Research Article

Full Rate Network Coding via Nesting Modulation Constellations

Suhua Tang,[1] Hiroyuki Yomo,[1,2] Tetsuro Ueda,[1] Ryu Miura,[1] and Sadao Obana[1]

[1] *ATR Adaptive Communications Research Laboratories, 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan*
[2] *Faculty of Engineering Science, Kansai University, 3-3-35 Yamate-cho, Suita, Osaka 564-8680, Japan*

Correspondence should be addressed to Suhua Tang, shtang@atr.jp

Network coding is an effective method to improving relay efficiency, by reducing the number of transmissions required to deliver data from source(s) to destination(s). However, its performance may be greatly degraded by rate mismatch, which is seldom touched in previous works and remains a challenge. In this paper, we reinterpret network coding as a mapping of modulation constellation. On this basis, we extend the mapping to support full rate network coding (FRNC), enabling simultaneous use of different modulations by nesting the low level constellation as a subset of the high level constellation. When relay links have different qualities, the messages of different flows are combined via network coding in such a way that for each relay link, its desired message is transmitted at its own highest rate. The limit in constellation size is also addressed. Compared with the state-of-the-art solutions to rate mismatch, the proposed scheme achieves the full rate of all relay links on the broadcast channel.

1. Introduction

Wireless communications suffer greatly from multipath fading where outages may degrade communication quality. Different schemes, such as adaptive modulation and coding and relay [1], have been exploited to mitigate this problem. The relay efficiency can be improved via network coding (NC) [2] if the traffic pattern and the *a priori* information are exploited. Typical transmission patterns suitable for applying NC include two-way relay [3–5], multihop forwarding [6], multiple access channel [7], multicast channel, and so forth. In addition, joint network and channel coding can further improve spectral efficiency [7, 8].

A two-way relay transmission typically consists of two stages: multiple access stage and broadcast stage. NC is applied in the second stage and has two types. The first type of NC is performed in the bit level [3]. Each node transmits its packet to the relay node successively. The relay node decodes each packet and combines them together via NC and forwards the coded packet later. The second type of NC is performed in the signal level. Typically two nodes transmit simultaneously their packets. The relay node regards the superposed signal as a NC signal and forwards it. Each

node recovers its desired signal from the NC signal with interference cancelation [4]. The NC signal is further refined in [5] by taking modulation constellations into account.

The performance of NC, especially bit level NC, is limited by several factors such as packet length mismatch (the short packets are zero padded), traffic rate mismatch (some packets cannot be network-coded due to lack of pairing packets) and transmit rate mismatch. The last factor is neglected in most previous works. In the two-way relay scenarios, the rate mismatch may be formed due to two factors. One is the relay position (which leads to relatively stable differences in link qualities) and the other is fading. Even though the relay lies exactly in the middle of two nodes, the two relay links may have different instantaneous qualities. Since the network-coded packet is intended to be received by both nodes, the minimal rate is chosen for the NC transmission [3]. In this way, the transmission at a low rate on the link supporting a high rate wastes channel bandwidth.

When NC is applied to multiple flows, the effect of rate mismatch becomes more obvious, since the minimal rate over more links only gets lower. One solution to this problem is to exploit opportunistic scheduling. Instead of transmitting network-coded packets to all potential nodes, only some

$M_1 \xleftarrow{\text{QPSK}} R \xrightarrow{\text{16 QAM}} M_2$

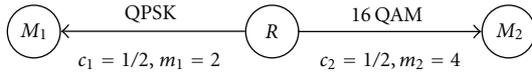$c_1 = 1/2, m_1 = 2 \qquad c_2 = 1/2, m_2 = 4$

FIGURE 1: Two-way relay, a special case of the general model.

of them are selected by taking the tradeoff between the number of links and the actual rate [9]. This opportunistic scheduling, however, cannot exploit the full power of NC.

Recently, some initial efforts were made to fully exploit rate adaptations in NC. Multiplicative NC was proposed in [10] to better adapt code and modulation rates to different links. However, it only applies to constant modulus signals. Unconventional 5-ary modulation was introduced to work together with QPSK in [11]. This makes modulation complex and is difficult to extend to other modulations. Compress and forward was studied in [12], with the assumption that the relay is much closer to the sink than nodes and has a much higher rate. Its application is limited to multiple access up-link. XOR-based NC and physical layer superposition coding were combined together to better exploit rate adaptation in [13], however, at the cost of significant power loss. Moreover, its performance is degraded and approaches that of NC when relay links have similar quality. Despite all these efforts, there is still no complete solution to the rate mismatch problem.

In this paper, we focus on the bit level NC and propose to achieve full rate network-coding (FRNC) on the broadcast channel via nesting modulation constellations. With a cross-layer design, we exploit physical layer method as well. The principle of dirty paper coding [14] indicates that *a signal known at the receiver is not interference* at all and with suitable coding the full capacity is achievable. As an analogy for NC, when the *a priori* information is available, NC should also achieve the highest rate over each link, in other words, achieve the *full rate* on the broadcast channel. This is our starting point. The basic idea is as follows: (i) at the relay node, in order to combine packets together and transmit them over links supporting different rates (modulations), the low level constellation points are nested in the high level constellation. In other words, *a subset of the high level constellation is used as the low level constellation, and this subset depends on the design of NC.* (ii) At the receiver side, nodes merely supporting low level modulation first find their constellation according to the *a priori* information and then perform demodulation and decoding. In this way, the highest rate of each link is used and the sum rate is achieved over the broadcast channel. We further study the effect of the limit in constellation size and suggest combining FRNC with superposition coding (SC). Both the analysis and simulation evaluation show that FRNC with SC is significantly superior to the state-of-the-art solutions to the rate mismatch problem.

The rest of the paper is organized as follows: the relay model is presented in Section 2 and the reinterpretation of NC as constellation mapping is addressed in Section 3. In Section 4, the detailed procedures for achieving full rate in network-coded transmissions are described. The performance of different schemes is analyzed in the two-way relay scenario in Section 5 and the related simulation evaluation is presented in Section 6. Finally, we conclude the paper with Section 7.

2. System Model

We consider a network with n nodes M_i, $i = 1, 2, \ldots, n$, and 1 relay R. For the simplicity of description, we assume that all nodes and R are synchronized and the transmissions are done in terms of TDMA (it is possible to extend the proposed scheme to other channel access methods such as CSMA.) There are n flows. The ith flow from M_i^S to M_i^D goes through the two-hop path M_i^S-R-M_i^D and the actual transfer of packets is via the relay R. The packet transfer is divided into two stages: (a) multiple access stage where R collects packets from all nodes. Each node M_i^S sends packet P_i to R in the ith slot, using its optimal rate (modulation and coding). After the data transmission, each node reports its receiving status of packets to R, in a similar way as the COPE scheme [15]. Based on such a feedback, R makes the NC scheduling, selecting a subset of n_1 nodes, each of which knows all packets involved in the NC except its own desired packet. Without loss of generality, in the following, we assume n_1 equals n. Then after n slots, M_i^D knows $P_1, \ldots, P_{i-1}, P_{i+1}, \ldots, P_n$. (b) Broadcast stage where R forwards packets to all nodes. R transmits $P_\Sigma = \oplus_i P_i$ ($\oplus$ represents the bitwise exclusive or (XOR) operation) to all nodes, and M_i^D recovers P_i by $(\oplus_{j \neq i} P_j) \oplus P_\Sigma$. Hence, n packets are exchanged in $n+1$ slots. A typical example for the model is the local area wireless networks where nodes exchange packets via their associated access point (R in the model). A special case is $n = 2$, corresponding to the two-way relay in Figure 1. In this special case, each node uses its transmitted packet as the *a priori* information and the feedback of receiving status of packets is unnecessary.

Over each link, the rate can be adjusted by modulation and coding. With coding rate c and modulation level m (constellation size $= 2^m$), on average $c \cdot m$ bits can be transmitted by each symbol. Generally, the modulation level determines a rate range, within which the coding scheme further fine adjusts the rate. Different modulation levels have constellations with different sizes. But these constellations all have the same normalized energy.

In the above model, NC is used at the second stage. We focus on this stage and exploit joint design of modulation and coding so that the highest rate of each link is realized in the NC transmission. We take the following assumptions: (i) R has collected enough data for each flow so that the zero-padding is unnecessary in the NC transmission, (ii) the *a priori* information required for network decoding is available at each node by recording the overheard packets, and (iii) the relay node knows the channel state information of all links. *The key problem is how to realize full rate on all links simultaneously.*

3. Reinterpretation of Network Coding

In conventional bit level NC schemes, bits from different flows are XORed together, channel coded, modulated, and
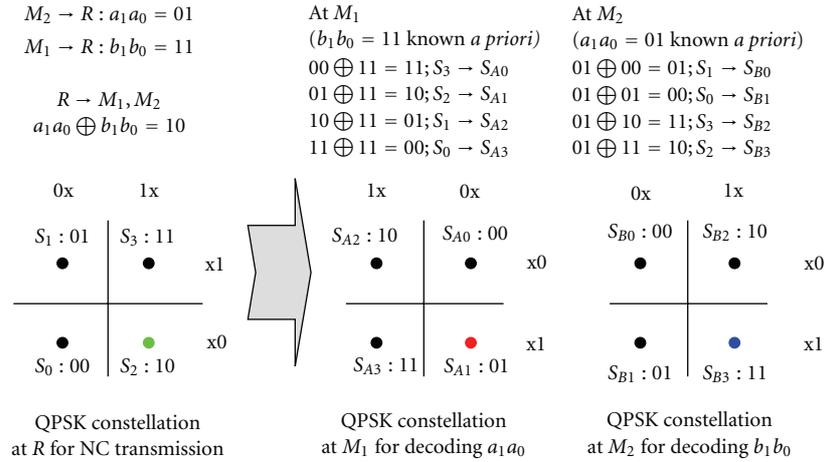
FIGURE 2: Reinterpretation of network-coding by modulation.

then transmitted. The receiver just works in the reverse way to recover its information bits. Due to the linearity of both channel coding and NC, their order can be exchanged [16]. In this paper, the NC operation is done after channel coding.

Although the NC is performed in the bit level, it can be reinterpreted as a function of constellation mapping. This is explained by an example shown in Figure 2 using the typical QPSK constellation with gray codes. Assume R relays $a_1 a_0$ ("01") from M_2 to M_1, and $b_1 b_0$ ("11") from M_1 to M_2, respectively. When relaying these bits, R combines them together as "10" by XOR and transmits $(S_2)_{\text{QPSK}}$. M_1 already knows $b_1 b_0$ ("11"). Exploiting this as the *a priori* information, for all possible bits of $a_1 a_0$, the NC bits and corresponding signals can be computed locally at M_1. These signals, when interpreted as the points for $a_1 a_0$, construct a new QPSK constellation $(S_{A0} S_{A1} S_{A2} S_{A3})_{\text{QPSK}}$. This constellation has the same size as $(S_0 S_1 S_2 S_3)_{\text{QPSK}}$ but with a different layout due to NC. In this way, the NC function actually provides a mapping between constellations. Instead of a fixed constellation in conventional modulations, such a mapping depends on the *a priori* information and changes for each symbol. The reinterpretation of NC can be summarized as follows:

(i) R transmits $F(P_{1,c}, P_{2,c}, \ldots, P_{n,c})$ where $P_{i,c}$ is the channel-coded packet of the ith flow, and F involves NC (XOR in conventional NC) and modulation.

(ii) At the receiver side, $F^{-1}(ap_i)$ ($ap_i = \oplus_{j \neq i} P_{j,c}$ in conventional NC) provides the constellation for demodulating $P_{i,c}$, where ap_i is the *a priori* information at the ith receiver.

In conventional NC schemes, the constellations for different flows have the same size and min-distance, and the latter is the main factor that decides the rate. This forces the relay node to transmit the XORed packet with the minimal rate of all links so that all nodes can correctly recover the XORed packet.

The above constellation mapping can be extended to support simultaneous use of modulations with different levels (size and min-distance). Specifically, we use constellation-compatible modulations and nest the low level constellation inside the high level constellation. For example, a subset of four 16 QAM constellation points can be used as the QPSK constellation so that over the broadcast channel QPSK is used for one link while 16 QAM is used for the other link in the two-way relay scenario. In other words, among the links supporting different modulations, the highest modulation level is always used as the container. Other low level modulations use a subset of the high level constellation as their constellations.

4. Full Rate Network Coding Protocol

In this section, we present the full rate network-coding (FRNC) protocol. First the basic idea is explained with a simple example. Then the idea is generalized. How to nest constellations, how to find the actual constellation under the NC operation, how to transmit at the relay, and how to receive at the nodes are successively described in detail.

4.1. An Example Revealing the Basic Idea. With the two-node ($n = 2$) scenario in Figure 1, we show how to use different rates over different links in the network-coded transmission. Assume that (i) links between R and M_1/M_2 support rates with $c_1 = 1/2$, $m_1 = 2$ (QPSK), $c_2 = 1/2$, $m_2 = 4$ (16 QAM), respectively. On average, R can forward $r_1 = c_1 \cdot m_1 = 1$ bit/symbol to M_1 and $r_2 = c_2 \cdot m_2 = 2$ bit/symbol to M_2. (ii) The slot length is $N = 2$ symbols. Then 2 bits to M_1 or 4 bits to M_2 can be transmitted in a single slot. (iii) The two bits from R to M_1 are $P_{1,u} = $ "10" and 4 bits from R to M_2 are $P_{2,u} = $ "1101".

The transmit procedure is shown in Figure 3. At R, $P_{1,u}$ and $P_{2,u}$ are channel-coded to $P_{1,c} = $ "1101" and $P_{2,c} = $ "11100010". The modulations for the two messages are QPSK and 16 QAM, respectively. To transmit the two messages together via NC, the QPSK constellation for $P_{1,c}$ is nested in the 16 QAM constellation used for $P_{2,c}$. The nesting is realized by postcoding. In this example, by repetition codes with rate = 1/2, $P_{1,c}$ is encoded to $P_1 = $ "11110011", with the

TABLE 1: Constellation conversion from 16 QAM to QPSK ($a_3a_2a_1a_0$ represents the *a priori* info, b_1b_0 are the info bits to be received).

a priori info $a_3a_2a_1a_0$	QPSK constellation $(a_3a_2a_1a_0 \oplus b_1b_1b_0b_0)$
0000	$(S_0, S_3, S_{12}, S_{15})_{16\,\text{QAM}}$
0001	$(S_1, S_2, S_{13}, S_{14})_{16\,\text{QAM}}$
0010	$(S_2, S_1, S_{14}, S_{13})_{16\,\text{QAM}}$
0011	$(S_3, S_0, S_{15}, S_{12})_{16\,\text{QAM}}$
0100	$(S_4, S_7, S_8, S_{11})_{16\,\text{QAM}}$
0101	$(S_5, S_6, S_9, S_{10})_{16\,\text{QAM}}$
0110	$(S_6, S_5, S_{10}, S_9)_{16\,\text{QAM}}$
0111	$(S_7, S_4, S_{11}, S_8)_{16\,\text{QAM}}$
1000	$(S_8, S_{11}, S_4, S_7)_{16\,\text{QAM}}$
1001	$(S_9, S_{10}, S_5, S_6)_{16\,\text{QAM}}$
1010	$(S_{10}, S_9, S_6, S_5)_{16\,\text{QAM}}$
1011	$(S_{11}, S_8, S_7, S_4)_{16\,\text{QAM}}$
1100	$(S_{12}, S_{15}, S_0, S_3)_{16\,\text{QAM}}$
1101	$(S_{13}, S_{14}, S_1, S_2)_{16\,\text{QAM}}$
1110	$(S_{14}, S_{13}, S_2, S_1)_{16\,\text{QAM}}$
1111	$(S_{15}, S_{12}, S_3, S_0)_{16\,\text{QAM}}$

TABLE 2: A comparison of the broadcast channel among three schemes, for the scenario shown in Figure 1.

scheme	rate	# transmitted bits
DF	$(r_1/2) + (r_2/2)$	3
NC (min rate)	$\min(r_1, r_2) \cdot 2$	4
FRNC (full rate)	$r_1 + r_2$	6

same length as $P_2 = P_{2,c} =$ "11100010". The XORed sum of P_1 and P_2 is $P_\Sigma =$ "00010001". Then P_Σ is modulated with the 16 QAM constellation (using gray codes) shown in Figure 4, and R transmits $x_\Sigma = (S_1 S_1)_{16\,\text{QAM}}$.

The receive procedure is shown in Figure 5. At the *i*th node, the signal received from R is $s_i(t)$. For simplicity, noise and channel fading are ignored. In the same way as the relay performs channel coding and post-coding, $P_1 =$ "11110011" is calculated from $P_{1,u}$ at M_2, and $P_2 =$ "11100010" is calculated from $P_{2,u}$ at M_1, and are used as the *a priori* information in the network decoding stage.

Since $s_2(t)$ has the same modulation level as the one supported by the quality of link M_2R, decoding at M_2 is the same as usual. At first, $P_\Sigma =$ "00010001" is demodulated from s_2. With $P_1 =$ "11110011" as the *a priori* information, $P_{2,c} = P_2 =$ "11100010" is obtained and then $P_{2,u} =$ "1101" is channel decoded. $s_1(t)$ has a higher modulation level than the one supported by the quality of link M_1R. Therefore, the decoding at M_1 is a little more complex. The QPSK constellation to be used at M_1 depends on the *a priori* information and has to be constructed from the 16 QAM constellation. With repetition codes used in the post-coding stage in this example, two bits b_1b_0 carried in a QPSK symbol are post-coded to $b_3'b_2'b_1'b_0' = b_1b_1b_0b_0$, corresponding to a 16-QAM symbol. With $a_3a_2a_1a_0$ as the *a priori* information, varying b_1b_0, the possible NC bits $a_3a_2a_1a_0 \oplus b_3'b_2'b_1'b_0'$ and the corresponding signals can be computed. Table 1 shows the derived QPSK constellations for demodulating b_1b_0, with the four *a priori* bits $a_3a_2a_1a_0$ as an index.

At M_1, with $P_2 =$ "11100010" known *a priori*, for the first symbol in x_Σ, $a_3a_2a_1a_0 =$ "1110", $(S_1)_{16\,\text{QAM}}$ is to be demodulated with the QPSK constellation $(S_{14}, S_{13}, S_2, S_1)_{16\,\text{QAM}}$

(refer to Table 1); for the second symbol in x_Σ, $a_3a_2a_1a_0 =$ "0010", $(S_1)_{16\,\text{QAM}}$ is to be demodulated with the QPSK constellation $(S_2, S_1, S_{14}, S_{13})_{16\,\text{QAM}}$. Here, $x_\Sigma = (S_1S_1)_{16\,\text{QAM}}$ logically corresponds to $(S_3S_1)_{\text{QPSK}}$. It is demodulated to $P_{1,c} =$ "1101" and converted to $P_{1,u} =$ "10" after channel decoding. In this way network decoding is realized by the constellation conversion.

A simple comparison on the broadcast channel, among decode-and-forward (DF), bit level NC with minimal rate, and FRNC, is summarized in Table 2. With DF, R uses one symbol for each node and thus transmits 3 bits in total. With NC, R transmits $(\min(r_i) \cdot 2) \cdot 2 = 4$ bits. With FRNC, R transmits $(\sum r_i) \cdot 2 = 6$ bits using two symbols.

Although superposition coding handles links with different qualities as well, the proposed FRNC scheme is quite distinct from it. Constellation nesting fully exploits the power on each link by using the *a priori* information in times of decoding. As a comparison, superposition coding does not exploit the *a priori* information for decoding the desired signal. Instead, the total power is divided into two parts, most power for the base layer signal and little power for the secondary layer signal, which results in significant power loss in transmitting the secondary layer signal.

In the following sections, we extend the above idea to more general cases and study the related post-coding scheme and the decoding method.

4.2. Nesting Constellations. We first consider how to nest N_1-QAM (we focus on the constellations in the form of grid, extension to other forms of constellations is also possible.) in N_2-QAM ($N_2 > N_1$, $N_k = (n_k)^2 = 2^{m_k}$, $k = 1, 2$). A simple way is to choose a subset of N_2-QAM constellation points as the N_1-QAM constellation. We construct the N_1-QAM constellation by dividing N_2-QAM into subsets. Let the min-distance of N_2-QAM be d_2. The points of N_1-QAM with a distance $d_1 = n_2/n_1 \cdot d_2$ to their neighbors are grouped into the same subset. In this way, the N_2-QAM constellation is divided into $(n_2/n_1)^2$ subsets, each of which having the same size N_1.

Choosing points for (non-QAM) BPSK requires one more step: after choosing a subset for QPSK from the QAM constellation, select two diagonal points from the QPSK constellation for BPSK.

A single point of the N_1-point constellation can only carry m_1 information bits. But after nesting it inside the N_2-constellation, its bit vector is extended to m_2 bits. There should be a bit-mapping. The only subset, which contains the all-zero bit-vector, is used for bit mapping from m_1 to m_2.

Figure 4 shows an example of dividing 16 QAM to find QPSK constellations, where $N_2 = 16$, $m_2 = 4$, $N_1 = 4$,
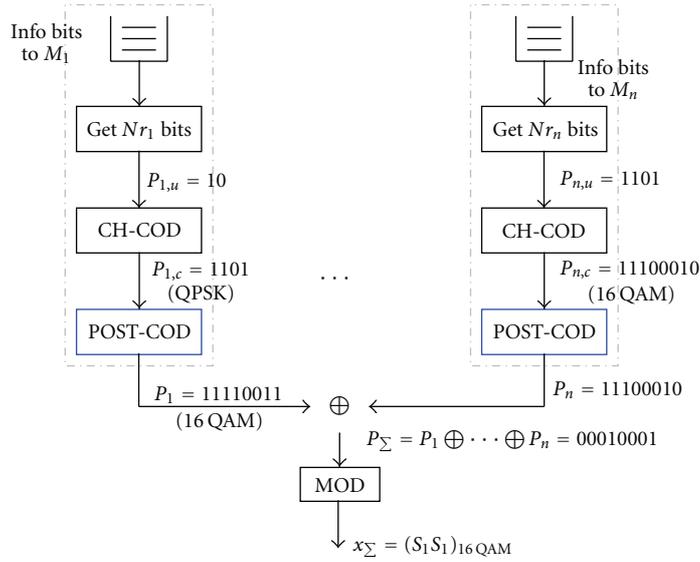
Info bits to M_1

Get Nr_1 bits

$P_{1,u} = 10$

CH-COD

$P_{1,c} = 1101$ (QPSK)

POST-COD

$P_1 = 11110011$ (16 QAM)

. . .

Info bits to M_n

Get Nr_n bits

$P_{n,u} = 1101$

CH-COD

$P_{n,c} = 11100010$ (16 QAM)

POST-COD

$P_n = 11100010$

$\oplus$

$P_\Sigma = P_1 \oplus \cdots \oplus P_n = 00010001$

MOD

$x_\Sigma = (S_1 S_1)_{16\,QAM}$

FIGURE 3: Coding and modulation at the relay node.

00xx (−0.948)	01xx (−0.316)	11xx (0.316)	10xx (0.948)	
$S_2 : 0010$	$S_6 : 0110$	$S_{14} : 1110$	$S_{10} : 1010$	xx10 (0.948)
$S_3 : 0011$	$S_7 : 0111$	$S_{15} : 1111$	$S_{11} : 1011$	xx11 (0.316)
$S_1 : 0001$	$S_5 : 0101$	$S_{13} : 1101$	$S_9 : 1001$	xx01 (−0.316)
$S_0 : 0000$	$S_4 : 0100$	$S_{12} : 1100$	$S_8 : 1000$	xx00 (−0.948)
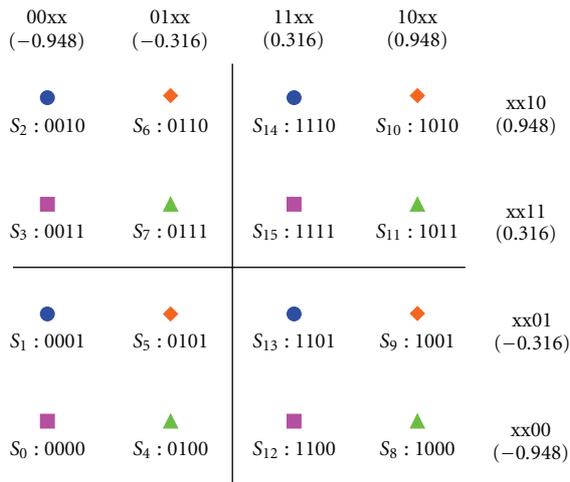
FIGURE 4: Nesting QPSK constellation in 16 QAM constellation.

TABLE 3: Some bit mapping methods.

Nesting Method	Bits in low-level constellation	Bits for sub-set in container constellation
BPSK in QPSK	0, 1	00, 11
BPSK in 16 QAM	0, 1	0000, 1111
QPSK in 16 QAM	00, 01	0000, 0011
	10, 11	1100, 1111
QPSK in 64 QAM	00, 01	000000, 000110
	10, 11	110000, 110110
16 QAM in 64 QAM	0000, 0001	000000, 000011
	0010, 0011	000101, 000110
	0100, 0101	011000, 011011
	0110, 0111	011101, 011110
	1000, 1001	101000, 101011
	1010, 1011	101101, 101110
	1100, 1101	110000, 110011
	1110, 1111	110101, 110110

$m_1 = 2$. The 16 QAM constellation is divided into four subsets: $CS_1 = (S_0, S_3, S_{12}, S_{15})$, $CS_2 = (S_1, S_2, S_{13}, S_{14})$, $CS_3 = (S_4, S_7, S_8, S_{11})$, $CS_4 = (S_5, S_6, S_9, S_{10})$. $CS = CS_1 \cup CS_2 \cup CS_3 \cup CS_4$ is the constellation for 16 QAM. QPSK may use any CS_i as its constellation point, although with a different layout under NC.

Table 3 shows some bit mapping methods, where the N_1 m_1-bit vectors are one-to-one mapped to N_1 m_2-bit vectors in the subset containing the all-zero vector. The left column represents the nesting method, the second column is the original bits to be transmitted with low level constellation, and the right column shows the bit vectors in the nested constellation. With the bit mapping, the bits of low-level constellations are modulated to the subsets of high-level modulation, and the function of post-coding in Figure 3 is realized.

The constellation nesting may have SNR loss since the min-distance of the nested constellation may be a little less than that of the standard one. For example, with normalized energy, the min-distance of two 16-QAM points equals 0.6325. When nesting QPSK inside 16 QAM, the min-distance equals $0.6325 \cdot 2 = 1.265$, which is 0.97 dB less than 1.414, the min-distance of normal QPSK constellation. Table 4 shows the SNR loss, where the horizontal and vertical labels stand for original constellations and container constellations, respectively. Although nesting QPSK in 16 QAM has
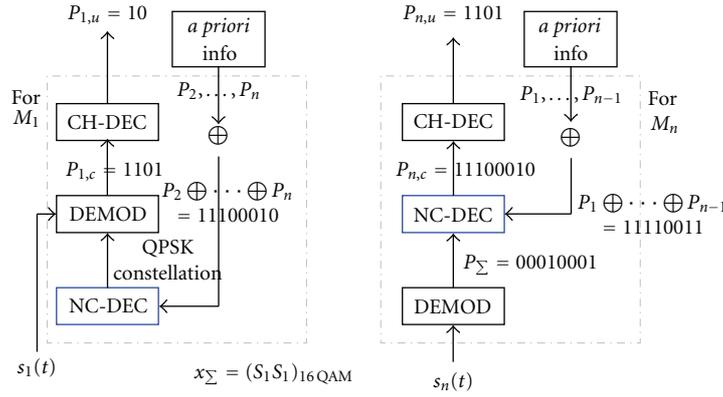
FIGURE 5: Recover information bits at the nodes.

TABLE 4: Potential SNR loss in constellation conversion.

	BPSK	QPSK	16 QAM	64 QAM
QPSK	0			
16 QAM	−0.97 dB	−0.97 dB		
64 QAM	−1.18 dB	−1.18 dB	−0.21 dB	
256 QAM	−1.23 dB	−1.23 dB	−0.26 dB	−0.05 dB

SNR loss of about 0.97 dB, nesting other constellations has little SNR loss (0.05 dB for 64 QAM in 256 QAM) or no SNR loss (0 dB for BPSK in QPSK) at all. The SNR loss is taken into account when choosing rate (modulation and coding) according to SNR.

4.3. Actual Constellation under Network Coding. The next important issue is to find the actual constellation layout under NC. We explain this with Figure 4 as an example. It can be easily verified that the division has the following property:

$$\forall P_1 \in CS_1, \ \forall ap_1 \in CS, \ \text{if } ap_1 \in CS_i, \\ \text{then } P_1 \oplus ap_1 \in CS_i. \tag{1}$$

Equation (1) shows that, for a point ap_1 in the high level constellation ($ap_1 \in CS$), if ap_1 is in the subset CS_i, it maps CS_1 to CS_i by the NC operation. The actual constellation layout of CS_i for demodulating P_1 is determined by $P_1 \oplus ap_1$, with ap_i known *a priori* at the receiver. Although a constellation under NC changes with the *a priori* information, the min-distance for N_1-QAM, under all the *a priori* information, remains the same: $d_1 = n_2/n_1 \cdot d_2$.

As for the example in Figure 1, using the third row of Table 3, two bits $b_1 b_0$ are post-coded to $b_3' b_2' b_1' b_0'$. With $a_3 a_2 a_1 a_0 \oplus b_3' b_2' b_1' b_0'$ being received and $ap_i = a_3 a_2 a_1 a_0$ known *a priori* at M_1, the QPSK constellation for demodulating $b_1 b_0$ is looked up in Table 1 by using $a_3 a_2 a_1 a_0$ as an index. For example, when $a_3 a_2 a_1 a_0 = $ "1110", $(S_{14}, S_{13}, S_2, S_1)_{16\,\text{QAM}}$ is equivalent to $(S_0, S_1, S_2, S_3)_{\text{QPSK}}$. Since the derived QPSK constellation depends on the *a priori* information, it changes for each symbol. Recovery of other constellations can be done in a similar way.

TABLE 5: SNR threshold for rate adaptation (for a message consisting of 4800 symbols).

SNR (dB)	Modulation and coding	Bit/Sym
	BPSK (1/2)	0.50
≥7.0	BPSK (3/4)	0.75
≥7.6	QPSK (1/2)	1.00
≥10.4	QPSK (3/4)	1.50
≥12.8	16 QAM (1/2)	2.00
≥17.0	16 QAM (3/4)	3.00
≥21.0	64 QAM (2/3)	4.00
≥23.4	64 QAM (3/4)	4.50
≥26.8	256 QAM (2/3)	5.33
≥28.0	256 QAM (3/4)	6.00

4.4. Encoding/Modulation at the Relay. Figure 3 shows the transmit procedure at relay R. For each flow f_i, according to the SNR of its relay link, R finds the transmit rate r_i from an empirical SNR-rate table shown in Table 5. SNR loss due to constellation nesting is considered in this process: the rate corresponding to SNR is lowered if SNR loss makes this rate improper.

Assume, without loss of generality, that rates, $r_1, r_2, \dots, r_n$, over links from R to $M_1, M_2, \dots, M_n$, are in the increasing order, that is, $r_1 \le r_2 \cdots \le r_n$. Each $r_i = c_i \cdot m_i$ corresponds to a coding rate c_i and a modulation level m_i. $m_i, i = 1, 2, \dots, n$, are also in the increasing order.

Transmission at R is done by the following steps.

(i) Every time R transmits a fixed number of symbols, N. For each flow f_i, the number of information bits that can be transmitted is $N \cdot r_i$. These information bits form a frame $P_{i,u}$. On $P_{i,u}$ channel coding with rate c_i is performed, which generates $P_{i,c}$. $P_{i,c}, i = 1, 2, \dots, n$, have different length in bits.

(ii) In order to transmit $P_{i,c}$, the constellation m_i should be nested in m_n. This is done by the bit mapping (POST-COD) according to Table 3, which maps m_i to m_n and encodes $P_{i,c}$ to P_i.

(iii) $P_i, i = 1, \ldots, n$ are XORed together as $P_\Sigma = \oplus_i P_i$. P_Σ, modulated to signal x_Σ with constellation m_n, is transmitted to all nodes with out-of-band rate information $r_i, i = 1, 2, \ldots, n$ (each node only records the information bits on overhearing packets from nearby nodes to the relay. With the rate information from the relay, the node performs the same channel coding/post coding as the relay and calculates the coded bits as the *a priori* information for network decoding.)

4.5. Demodulation/Decoding at the Receiver.

Figure 5 shows the demodulation and decoding procedure at all nodes. At the ith node, the signal received from R is

$$s_i(t) = h_i \cdot x_\Sigma(t) + n_i(t), \qquad (2)$$

where h_i is the channel gain and $n_i(t)$ is zero mean additive white Gaussian noise (AWGN).

A node M_i, supporting the used constellation ($m_i = m_n$), performs soft demodulation and calculates symbol log-likelihood ratio (LLR) [17] and then converts to bit LLR. The bit LLR corresponds to XORed bits from all flows. With the *a priori* information bits ($ap_i = \oplus_{j \neq i} P_j$) known in advance, the LLR of desired bits can be recovered and then channel decoding is performed. The whole procedure is shown in the right side of Figure 5.

For a receiver M_i requiring a lower constellation ($m_i < m_n$), at first the low-level constellation is derived by exploiting the *a priori* information, as described in Section 4.3. This derivation of constellation is actually network decoding. Then the received signal is demodulated with the derived constellation and later channel decoded to recover the bits, as shown in the left side of Figure 5.

4.6. Discussion: Constellation Size Limit.

FRNC requires that constellation size should be large enough so that high rate can be used at high SNR. In practical systems, there is a constraint on the constellation size which restricts the maximal rate. The maximum of constellation size, referred to as the *constellation size limit* hereafter, confines the performance of FRNC. In such cases, FRNC can be used together with superposition coding (SC) to fully exploit the transmit power. NC is already used together with SC in [13], where the fine scheduling is used to combine links with almost the same gain and apply NC to them. For links with quite different gains, SC is applied. But such a scheduling heavily depends on the actual topology. As a comparison, we replace the NC in [13] with FRNC and suggest FRNC + SC, which is analyzed in Section 5.

5. Performance Analysis of Two-Way Relay

In this section, we analyze the performance of different schemes under the two-way relay scenario shown in Figure 1. Six schemes are compared here. (i) DF, the decode-and-forward scheme, (ii) NC, the normal bit level NC scheme with minimal rate constraint, (iii) SC, the superposition coding scheme, (iv) NC + SC, the iPack scheme in [13], (v) FRNC, and (vi) FRNC + SC. In the analysis, we assume (i) each channel has zero mean AWGN noise with variance σ_n^2; (ii) channel gain of the link M_i-R is h_i; (iii) each symbol has normalized energy and the SNR over each link is $\gamma_i = |h_i|^2/\sigma_n^2$, (iv) the packet length is infinite.

5.1. Capacity without Fading.

With FRNC, the capacity of the broadcast channel reaches the sum rates of the two links (here, we ignore SNR loss in constellation nesting for simplicity. the SNR loss is taken into account in the simulation evaluation),

$$c_{\text{FRNC}}(\gamma_1, \gamma_2) = \log_2(1 + \gamma_1) + \log_2(1 + \gamma_2). \qquad (3)$$

The capacity of DF is half of that of FRNC,

$$c_{\text{DF}} = \frac{c_{\text{FRNC}}}{2}. \qquad (4)$$

The capacity of NC is

$$c_{\text{NC}}(\gamma_1, \gamma_2) = 2 \cdot \log_2(1 + \min(\gamma_1, \gamma_2)). \qquad (5)$$

Next we calculate the throughput of NC + SC, SC, FRNC + SC, where SC is involved. Assume, without loss of generality, that $\gamma_2 \geq \gamma_1$. Part ($0 \leq \alpha \leq 1$) of the power is used to transmit the base layer signal x_Σ (the NC coded message in NC + SC, the plain message in pure SC, the FRNC coded message in FRNC + SC), and the remaining power ($1 - \alpha$) is used to transmit the secondary signal x_2 to M_2. The transmitted signal is

$$x(t) = \sqrt{1 - \alpha} \cdot x_2 + \sqrt{\alpha} \cdot x_\Sigma. \qquad (6)$$

At M_i, SNR of the received base layer signal (x_Σ) is

$$\gamma_i' = \frac{\alpha \cdot |h_i|^2}{(1 - \alpha) \cdot |h_i|^2 + \sigma_n^2} = \frac{\alpha \cdot \gamma_i}{(1 - \alpha) \cdot \gamma_i + 1}. \qquad (7)$$

Since γ_i' is an increasing function of γ_i, $\min(\gamma_1', \gamma_2') = \gamma_1'$ and the rate used for the NC coded message is determined by γ_1'. At M_2, after perfect interference cancellation, the SNR of the secondary layer signal is

$$\gamma_2'' = \frac{(1 - \alpha) \cdot |h_2|^2}{\sigma_n^2} = (1 - \alpha) \cdot \gamma_2. \qquad (8)$$

The throughput of NC + SC under given γ_1, γ_2, and α, is

$$\begin{aligned} c_{\text{NC}+\text{SC}}(\gamma_1, \gamma_2, \alpha) &= 2\log_2(1 + \gamma_1') + \log_2(1 + \gamma_2'') \\ &= 2\log_2(1 + \gamma_1) - 2\log_2(1 + (1 - \alpha) \cdot \gamma_1) \\ &\quad + \log_2(1 + (1 - \alpha) \cdot \gamma_2). \end{aligned} \qquad (9)$$

By differentiating $c_{\text{NC}+\text{SC}}(\gamma_1, \gamma_2, \alpha)$ with respect to α, its maximum can be obtained at $\alpha = 1 - (1/\gamma_1 - 2/\gamma_2)$. To achieve

the maximal capacity, the power allocation should be done as follows:

$$\alpha = 1, \qquad 2\gamma_1 > \gamma_2 \geq \gamma_1,$$

$$\alpha = 1 - \left(\frac{1}{\gamma_1} - \frac{2}{\gamma_2}\right), \quad \gamma_1 \geq 1, \ \gamma_2 \geq 2\gamma_1 \text{ or}$$

$$1 > \gamma_1 \geq 0, \ \frac{2\gamma_1}{(1-\gamma_1)} > \gamma_2 \geq 2\gamma_1 \qquad (10)$$

$$\alpha = 0, \quad 1 > \gamma_1 \geq 0, \ \gamma_2 \geq \frac{2\gamma_1}{(1-\gamma_1)}.$$

The capacity of the pure SC is as follows:

$$\begin{aligned} c_{SC}(\gamma_1, \gamma_2, \alpha) &= \log_2(1 + \gamma_1') + \log_2(1 + \gamma_2'') \\ &= \log_2(1 + \gamma_1) - \log_2(1 + (1-\alpha) \cdot \gamma_1) \\ &\quad + \log_2(1 + (1-\alpha) \cdot \gamma_2). \end{aligned}$$
$$(11)$$

This is a decreasing function of α and reaches its maximum when α approaches 0. With the constellation size limit in practical systems, over the link M_2-R, all SNR greater than a certain threshold γ_{max} supports the maximal rate. When γ_2 is large enough, it is sufficient to choose α so as to satisfy $(1-\alpha) \cdot \gamma_2 \geq \gamma_{max}$. The rest of the power can be used for the SC transmission over the link M_1-R.

As for FRNC + SC, the FRNC coded message replaces x_Σ in (6), and its capacity is as follows:

$$\begin{aligned} &c_{FRNC+SC}(\gamma_1, \gamma_2, \alpha) \\ &= \log_2(1 + \gamma_1') + \log_2(1 + \gamma_2') + \log_2(1 + \gamma_2'') \\ &= \log_2(1 + \gamma_1) - \log_2(1 + (1-\alpha) \cdot \gamma_1) + \log_2(1 + \gamma_2). \end{aligned}$$
$$(12)$$

It is interesting to see that the power allocation has no capacity loss over the link R-M_2. The only loss compared with FRNC is the part $\log_2(1+(1-\alpha) \cdot \gamma_1)$ over the R-M_1 link, which approaches 0 as α approaches 1. $c_{FRNC+SC}(\gamma_1, \gamma_2, \alpha)$ is an increasing function of α. Without constellation size limit, α should be set to 1, and FRNC + SC degenerates to FRNC. With the constellation size limit, devoting full power to transmitting the FRNC coded packet is unnecessary. Therefore, the SC coding should be used and γ_2 is divided into two parts, γ_2' and γ_2''. The optimal power allocation policy is as follows:

$$\alpha = 1, \qquad\qquad \gamma_2 \leq \gamma_{max}$$

$$\alpha = \frac{\gamma_{max}}{\gamma_{max} + 1} \cdot \frac{\gamma_2 + 1}{\gamma_2}, \qquad \text{otherwise} \qquad (13)$$

$$\alpha = 1 - \frac{\gamma_{max}}{\gamma_2}, \qquad \gamma_2 \geq \gamma_{max}^2 + 2 \cdot \gamma_{max}.$$

The power allocation can be explained as follows: (i) when γ_2 is small enough ($\gamma_2 \leq \gamma_{max}$), all power ($\alpha = 1$) should be used for FRNC since its rate is not saturated yet. (ii) As γ_2 gets greater than γ_{max}, α should be set to satisfy $\gamma_2' = \gamma_{max}$.
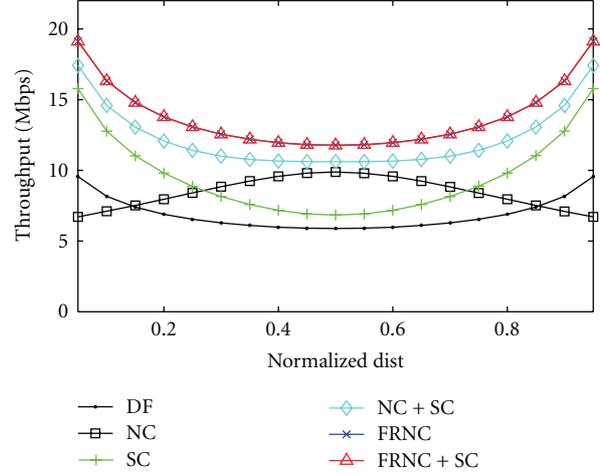

FIGURE 6: Throughput achieved by different schemes on the broadcast channel-effect of relay position for two way relay (theoretical calculation, no constellation size limit).

The maximal rate is used over link M_2-R for FRNC, and the extra power is used for the SC transmission over the link M_2-R. (iii) If γ_2 is very large, both the FRNC and SC transmission reach the maximal rate over the link M_2-R, and α is chosen to satisfy $(1-\alpha) \cdot \gamma_2 \geq \gamma_{max}$ for the SC transmission. The rest power is used in improving the FRNC rate over the link M_1-R.

5.2. Capacity with Fading. Next we consider the effect of fading and assume each channel experiences block Rayleigh fading. γ_i follows the exponential distribution: $f_{\bar{\gamma}_i}(\gamma_i) = 1/\bar{\gamma}_i \cdot e^{-\gamma_i/\bar{\gamma}_i}$, and the joint distribution is $f(\gamma_1, \gamma_2) = f_{\bar{\gamma}_1}(\gamma_1) \cdot f_{\bar{\gamma}_2}(\gamma_2)$. The average throughput can be calculated by numerical integration.

With a two-way relay scenario similar to the one shown in Figure 1, we study how the position of the relay node affects the system performance. Adjusting the position of R between M_1 and M_2 changes the normalized distance $d_{M_1R}/d_{M_1M_2}$. Average SNR ($\bar{\gamma}_i$) of links M_1R and M_2R is calculated from the normalized distance $d_{M_1R}/d_{M_1M_2}$ according to the two-ray model [18] with the path loss exponent (equaling 3 in the simulation). When R lies in the middle of M_1 and M_2, the average SNR of both relay links equals 20 dB. α is set to the optimal value for schemes employing SC.

Figure 6 shows the average throughput of different schemes, where there is no limit on constellation size. The curve of FRNC + SC overlaps with that of FRNC. Both outperform other schemes. As analyzed before, SC in NC + SC only works under certain conditions. When the relay is near the middle point, the difference in link quality is not very large. NC + SC degenerates to NC, as is clear when the distance equals 0.5. Without constellation size limit, the best link is always chosen in SC, and the two-way communication becomes unidirectional. The performance of NC is greatly affected by the min-rate, especially when the relay is away from the middle point and the difference in link quality becomes large. Due to the effect of fading, two links with
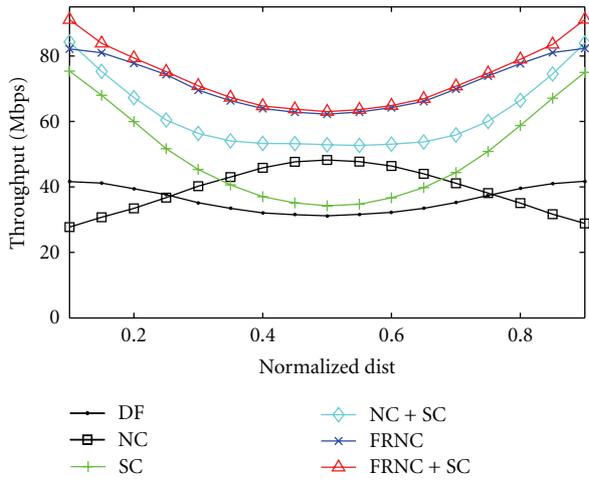
FIGURE 7: Throughput achieved by different schemes on the broadcast channel-effect of relay position for two way relay (simulation results, largest constellation is 256 QAM).

FIGURE 9: Throughput achieved by different schemes on the broadcast channel-effect of relay position for two way relay.
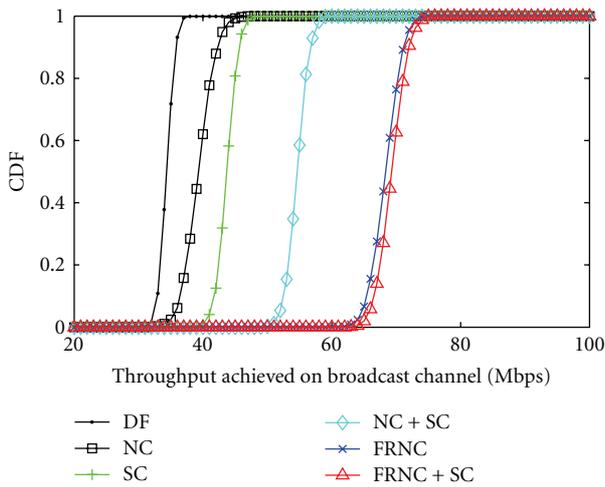

FIGURE 8: Cumulative density function of throughput achieved on the broadcast channel (normalized distance = 0.3 in Figure 7).

the same average SNR have different instantaneous SNR. Therefore, FRNC/FRNC + SC outperform NC and NC + SC even when the normalized distance equals 0.5.

6. Numerical Results

In this section, we evaluate the proposed FRNC and FRNC + SC schemes using Monte-Carlo simulations. Each slot consists of 4800 symbols. Messages are coded by a 4-state recursive systematic convolutional (RSC) code with the generator matrix (1, 5/7). Modulation and coding schemes shown in Table 5 are used. Altogether, 10 different transmit rates can be supported. The number of information bits in a message varies from 2400 bits to 28800 bits. Messages are transmitted from R to nodes via different schemes and decoded accordingly. In the evaluation, we focus on the
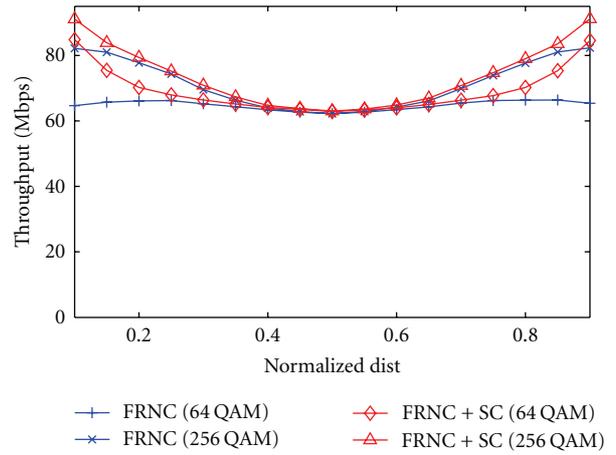
broadcast channel, and compare FRNC, FRNC + SC against DF, NC [3], NC with opportunistic scheduling (NCSched) [9], SC, and NC + SC [13]. SNR loss in FRNC is taken into account when choosing rates for transmissions. It is assumed that each link experiences independent block Rayleigh fading. Modulation constellations are adopted from IEEE 802.11a and the related parameters (symbol period, number of subcarrier) are used in calculating throughput [19].

With the two-way relay scenario shown in Figure 1 and the same setting as in Section 5.2, we compare the actual throughput achieved by different schemes, with the practical constellation size limit. Figure 7 shows the total throughput of different schemes on the broadcast channel with respect to the normalized distance, where the largest constellation is 256 QAM. Generally speaking, Figure 7 shows similar trend as Figure 6. But with the limit in constellation size, some differences do occur: (i) FRNC + SC outperforms FRNC, (ii) at a small distance, FRNC and NC + SC have similar performances, and (iii) the difference between FRNC + SC and NC + SC gets larger than that in Figure 6. By the optimal allocation of power between FRNC and SC, the best performance is achieved in FRNC + SC under all distances. When the distance equals 0.30, FRNC + SC reaches the largest throughput gain, 25.8%, against NC + SC. At this distance, FRNC + SC achieves a much larger gain, 74.2%, against NC. The cumulative density function of the throughput at this distance is shown in Figure 8. The superiority of FRNC and FRNC + SC over other schemes is very clear.

Figure 9 shows the effect of constellation size limit. When the largest constellation is constrained to 64 QAM instead of 256 QAM, the performance of both FRNC + SC and FRNC is degraded. But the performance of FRNC + SC is less affected, where the extra-power is used in SC transmission than being wasted in FRNC.

Next the effect of the number of nodes, n, is evaluated. Average SNR of all relay links is fixed at 20 dB. In such
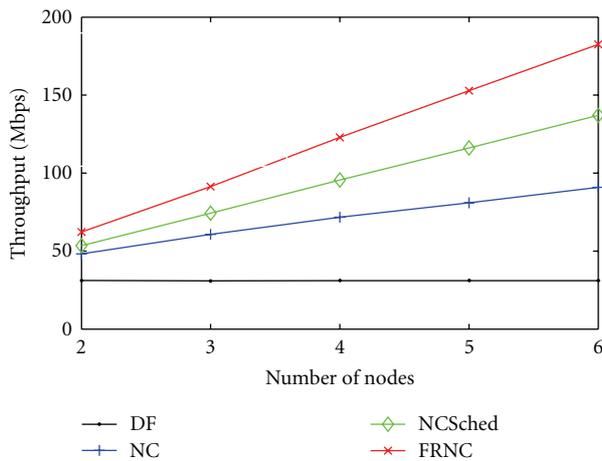
FIGURE 10: Throughput achieved by different schemes on the broadcast channel, effect of the number of nodes. (simulation results, largest constellation is 256 QAM).

scenarios, SC can hardly be used. Therefore, only the schemes without using SC are compared. Figure 10 shows the throughput on the broadcast channel. DF transmits in a TDMA manner. Therefore, it cannot benefit from the increase in nodes and its throughput is almost a constant value. On the other hand, NC, NCSched and FRNC all benefit from the increase in flows more or less. Due to the different capability in handling rate mismatch, the slopes of three curves differ greatly. FRNC always has the highest throughput because the rate mismatch problem is completely solved and full rate is achieved.

7. Conclusions

Recently, network coding is widely studied for improving the relay efficiency in wireless networks. Its performance, however, is greatly limited by factors such as rate mismatch. In this paper, we reinterpreted network coding as a mapping between modulation constellations and extended this mapping to enable simultaneous use of different modulations in network-coded transmissions. In this way, the highest rate over each link can be used and the sum rate can be achieved over the broadcast channel. As a result, the rate mismatch problem is completely solved. The only shortcoming of the proposed scheme is its SNR loss in nesting constellations. This little SNR loss is acceptable if the throughput gain is taken into account. We will further study the effect of the direct link and the potential errors at relay node.

Acknowledgment

References

[1] F. H. P. Fitzek and M. D. Katz, *Cooperation in Wireless Networks: Principles and Applications*, Springer, New York, NY, USA, 1st edition, 2006.

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

[3] P. Larsson, N. Johansson, and K. E. Sunell, "Coded bi-directional relaying," in *Proceedings of the IEEE 63rd Vehicular Technology Conference (VTC '06)*, vol. 2, pp. 851–855, July 2006.

[4] P. Popovski and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," in *Proceedings of the IEEE 63rd Vehicular Technology Conference (VTC '06)*, vol. 2, pp. 588–593, July 2006.

[5] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: physical-layer network coding," in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking (MOBICOM '06)*, pp. 358–365, September 2006.

[6] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *Proceedings of the ACM : Conference on Computer Communications (SIGCOMM '07)*, pp. 397–408, August 2007.

[7] C. Hausl and J. Hagenauer, "Iterative network and channel decoding for the two-way relay channel," in *Proceedings of the IEEE International Conference on Communications (ICC '06)*, vol. 4, pp. 1568–1573, July 2006.

[8] S. Tang, J. Cheng, C. Sun, R. Miura, and S. Obana, "Joint channel and network decoding for XOR-based relay in multi-access channel," *IEICE Transactions on Communications*, vol. E92-B, no. 11, pp. 3470–3477, 2009.

[9] H. Yomo and P. Popovski, "Opportunistic scheduling for wireless network coding," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, Article ID 5089949, pp. 2766–2770, 2009.

[10] P. Larsson, "A multiplicative and constant modulus signal based network coding method applied to CB-Relaying," in *Proceedings of the IEEE 67th Vehicular Technology Conference (VTC '08)*, pp. 61–65, May 2008.

[11] T. Koike-Akino, P. Popovski, and V. Tarokh, "Denoising maps and constellations for wireless network coding in two-way relaying systems," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '08)*, pp. 3790–3794, December 2008.

[12] G. Zeitler, R. Koetter, G. Bauch, and J. Widmer, "On quantizer design for soft values in the multiple-access relay channel," in *Proceedings of the IEEE International Conference on Communications (ICC '09)*, June 2009.

[13] R. Alimi, LI. Li, R. Ramje, H. Viswanathan, and Y. R. Yang, "IPack: in-network packet mixing for high throughput wireless mesh networks," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 66–70, April 2008.

[14] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. IT-29, no. 3, pp. 439–441, 1983.

[15] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," in *Proceedings of the Conference on Computer Communications (SIGCOMM '06)*, pp. 243–254, 2006.

[16] L. Xiao, T. E. Fuja, J. Kliewer, and D. J. Costello Jr., "Nested codes with multiple interpretations," in *Proceedings of the 40th Annual IEEE Conference on Information Sciences and Systems (CISS '06)*, pp. 851–856, 2007.

[17] J. Hagenauer, L. Papke, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 429–445, 1996.

[18] A. Goldsmith, *Wireless Communications*, Cambridge University Press, New York, NY, USA, 2005.

[19] IEEE Computer Society LAN MAN Standards Committee, "LAN Medium Access Protocol (MAC) and Physical Layer (PHY) Specification," IEEE Std 802.11-2007, IEEE, 2007.