analyze the performance of the considered retransmission schemes in Section IV, while numerical results are presented in Section V.

II. System Model

We consider a BS with N information packets, $\mathbf{I}_1, \mathbf{I}_2, ..., \mathbf{I}_N$, to be broadcasted to M users with the

Manuscript received February 15, 2011; revised May 15, 2011; accepted September 30, 2011.
The research leading to these results has received funding from the European Research Council under the European Community Seventh Framework Programme (FP7/2007- 2013) / ERC grant agreement n° 228044. The work has further been supported in parts by VINNOVA under the "Joint Sweden-China Strategic Cooperation Program," the ARC Grant DP0986089, and the VR grants 621-2008-4349 and 621-2009-4666.

doi:10.4304/jcm.6.8.610-617

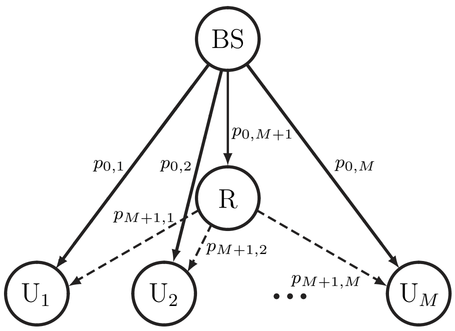

Figure 1. System model with one BS, one relay and M users.

assistance of a relay, as shown in Fig. 1. The BS is referred to as node 0, node $i = 1, 2, ..., M$ refers to user i, and node $M + 1$ represents the relay. The links between nodes are modelled as packet-erasure channels, where the erasure probability of the channel between node i and node j is denoted by $p_{i,j}$, with $i = 0, M + 1$, $j = 1, 2, ..., M + 1$, and $i \neq j$. We assume that the BS-to-relay channel is better than the BS-to-user channels, i.e., $p_{0,M+1} < p_{0,i}, i = 1, 2, ..., M$, and the relay-to-user channels are better than the BS-to-user channels, i.e., $p_{0,i} > p_{M+1,i}, i = 1, 2, ..., M$. As discussed in the introduction, we consider time-division transmissions; thus only one packet is transmitted per time slot either by the BS or by the relay.

To evaluate the status of received packets at each node, we let $\mathcal{H}_i, i = 0, 1, ..., M + 1$ denote the set of packets that have been received by the i-th node. Similarly, we let $\mathcal{L}_i, i = 0, 1, ..., M + 1$ denote the complement set, i.e., the set of packets which are yet to be received by the i-th node. Thus, if user i receives $\mathbf{I}_1, \mathbf{I}_3$ and $N = 5$, then $\mathcal{H}_i = \{1, 3\}$ and $\mathcal{L}_i = \{2, 4, 5\}$. In general we have $\mathcal{H}_i \bigcup \mathcal{L}_i = \Omega = \{1, 2, ..., N\}$, the set of all information packets, and $\mathcal{H}_i \bigcap \mathcal{L}_i = \emptyset$, the empty set.

When transmission starts, N information packets are at the BS; that is $H_0 = \Omega$ and $\mathcal{H}_i = \emptyset, i = 1, 2, ..., M+1$. In Phase 1 the BS broadcasts N packets to the relay and all the user nodes. Obviously due to the packet-erasure links, some packets are lost. Following Phase 1, each user feeds back the set of indices of lost packets to the BS and the relay. The relay also feeds back the relevant information to the BS. For simplicity, we assume instantaneous and error-free feedback channels. For retransmission, the BS and the relay evaluate $\mathcal{L}_i, i = 1, 2, ..., M$ based on the feedback received from the user nodes. The BS also keeps track of the packets lost by the relay, $\mathcal{L}_{M+1}$. The retransmission process is divided into two phases. In Phase 2 the BS conducts rounds of retransmissions, which are each followed by feedback updates from user nodes and the relay. This process is repeated until all packets are received by all users, or until all packets still missing at the user nodes are available at the relay. In Phase 3 the relay retransmits, again followed by feedback updates from the users nodes. As in Phase 2 this process is repeated until all user nodes have successfully received all information packets.

During the retransmission process, we exploit the benefits of the relay and of instantaneous-decodable network coding. Since the relay-to-user channels are better than the BS-to-user channels, the benefits of the relay are obvious. To illustrate the benefit of network coding, we provide an intuitive example here. Consider a two-user broadcast system without a relay and assume $N = 2$, $\mathcal{L}_1 = \{1\}$ and $\mathcal{L}_2 = \{2\}$. With ARQ, packets $\mathbf{I}_1$ and $\mathbf{I}_2$ are retransmitted separately, making a total of two retransmitted packets even if the retransmission process is error-free. Conversely, with network coding we can manage to retransmit only one network-coded packet $\mathbf{I}_1 \bigoplus \mathbf{I}_2$ to decrease the total number of retransmission. The first user can obtain $\mathbf{I}_1$ as $(\mathbf{I}_1 \bigoplus \mathbf{I}_2) \bigoplus \mathbf{I}_2$, while the second user can get $\mathbf{I}_2$ as $(\mathbf{I}_1 \bigoplus \mathbf{I}_2) \bigoplus \mathbf{I}_1$.

To enable instantaneous decodability we define an encoding rule as follows. Two information packets cannot be jointly encoded if they have both been simultaneously requested for retransmission by the same user. That is, $\mathbf{I}_i$ and $\mathbf{I}_j$ cannot be encoded jointly if $i \in \mathcal{L}_k$ and $j \in \mathcal{L}_k$, $i \neq j = 1, 2, ..., N, k = 1, 2, ..., M$. Instantaneous decodability obviously allows for fast decoding of individual packet, and since binary XOR operations suffice in the encoding and decoding processes, low complexity is ensured. Note however that we do not require instantaneous decodability at the relay node, since individual information packets are only required at the user nodes. If the relay receives $\mathbf{I}_{k_1} \bigoplus \mathbf{I}_{k_2} \bigoplus ... \bigoplus \mathbf{I}_{k_r}$ correctly, we say all the packets $\mathbf{I}_{k_1}$ to $\mathbf{I}_{k_r}$ are available at the relay. To facilitate analysis, we consider a network-coded packet to be *innovative* for user i if a new information packet can be obtain from the received coded packet.

A similar scheme was considered [11]; however, our proposed scheme is more efficient in the retransmission phases. Since packets are lost at the relay due to packet-erasures in the BS-relay channel, the relay node can also get innovative packets from the BS during the corresponding retransmission phase. Thus, in contrast to the protocol in [11], we allow the BS to retransmit first in Phase 2 before the relay gets to retransmit in Phase 3. Allowing for feedback updates during both Phase 2 and Phase 3 transmission efficiency is improved.

To measure the efficiency of the transmission schemes, we define the overhead η as follows,

$$\eta = \frac{X}{N}, \tag{1}$$

where X is the number of time slots used for transmission until all the users get all the N information packets. The efficiency increases as the overhead decreases.

III. Relay-aided Network Coding

In this part we describe the relay-aided ARQ protocol and our proposed relay-aided network coding protocol, where the objective of our scheme is to minimize the overhead η.

In the retransmission phase, the packets $\mathbf{I}_j, j \in \bigcup_{i \in \{1,2,...,M\}} \mathcal{L}_i$, are requested by the set of users, and will be retransmitted by the BS or the relay node. The lost packets can be divided into two subsets conditioned on whether the packet is received by the relay node or not. The first subset contains the packets $\mathbf{I}_j, j \in (\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i) \bigcap \mathcal{H}_{M+1}$, which have been received correctly by the relay node. The remaining subset of packets $\mathbf{I}_j, j \in (\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i) \bigcap \mathcal{L}_{M+1}$, are only available at the BS, and thus can only be retransmitted by the BS.

A. Relay-aided ARQ

In Phase 2, the BS retransmits the information packets in the second subset until $(\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i) \bigcap \mathcal{H}_{M+1} = \emptyset$. Then, in Phase 3, the relay retransmits the remaining packets until they are all successfully received by all the users.

B. Relay-aided Network Coding

In contrast to the ARQ protocol the retransmitted packets in our proposed scheme are not the originally lost information packets, but instead appropriately network-encoded packets. As for the ARQ protocol, the lost packets in the second subset are retransmitted in Phase 2, which ends when $(\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i) \bigcap \mathcal{H}_{M+1} = \emptyset$. In Phase 3, the relay node assists to retransmit the remaining lost packets using network coding to complete the transmission.

The algorithm for Phase 2 is outlined in Algorithm 1. In order to increase efficiency, the algorithm seeks to maximize the number of users capable of receiving an innovative network-coded packet in step b. After Phase 2 all the packets still lost by the users can be retransmitted by the relay, since $(\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i) \bigcap \mathcal{L}_{M+1} = \emptyset$. Note that we do not require that the network-coded packets are instantaneously decodable at the relay node. To illustrate this observation we use a simple example. Let $M = 2$, and $\mathcal{L}_1 = \{1\}, \mathcal{L}_2 = \{2\}, \mathcal{L}_3 = \{1, 2\}$ following Phase 1. The BS then retransmits $\mathbf{I}_1 \bigoplus \mathbf{I}_2$, which is innovative for both users. If the relay successfully receives the coded packet in Phase 2, while one or both users do not, then the relay can retransmit the received coded packet as is, $\mathbf{I}_1 \bigoplus \mathbf{I}_2$, without attempting decoding first. However, if we require that the packet is also instantaneously decodable for the relay, then the BS can only retransmit $\mathbf{I}_1$ and $\mathbf{I}_2$ separately since the relay has lost both packets in Phase 1. Obviously, this will affect the efficiency.

In Phase 3, the system is the same as a one-source broadcast system without a relay, as studied in [2]–[5] for binary network coding schemes. We therefore modify the algorithm proposed in [5] by considering that the relay may use network-coded packet from the BS directly in the relay encoding process. For example, assume the relay node has received $\mathbf{I}_j \bigoplus \mathbf{I}_k$, but has not previously been able to decode $\mathbf{I}_j$, $\mathbf{I}_k$. Now further assume that the network-coded packet required based on user feedback has to be in a form of $\mathbf{I}_j \bigoplus \mathbf{I}_k \bigoplus \mathbf{I}_{m_1} \bigoplus ... \bigoplus \mathbf{I}_{m_t}$, in which $m_1 \neq ... \neq m_t \neq j \neq k$, unless $\mathbf{I}_j \notin (\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i)$ or $\mathbf{I}_k \notin (\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i)$. To keep the instantaneous decodability property of the users in Phase 3, the relay simply treats its undecodable packet from the BS as one source packet for encoding. After Phase 2, we assume there are n_r packets at the relay node, including the received uncoded packets in Phase 1 and coded packets in Phase 2. They are denoted as $\mathbf{J}_i, (i = 1, 2, ..., n_r)$. We define $\mathcal{N}_i$ to record the indices of the information packets which are contained in $\mathbf{J}_i$. If $\mathbf{J}_i$ is a uncoded packet $\mathbf{I}_j$, $\mathcal{N}_i = \{j\}$. If $\mathbf{J}_i$ is a coded packet, say $\mathbf{I}_{k_1} \bigoplus \mathbf{I}_{k_2} \bigoplus ... \bigoplus \mathbf{I}_{k_r}$, $\mathcal{N}_i = \{k_1, ..., k_r\}$. The algorithm for Phase 3 is outlined in Algorithm 2.

Algorithm 1 Retransmission algorithm for Phase 2

a: *Initialization*:
- Let $\mathcal{R} = (\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i) \bigcap \mathcal{L}_{M+1}$.
- Define a set $\mathcal{C} = \emptyset$ to record the indices of packets to be jointly encoded and transmitted.

b: *Network-coded packet formulation*:
- Define $\mathcal{Q}_j, j = 1, 2, ..., N$ to record the indices of the users that did not receive information packet $\mathbf{I}_j$.
- Determine the packet index $k \in \mathcal{R}$ for which,
$$k = \arg\max_{j\in\mathcal{R}} |\mathcal{Q}_j|, \quad (2)$$
where $|\mathcal{Q}_j|$ denotes the cardinality of $\mathcal{Q}_j$.
- Let $\mathcal{C} = \mathcal{C} \bigcup \{k\}$, and determine the set $\mathcal{O} \subset \mathcal{R}$ that does not violate the encoding principle based on $\mathcal{C}$.
- Update $\mathcal{R} = \mathcal{O}$.
- Repeat step **b** until $\mathcal{R} = \emptyset$.

c: *Retransmissions*:
- Encode (binary add) information packets whose indices are in $\mathcal{C}$. If $\mathcal{C} = \{1, 2\}$, the encoded packet is $\mathbf{I}_1 \bigoplus \mathbf{I}_2$.
- Retransmit the resulting encoded packet.

d: *Status updates*:
- If the relay node receives the network-coded packet then update the relay status as $\mathcal{L}_{M+1} := \mathcal{L}_{M+1} \backslash \mathcal{C}$.
- If user i receives the coded packet and it is innovative for user i, one new information packet, say $\mathbf{I}_t$, can be retrieved as an innovation. Then the status for user i is updated as $\mathcal{L}_i := \mathcal{L}_i \setminus \{t\}$. Update $\mathcal{L}_i$ for all the users, respectively.
- Let $\mathcal{R} = (\bigcup_{i\in\{1,2,...,M\}} \mathcal{L}_i) \bigcap \mathcal{L}_{M+1}$, and $\mathcal{C} = \emptyset$.
- Repeat step **b** until $\mathcal{R} = \emptyset$.

IV. Performance Analysis

A. Relay-Aided ARQ

For the ARQ case the transmission process of each packet is independent and has the same statistical characteristics. Thus, to evaluate the performance, we only need to consider the transmission of one arbitrary packet.

Algorithm 2 Retransmission algorithm for Phase 3

a: *Initialization*:

- Let $\mathcal{R} = \{i|\mathcal{N}_i \bigcap (\bigcup_{k=\{1,...,M\}} \mathcal{L}_k) \neq \emptyset\}$.
- Define a set $\mathcal{C} = \emptyset$ to record the indices of packets to be jointly encoded and transmitted.

b: *Network-coded packet formulation*:

- Let n_i record the number of users that have lost one or more information packets in $\mathcal{N}_i$,

$$n_i = \sum_{k=1}^{M} f\left(\mathcal{N}_i \bigcap \mathcal{L}_k\right) \tag{3}$$

where

$$f(x) = \begin{cases} 1 & x \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

- Determine the packet index $i^* \in \mathcal{R}$ for which,

$$i^* = \arg\max_{i \in \mathcal{R}} n_i. \tag{5}$$

- Let $\mathcal{C} = \mathcal{C} \bigcup \mathcal{N}_{i^*}$, and determine the set $\mathcal{O} \subset \mathcal{R}$ that does not violate the encoding principle based on $\mathcal{C}$.
- Update $\mathcal{R} = \mathcal{O}$.
- Repeat step **b** until $\mathcal{R} = \emptyset$.

c: *Retransmissions*:

- Encode (binary add) information packets whose indices are in $\mathcal{C}$. If $\mathcal{C} = \{1, 2\}$, the encoded packet is $\mathbf{I}_1 \bigoplus \mathbf{I}_2$.
- Retransmit the resulting encoded packet.

d: *Status updates*:

- If user i receives the coded packet and it is innovative for user i, one new information packet, say $\mathbf{I}_t$, can be retrieved as an innovation. Then the status for user i is updated as $\mathcal{L}_i := \mathcal{L}_i \setminus \{t\}$. Update $\mathcal{L}_i$ for all the users, respectively.
- If $\mathbf{I}_j$ has been received correctly by all the users, remove j from $\mathcal{N}_i, (i = 1, 2, ..., n_r)$.
- Let $\mathcal{R} = \{i|\mathcal{N}_i \bigcap (\bigcup_{k=\{1,...,M\}} \mathcal{L}_k) \neq \emptyset\}$ and $\mathcal{C} = \emptyset$.
- Repeat from step **b** until $\mathcal{R} = \emptyset$.

For the performance analysis we introduce an absorbing Markov chain to describe the transmission process.

We define the state of the broadcast system as a vector of length $M + 1$. If node i receives the packet correctly, the element i is set to one. Otherwise, it is set to zero. Thus, there are $2^{(M+1)}$ possible states, where state j is expressed as,

$$S^j = \left[s_1^j, s_2^j, ..., s_{M+1}^j\right]. \tag{6}$$

We then define the probability transition matrix $\mathbf{Q}$ of dimensions $2^{(M+1)} \times 2^{(M+1)}$. The state transition probability from state S^i to S^j is denoted by $q_{i,j}$, which can be computed based on the corresponding channel erasure probabilities. If $s_{M+1}^i = 1$, the relay retransmits the packet. We have

$$q_{i,j} = \begin{cases} 0 & \exists s_k^j < s_k^i \\ \prod_{k=1}^{M} p_{M+1,k}^{e_{k,0}} (1 - p_{M+1,k})^{e_{k,1}} & \text{otherwise} \end{cases} \tag{7}$$

where $e_{k,0} = I(s_k^i)[1 - (s_k^j - s_k^i)]$ and $e_{k,1} = I(s_k^i)(s_k^j - s_k^i)$, $k = 1, 2, ..., M+1$, and $I(\cdot)$ is the indicator function defined as,

$$I(s_k^i) = \begin{cases} 1 & \text{if } s_k^i \neq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

If instead $s_{M+1}^i = 0$, then the BS retransmits the packet. In this case,

$$q_{i,j} = \begin{cases} 0 & \exists s_k^j < s_k^i \\ \prod_{k=1}^{M+1} p_{0,k}^{e_{k,0}} (1 - p_{0,k})^{e_{k,1}} & \text{otherwise} \end{cases} \tag{9}$$

where $k = 1, 2, ..., M + 1$. Note that $q_{i,j}$ only depends on the current state, but not the previous states. Thus, the entire transmission can be modelled as a Markov chain. Moreover, if the system enters the state S^j, where $s_k^j = 1$ for $k = 1, 2, ..., M$, the transmission terminates and the system is unable to leave the state. There are $r = 2$ such states which are referred to as absorbing states. $t = 2^{(M+1)} - 2$ states are referred as transient states [14]. Thus, the transmission is modelled as an absorbing Markov chain.

We then use the probability transition matrix $\mathbf{Q}$ to analyze this absorbing Markov chain. To easy notation, we first reorder the states according to transient states and absorbing states. The transition matrix have the following canonical form,

$$\mathbf{Q} = \begin{array}{c} \\ \text{transient} \\ \text{absorbing} \end{array} \begin{array}{c} \begin{array}{cc} \text{transient} & \text{absorbing} \end{array} \\ \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \end{array} \tag{10}$$

where $\mathbf{I}$ is an r-by-r identity matrix, $\mathbf{0}$ is an r-by-t zero matrix, $\mathbf{Q}_1$ is a t-by-t matrix and $\mathbf{Q}_2$ is a nonzero t-by-r matrix. The entries of the submatrix $\mathbf{Q}_1$ are the probabilities for being in each of the transient states after one transmission for each possible current transient state. Let m_j be the expected number of transitions before the chain is absorbed, given that the chain starts in state S^j, and let $\mathbf{m} = [m_1, m_2, ..., m_{(N+1)^{M+1}}]^{\mathrm{T}}$. Then,

$$\mathbf{m} = (\mathbf{I} - \mathbf{Q_1})^{-1}\mathbf{c}, \tag{11}$$

where $\mathbf{c}$ is a column vector with all-one entries (theorem 11.4 and 11.5 in [14]). Our system starts with the all-zero state, and we can get the expected value of η directly from (11). For example, if $S^1 = [0, 0, ..., 0]$, then, $\eta = m_1$ where m_1 is the first element of the vector $\mathbf{m}$.

B. Relay-Aided Network Coding

For our proposed scheme the analysis is based on a probabilistic approach. Without loss of generality, we assume that $p_{0,1} \geq p_{0,2} \geq ... \geq p_{0,M} \geq p_{0,M+1}$. We first derive the expected number of transmissions for a one-hop broadcast system; a result we will use for the analysis of Phase 3. When there is no relay, the retransmission algorithm is based on Algorithm 2, where $\mathcal{N}_i = \{i\}$ for $i = 1, 2, ..., N$. We begin with the case of $M = 2$ before generalizing the results to arbitrary M. According to the algorithm, $\mathbf{I}_j$, $j \in \mathcal{L}_1 \bigcap \mathcal{L}_2$ will be retransmitted without coding. If $\mathbf{I}_j$, $j \in \mathcal{L}_1 \bigcap \mathcal{H}_2$, the BS will locate a packet $\mathbf{I}_m$, where $m \in \mathcal{H}_1 \bigcap \mathcal{L}_2$ and retransmit the coded packet $\mathbf{I}_j \bigoplus \mathbf{I}_m$. If $\mathcal{H}_1 \bigcap \mathcal{L}_2 = \emptyset$, $\mathbf{I}_j$, $j \in \mathcal{L}_1 \bigcap \mathcal{H}_2$, will be retransmitted directly. For packet $\mathbf{I}_j$, $j \in \mathcal{H}_1 \bigcap \mathcal{L}_2$ the BS will locate a packet $\mathbf{I}_m$, where $m \in \mathcal{H}_1 \bigcap \mathcal{L}_2$ and retransmit the coded packet $\mathbf{I}_j \bigoplus \mathbf{I}_m$. If $\mathcal{L}_1 \bigcap \mathcal{H}_2 = \emptyset$, $\mathbf{I}_j$, $j \in \mathcal{H}_1 \bigcap \mathcal{L}_2$, will be retransmitted directly. Since $p_{0,1} \geq p_{0,2}$, then $E[|\mathcal{L}_1 \bigcap \mathcal{H}_2|] \geq E[|\mathcal{H}_1 \bigcap \mathcal{L}_2|]$ where $E[|\mathcal{L}|]$ denotes the expected cardinality of the set $\mathcal{L}$. It follows that, on average, every retransmitted packet is innovative for user 1. Thus, the number of retransmissions equals the number of transmissions required to complete the transmission for user 1. At the end of Phase 1, we have $E[|\mathcal{L}_1|] = Np_{0,1}$. Since the erasure probability of the channel is still $p_{0,1}$ during the retransmissions, the expected number of retransmissions is

$$E[N_r] = Np_{0,1} + Np_{0,1}^2 + ... + Np_{0,1}^{\infty} = \frac{Np_{0,1}}{1 - p_{0,1}}. \quad (12)$$

Since this result is based on every retransmitted packet being innovative to user 1, it is a lower bound on the number of retransmissions.

We then extend the results to the general case of $M \geq 2$ by induction, showing that all retransmitted packets are innovative to user 1. We have already shown that the statement is true when $M = 2$. We now assume that the statement is true for $M = k - 1$, and show that it is also true for $M = k$. Consider the first $k - 1$ users out of k. From our induction assumption, all the generated encoded packets are innovative for user 1. At the kth user, the expected number of lost packets is $Np_{0,k}$. With the assumption of independent erasure channels, the number of lost packets at the kth users which are also lost at one or more other users is $T_1 = Np_{0,k} - Np_{0,k} \prod_{i=1}^{k-1}(1 - p_{0,i})$. These packets are already processed into encoded packets. The remaining $T_2 = Np_{0,k} \prod_{i=1}^{k-1}(1 - p_{0,i})$ packets are only lost by the kth user. These T_2 packets can only be encoded with $T_3 = Np_{0,1} - T_1$ packets without violating the encoding principle. Since $T_3 - T_2 = Np_{0,1} - N_{0,k} \geq 0$, the encoded packets are always innovative for user 1. Thus, the average number of retransmissions is still the same as before (12), which is also a lower bound for the number of retransmissions.

Next, we analyze the performance for the relay-aided system, where the results for the single-hop broadcast system will be used in the analysis of Phase 3. We again begin with the case of $M = 2$ before generalizing the results to arbitrary $M \geq 2$. At the end of Phase 1 we have $E[|\mathcal{L}_1 \bigcap \mathcal{L}_3|] = Np_{0,1}p_{0,3} \geq Np_{0,2}p_{0,3} = E[|\mathcal{L}_2 \bigcap \mathcal{L}_3|]$. So the expected number of packets lost by both user 1 and the relay are greater than the expected number of packets lost by both user 2 and the relay. In Phase 2, the packets $\mathbf{I}_{x_1}, x_1 \in \mathcal{L}_1 \bigcap \mathcal{L}_2 \bigcap \mathcal{L}_3$ will be retransmitted first, and then, network encoded packets $\mathbf{I}_{x_1} \bigoplus \mathbf{I}_{x_2}, x_1 \in \mathcal{L}_1 \bigcap \mathcal{H}_2 \bigcap \mathcal{L}_3, x_2 \in \mathcal{H}_1 \bigcap \mathcal{L}_2 \bigcap \mathcal{L}_3$ will be transmitted. As before, since $p_{0,1} \geq p_{0,2}$ we always have $E[|\mathcal{L}_1 \bigcap \mathcal{L}_3|] \geq E[|\mathcal{L}_2 \bigcap \mathcal{L}_3|]$. So assuming average behavior, the packets $\mathbf{I}_{x_1}, x_1 \in \mathcal{L}_1 \bigcap \mathcal{H}_2 \bigcap \mathcal{L}_3$ will conclude the retransmissions in Phase 2. We observe that in the average sense only a subset of the transmitted packets in Phase 2 are innovative for user 2; however, all the transmitted packets are innovative for user 1. Thus, the transmissions in Phase 2 are divided into two sub-phases: 2.(A), where the transmitted packets are innovative for both the users; and 2.(B), where the transmitted packets are innovative only to user 1.

In Phase 2.(A), the transmitted packets are based on network-encoded packets from the set $\mathbf{I}_j, j \in J_{2.(A)} = \{j | j \in \mathcal{L}_2 \bigcap \mathcal{L}_3\}$. The transmission of one such packet is considered successful if the relay or user 2 receives it correctly. For an arbitrary lost packet in this phase, the probability that the number of transmissions required for is successful reception is equal to k, $T_{2.(A)} = k$ is

$$\begin{aligned} P\{T_{2.(A)} = k\} &= p_{0,3}^{k-1}(1 - p_{0,3})p_{0,2}^{k} + p_{0,3}^{k}p_{0,2}^{k-1}(1 - p_{0,2}) \\ &\quad + p_{0,3}^{k-1}(1 - p_{0,3})p_{0,2}^{k-1}(1 - p_{0,2}) \\ &= p_{0,3}^{k-1}p_{0,2}^{k-1}(1 - p_{0,3}p_{0,2}). \end{aligned} \quad (13)$$

Thus, the average number of transmission packets in this phase for one outstanding packet to be successfully received is,

$$E[T_{2.(A)}] = \sum_{k=1}^{\infty} kP\{T_{2.1} = k\} = \frac{1}{1 - p_{0,3}p_{0,2}} (14)$$

It follows that the average number of transmissions in Phase 2.(A) is,

$$E[N_{2.(A)}] = N\frac{p_{0,3}p_{0,2}}{1 - p_{0,3}p_{0,2}}. \quad (15)$$

Similarly, we can get the average number of transmissions for user 1 in Phase 2, denoted as N_2. Since the packets $J_2 = \{j | j \in \mathcal{L}_1 \bigcap \mathcal{L}_3\}$ are transmitted during the entire Phase 2, it is also the average number of transmissions. We have,

$$E[N_2] = N\frac{p_{0,3}p_{0,1}}{1 - p_{0,3}p_{0,1}}. \quad (16)$$

In order to determine the expected number of transmissions in Phase 3, we need to know $E[|\mathcal{L}_i|], i = 1, 2$ at the completion of Phase 2. We already know that $E[|\mathcal{L}_i|] = Np_{0,i}, i = 1, 2$ at the beginning of Phase 2. Thus we only need to determine the expected number of packets received in Phase 2 that are innovative for each user. Since the transmission of one packet in Phase 2 is successful if either the relay or a user receives it , the ratios of packets received by respective users are

$$f_{2.(A)} = \frac{P\{\text{node 2 receives the packet when } T_{2.(A)} = k\}}{P\{T_{2.(A)} = k\}} = \frac{p_{0,3}^k p_{0,2}^{k-1}(1-p_{0,2}) + p_{0,3}^{k-1}(1-p_{0,3})p_{0,2}^{k-1}(1-p_{0,2})}{p_{0,3}^{k-1}p_{0,2}^{k-1}(1-p_{0,3}p_{0,2})} = \frac{1-p_{0,2}}{1-p_{0,3}p_{0,2}}, \quad (17)$$

and,

$$f_2 = \frac{P\{\text{node 1 receives the packet when } T_2 = k\}}{P\{T_2 = k\}} = \frac{p_{0,3}^k p_{0,1}^{k-1}(1-p_{0,1}) + p_{0,3}^{k-1}(1-p_{0,3})p_{0,1}^{k-1}(1-p_{0,1})}{p_{0,3}^{k-1}p_{0,1}^{k-1}(1-p_{0,3}p_{0,1})} = \frac{1-p_{0,1}}{1-p_{0,3}p_{0,1}}. \quad (18)$$

Based on these results, the expected cardinality of $\mathcal{L}_i$ at the end of Phase 2 (p2) is,

$$E[|\mathcal{L}_1||\text{p2 ends}] = N - N(1-p_{0,1}) - \frac{Np_{0,1}p_{0,3}(1-p_{0,1})}{1-p_{0,1}p_{0,3}} = Np_{0,1}\frac{1-p_{0,3}}{1-p_{0,1}p_{0,3}}, \quad (19)$$

$$E[|\mathcal{L}_2||\text{p2 ends}] = N - N(1-p_{0,2}) - \frac{Np_{0,2}p_{0,3}(1-p_{0,2})}{1-p_{0,2}p_{0,3}} = Np_{0,2}\frac{1-p_{0,3}}{1-p_{0,2}p_{0,3}}. \quad (20)$$

The transmissions in Phase 3 are similar to a one-hop broadcast system with network coding. At the beginning of Phase 3, the relay nodes have some original information packets and some network-coded packets which cannot be decoded by the relay. There are two different cases for the original information packets: (1) $\mathbf{I}_k$, $k \in \mathcal{L}_1 \bigcap \mathcal{L}_2$, which is lost by both users; (2) $\mathbf{I}_k$, $k \in \mathcal{L}_1 \bigcap \mathcal{H}_2$ or $k \in \mathcal{H}_1 \bigcap \mathcal{L}_2$, which is lost by one of the users. For the un-decoded network-encoded packets, there are two different cases as well, denoted as cases (3) and (4), respectively. In case (3) each user requires one of the respective packets. In case (4), only one user requires one of the packets. According to our transmission scheme, a packet belonging to cases (1) or (3) (innovative for both users) will be retransmitted directly until any, or both, of the users successfully receive it. If both users are successful the packet is no longer lost. If only one user is successful then the packet is now in one of the cases (2) or (4). For a packet in case (2), if $\mathbf{I}_k$, $k \in \mathcal{L}_1 \bigcap \mathcal{H}_2$, the relay will locate one packet $\mathbf{I}_m$, $m \in \mathcal{H}_1 \bigcap \mathcal{L}_2$ or $\mathbf{I}_m \bigoplus \mathbf{I}_t$, $m \in \mathcal{H}_1 \bigcap \mathcal{L}_2, t \in \mathcal{H}_1 \bigcap \mathcal{H}_2$ and generate $\mathbf{I}_k \bigoplus \mathbf{I}_m$ or $\mathbf{I}_k \bigoplus \mathbf{I}_m \bigoplus \mathbf{I}_t$, respectively. A similar procedure is executed for a packet in case (4). Unless $\mathcal{L}_i = \emptyset$, $\forall i \in \{1, 2\}$, our scheme can always generate an innovative packet for both users. Based on the analysis, we can conclude that the transmissions in Phase 3 are dominated by the user that has more transmissions to complete. Thus, the expected number of transmissions in Phase 3 is,

$$E[N_3] = \max\left(\frac{E[|\mathcal{L}_1||\text{p2 ends}]}{1-p_{3,1}}, \frac{E[|\mathcal{L}_2||\text{p2 ends}]}{1-p_{3,2}}\right). \quad (21)$$

Thus, for $M = 2$ the expected efficiency for the entire transmission period is,

$$\eta_2 = \frac{N + E[N_2] + E[N_3]}{N}. \quad (22)$$

It is challenging to derive a closed-form solution for the general case of $M \geq 2$. Instead we follow a similar approach as for $M = 2$ and the results for the one-hop system to derive a lower bound. We conclude that the transmissions in Phase 2 is dominated by the first user, which has the highest erasure probability. We can therefore obtain a lower bound for Phase 2 under the assumption that every packet is innovative for user 1, and thus,

$$E_L[N_2] = \frac{Np_{0,1}p_{0,M+1}}{1-p_{0,1}p_{0,M+1}}. \quad (23)$$

At the end of this phase, we have

$$E_L[|\mathcal{L}_i||\text{p2 ends}] = Np_{0,i}\frac{1-p_{0,M+1}}{1-p_{0,i}p_{0,M+1}}. \quad (24)$$

Then, based on the results for the system without a relay, the lower bound for Phase 3 is,

$$E_L[N_3] = \max_i\left(\frac{E_L[|\mathcal{L}_i||\text{p2 ends}]}{1-p_{M+1,i}}\right). \quad (25)$$

Using (23), (24) and (25), we have

$$\eta_M \geq \frac{N + E_L[N_2] + E_L[N_3]}{N} = \frac{1}{1-p_{0,1}p_{0,M+1}} + \max_i\left(\frac{p_{0,i}\frac{1-p_{0,M+1}}{1-p_{0,i}p_{0,M+1}}}{1-p_{M+1,i}}\right) \quad (26)$$

V. Numerical results

The performance of the proposed schemes is contrasted against the performance of ARQ and the network coding scheme in [11]. The simulation results are based on 50000 runs of system realizations.

In Fig. 2, we show the overhead η_M as a function of the number of users M for the case of $N = 1000$, $p_{0,1} = 0.3$, $p_{0,i} = 0.2$, ($i = 2, 3, ..., M$), $p_{0,M+1} = 0.1$, $p_{M+1,1} = 0.2$ and $p_{M+1,i} = 0.1$ where $i = 2, 3, ..., M$. We observe that the efficiency of our proposed scheme is not affected by increasing M and approaches the lower bound in (26). This confirms our analytical results since the lower bound of the general case does not depend on M. We obtain similar results for other system parameters. In contrast the performance of the scheme in [11] degrades with increasing M. Thus, the performance improvements of our proposed scheme over the scheme in [11] grows with increasing M.

Since our proposed algorithm shows minimal dependence of M we will consider a system with $M = 2$

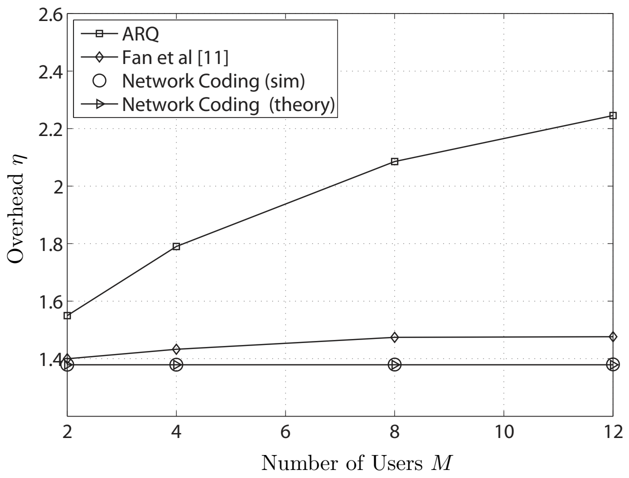

Figure 2. Overhead versus the number of users M, where $N = 1000$, $p_{0,1} = 0.3$, $p_{0,i} = 0.2$, $(i = 2, 3, ..., M)$, $p_{0,M+1} = 0.1$, $p_{M+1,1} = 0.2$ and $p_{M+1,i} = 0.1$ where $i = 2, 3, ..., M$.

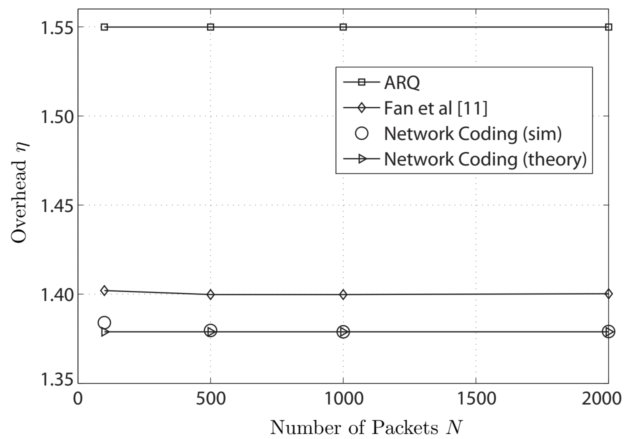

Figure 3. Overhead versus the number of information packets N, where $M = 2$, $p_{0,1} = 0.3$, $p_{0,2} = 0.2$, $p_{0,3} = 0.1$, $p_{3,1} = 0.2$ and $p_{3,2} = 0.1$.

in the following to limited the number of independent variables, and thus more clearly demonstrate the impact of other system parameters. Moreover, a lower bound on the potential performance gains over the scheme in [11] can be provided by the $M = 2$ system.

In Fig. 3 the overhead is shown as a function of the number, N, of information packets to be broadcast for the case of $M = 2$, $p_{0,1} = 0.3$, $p_{0,2} = 0.2$, $p_{0,3} = 0.1$, $p_{3,1} = 0.2$ and $p_{3,2} = 0.1$. We observe that the performances of all the schemes are virtually unaffected by N. For the proposed scheme, we note a minor gap between the analytical overhead and the simulated performance when N is small, i.e. $N < 300$. With increasing N, the discrepancy vanishes. We further observe that the networked-coded schemes are far superior to a traditional ARQ strategy. The modifications of our proposed scheme as compared to the scheme in [11] leads to the observed improvements in efficiency.

In Fig. 4, the overhead is shown as a function of the erasure probability in the BS-to-relay link. As expected we note that the performance of the system degrades with increasing erasure probability. It is clear that the gain of using a relay decreases as the quality of the BS-to-relay channel deteriorates. We observe however, that our proposed scheme always gains significantly over the ARQ scheme and the scheme in [11].

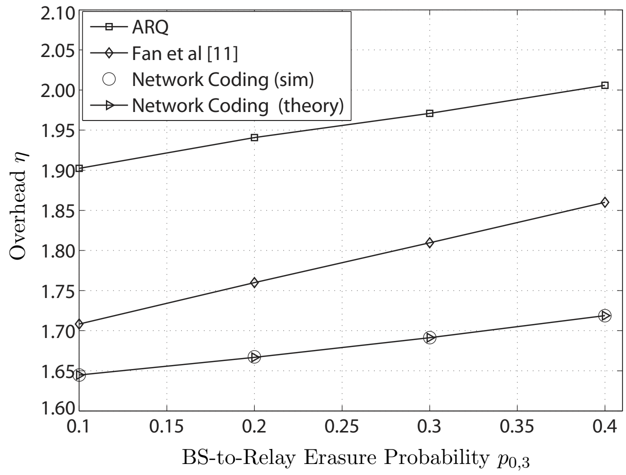

Figure 4. Overhead versus the BS-to-relay erasure probability $p_{0,3}$, where $M = 2$, $N = 1000$, $p_{0,1} = 0.5$, $p_{0,2} = 0.4$, $p_{3,1} = 0.2$ and $p_{3,2} = 0.1$.

Fig. 5 shows the overhead as a function of the erasure probability of the relay-to-user 2 link. Again, the performance of the system decreases with increasing erasure probability. Moreover, our proposed scheme gains over the ARQ scheme the the scheme in [11].

VI. Conclusion

We propose an efficient retransmission scheme based on instantaneously-decodable network coding for relay-aided broadcast systems. In our scheme the base station first broadcasts the source information packets and then, following user feedback, retransmits lost packets using instantaneously-decodable network coding. This process is repeated until the relay has successfully received all remaining packets lost by the users. The relay subsequently retransmits the remaining lost packets using the same network coding strategy to complete the broadcast. The proposed binary coding scheme has the merits of low-delay and low-complexity. Our main contributions are an efficient coding algorithm to improve the transmission efficiency, and an analytical framework for deriving the expected overhead. Numerical results meet the theoretical analysis and demonstrate that our scheme offers improved efficiency as compared to traditional ARQ and previously proposed relay-aided broadcast systems based on network coding strategies.

References

[1] H. Djandji, "An efficient hybrid ARQ protocol for point-to-multipoint communication and its throughput performance," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1688–1698, Sep. 1999.

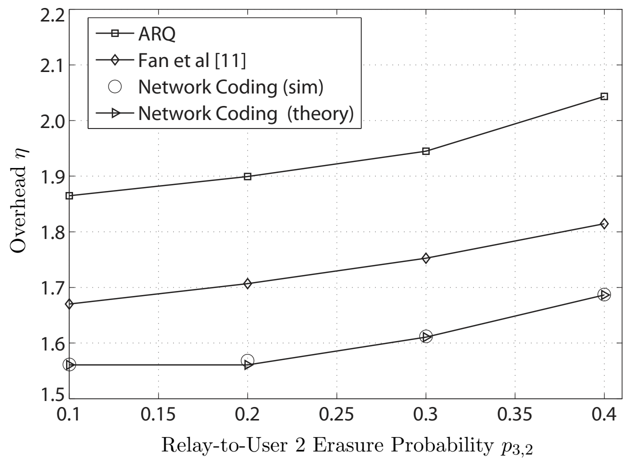

Figure 5. Overhead versus the relay-to-user 2 erasure probability $p_{3,2}$, where $M = 2$, $N = 1000$, $p_{0,1} = 0.45$, $p_{0,2} = 0.4$, $p_{0,3} = 0.3$ and $p_{3,1} = 0.1$.

[2] S. Sorour and S. Valaee, "On Minimizing Broadcast Completion Delay for Instantly Decodable Network Coding ," *IEEE Int. Conf. Commun.*, May 2010.

[3] P. Sadeghi, R. Shams, and D. Traskov, "An Optimal Adaptive network Coding Scheme for Minimizing Decoding Delay in Broadcast Erasure Channels," *EURASIP J. Wireless Commun. Networking*, Jun. 2010.

[4] L. Lu, M. Xiao, M. Skoglund, L. K. Rasmussen, G. Wu, and S. Li, "Efficient Network Coding for Wireless Broadcasting," *IEEE Wireless Commun. & Networking Conf.*, May 2010.

[5] D. Nguyen, T. Tran, T. Nguyen, and B. Bose, "Wireless broadcast using network coding," *IEEE Trans. Veh. Technol.*, Vol. 58, pp. 914–925, Feb. 2009.

[6] T. Cover and A. EL Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, 1979.

[7] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3037–3063, Sep. 2005.

[8] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Trans. Inform. Theory*, vol. 53, pp. 900–928, Mar. 2007.

[9] D. M. Pozar, *Microwave Engineering*. John Wiley & Sons, 1998.

[10] M. L. Honig (editor), *Advances in Multiuser Detection*. John Wiley & Sons, 2009.

[11] P. Fan, C. Zhi, C. Wei, K. B. Letaief, "Reliable Relay Assisted Wireless Multicast Using Network Coding," *IEEE J. Sel. Areas Commun.*, Vol. 27, pp. 749–762, Jun. 2009.

[12] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inform. Theory*, vol. 52, pp. 4413–4430, Oct. 2006.

[13] L. Lu, M. Xiao, and L. K. Rasmussen, "Relay-aided Broadcasting with Instantaneously Decodable Binary Network Codes," *Int. Conf. Comp. Commun. and Networks*, Aug. 2011.

[14] Charles. M. Grinstead and J. Laurie Snell, *Introduction to Probability, second revised edition*. American Mathematical Society, 1998.