10 Network coding in relay-based networks

Hong Xu and Baochun Li

10.1 Introduction

Since its inception in information theory, network coding has attracted a significant amount of research attention in recent years. Shortly after theoretical explorations in wired networks, the use of network coding in wireless networks towards improving throughput has been widely recognized. In this chapter, we present a survey of recent advances in relay-based cellular networks with network coding. We begin with an introduction of network coding theory with a focus on wireless networks. We discuss various network coded cooperation schemes that apply network coding on digital bits of packets or channel codes, in terms of, for example, outage probability and diversity-multiplexing tradeoff. We also consider physical-layer network coding that operates on the electromagnetic waves and its application in relay-based networks. Then we take a networking perspective, and present in details some scheduling and resource allocation algorithms to improve throughput using network coding in relay-based networks with a cross-layer perspective. Finally, we conclude the chapter with an outlook into the future.

Network coding is first proposed in [1] for noiseless wireline communication networks to achieve the multicast capacity of the underlying network graph. The essential idea of network coding is to allow coding capability at network nodes (routers, relays, etc.) in exchange for capacity gain, i.e., an alternative tradeoff between computation and communication. This can be understood by the classic "butterfly" network example. In Figure. 10.1, suppose the source S wants to multicast two bits a and b to two sinks $D1$ and $D2$ simultaneously. Each of the links in the network is assumed to have a unit capacity of 1 bit per time slot (bps). With traditional routing, each relay node between S and the two sinks simply forwards a copy of what it receives. It is then impossible to achieve the theoretical multicast capacity of 2 bps for both sinks, since the thick link in the middle can only transmit either a or b at a time. However, with network coding, the intermediate relay node (darkened in the figure) can perform *coding*, in this case a bitwise exclusive-or operation, upon the two information bits and generate $a + b$ to multicast towards its two outgoing links. $D1$ receives a and $a + b$, and recovers b as $b = a + (a + b)$. Similarly, $D2$ receives b and $a + b$ and

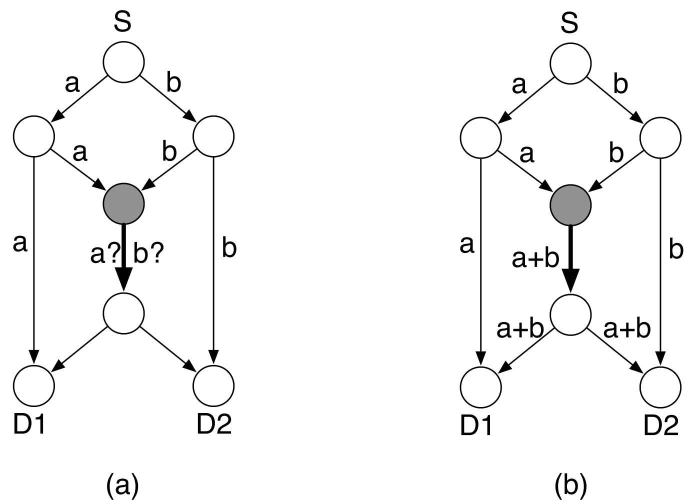


Figure 10.1: The "butterfly" network example of network coding. (a) With traditional routing, the link in the middle can only transmit either a or b at a time. (b) With network coding, the relay node can mix the bits together and transmit $a + b$ to achieve the multicast capacity of 2 bps.

can recover a. Both sinks are therefore able to receive at 2 bps, achieving the multicast capacity.

In the above example, the network coding operation is bitwise exclusive-or, which can be viewed as linear coding over the finite field $GF(2)$. Following the seminal work of [1], Li *et al.* [2] showed that a linear coding mechanism suffices to achieve the multicast capacity. Ho *et al.* [3, 4] further proposed a distributed random linear network coding approach, in which nodes independently and randomly generate linear coefficients from a finite field to apply over input symbols without *a priori* knowledge of the network topology. They proved that receivers are able to decode with high probability provided that the field size is sufficiently large. These works lay down a solid foundation for the practical use of network coding in a diverse set of applications.

After the initial theoretical studies in wireline networks, the applicability and advantages of network coding in wireless networks were soon identified and investigated extensively [5]. Though the noiseless assumption no longer holds for wireless communications, the wireless medium does provide a unique characteristic conducive for network coding operations — the inherent broadcasting capability. Again this can be best understood by another classic example of the "Alice and Bob" topology as in Figure 10.2. Assume that Alice and Bob want to exchange their information represented by bits a and b, respectively, and each link has a static capacity of 1 bps. It can be readily seen that under the traditional routing paradigm, four time slots are needed to exchange the bits through relay R, which sequentially forwards one bit at a time. On the contrary, with network coding, R

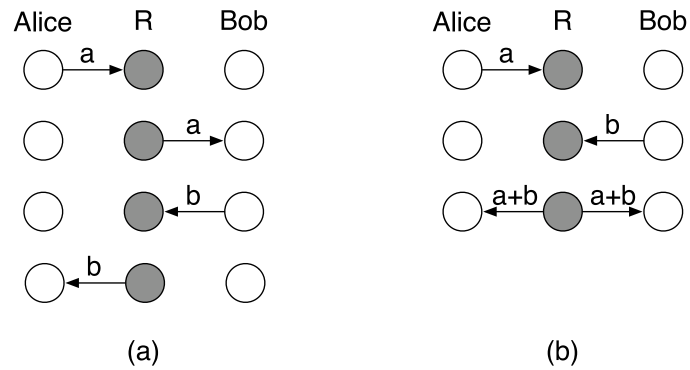


Figure 10.2: The benefits of network coding in wireless networks. (a) Four time slots are needed for Alice and Bob to exchange information bits through relay R by plain routing. (b) With network coding, R can XOR the bits and broadcast the coded bit to both parties simultaneously, reducing it to three time slots.

can XOR the two bits together and transmit the coded bit. Because of the broadcast nature of wireless medium, this transmission can be heard by both Alice and Bob. Alice then receives $a + b$, and recovers Bob's bit b as $b = a + (a + b)$. Similarly Bob can recover a. Therefore only three time slots are needed in this case, which represents a 25% throughput improvement for both parties.

Inspired by the mathematical simplicity and practical potential of network coding, the communications and networking communities have devoted a significant amount of research efforts to utilize it in a number of wireless applications, ranging from opportunistic routing in mesh networks to distributed storage and link inference in sensor networks. Our focus in this chapter is on relay-based networks, which essentially generalizes the example above in various ways depending on the network model. Here, the relays can refer to dedicated stations solely providing traffic relaying for others' benefits, or to users relaying one another's signal, i.e., user cooperation. The purpose of relaying can be merely to extend the coverage of a cellular network when the base station is too distant from the mobile station (multi-hopping), or to combat fading by providing additional cooperative diversity with more advanced receiver hardware. Our discussion will assume that cooperative diversity is exploited whenever possible, i.e., whenever the receiver receives multiple transmissions that contain the same data, and relaying amounts to mere multi-hopping only when the receiver obtains one copy of the data.

10.2 Network coded cooperation

In both wireline and wireless networks, network coding usually operates over bits, or *symbols*. The general use of this conventional form of network coding in cooperative diversity is usually termed *network coded cooperation* in the lit-

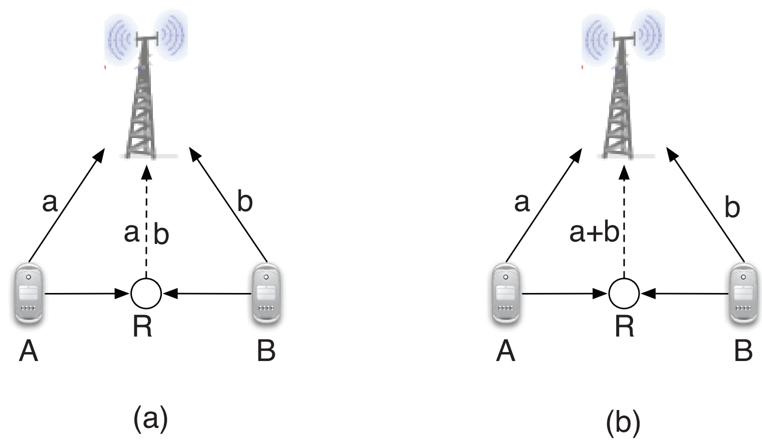


Figure 10.3: (a) Plain relaying without network coding. (b) Network coded cooperation with one dedicated relay.

erature. As was originally invented for wireline networks, network coding can work at or above MAC layer over the data bits of packets when applied to relay-based wireless networks, assuming the link layer delivers error-free data. In other words, it serves as an "add-on" component to the existing lower layer techniques. Alternatively, it can also be applied at the link layer, by a more dedicated joint design with channel coding/decoding.

10.2.1 Simple network coded cooperation

Consider a simple network model where there are two mobile stations (MS) transmitting on the uplink to the base station. The simplest and naive way to attain cooperative diversity, shown in Figure. 10.3(a), is to replicate the classic three-terminal model in information theory by assigning a relay to users on orthogonal channels (can be achieved by time, frequency or code division). The cooperative transmission progresses in two phases [6]. In the first phase, each MS transmits its own data on orthogonal channels while its relay receives and decodes the data. In the second phase the relay forwards the data to the base station, again on orthogonal channels. The base station receives both the original and relayed signals which results in a diversity order of 2 for each user. If time-division is used to orthogonalize channels, a total of 4 time slots are needed for two MS, 2 for the first phase and 2 for the second phase.

Chen *et al.* [7] are among the first to point out that same diversity gain can be achieved with less spectrum cost by network coding as shown in Figure 10.3(b), where the relay assists both users at the same time by transmitting the XOR-ed version of information from both users in the second phase. If any two of the three transmissions succeed, the base station can still recover both a and b, therefore the diversity order is 2 for both users. By a probabilistic analysis, [7] shows that the above network coded cooperation also provides a lower system outage probability at high SNR if the total power consumption for the system is fixed, since only 3 transmissions are required for a complete round of cooperative transmission.

The above network coded cooperation scheme has the following basic intuition. It allows the relay node to first combine information overheard from multiple sources with linear network coding, and then forward the coded data to the destination. This provides diversity for multiple sources with better spectral efficiency because the relaying bandwidth is suppressed. Though the analysis in [7] is fairly preliminary, this key idea demonstrates the relevance of network coding to cooperative diversity, and is largely followed in the community.

Now consider a more general model of a cellular network with $N \geq 2$ users and $M \geq 1$ relays communicating to the base station. Following [7], Peng *et al.* [8, 9] proposed an extended network coded cooperation for multi-user networks. Although their scheme was designed for multiple unicast sessions with N distinct destinations, it readily applies to a multi-user cellular network as we outlined above with one common destination, the base station. In their scheme, each user still transmits their own data in the first phase on orthogonal channels achieved by time-division. Then in the second phase, a single "best" relay is selected from the M candidates that maximizes the worst instantaneous channel conditions of links from users to the relay and from the relay to the base station, and broadcasts the XOR-ed version of data received from each user to the base station. The idea of using a single best relay instead of all available ones can find its root in *opportunistic relaying* first proposed in [10].

Recall the conventional cooperation protocols mandate that a relay transmission must be coupled with each source transmission, no matter whether the relay transmission uses distributed space-time coding across multiple relays or opportunistically utilizes a single best relay [10]. Thus with time-division, a total of $2N$ time slots are needed with N time slots in the first phase and N time slots in the second phase for N users. On the contrary, the network coded cooperation in [9] only takes one time slot in the second phase to broadcast the XOR-ed message for all N users, and only $N + 1$ time slots are required. Therefore, intuitively, the multiplexing gain of the system can be improved by the proposed scheme. The question is, however, can it also maintain the full diversity gain of $M + 1$ provided by the $M + 1$ possible paths to the base station as the conventional schemes?

A comprehensive diversity-multiplexing tradeoff analysis of the selection-based network coded cooperation is offered in [9]. Before we present the results, let us recall the definition of diversity-multiplexing tradeoff first.

Definition 10.1. *A scheme is said to achieve spatial multiplexing gain R_{norm} and diversity gain D if the data rate R satisfies*

$$\lim_{\rho\to\infty} R(\rho)/\log\rho = R_{norm} \tag{10.1}$$

and the average error probability p_e satisfies

$$-\lim_{\rho\to\infty} \log p_e(\rho)/\log\rho = D \tag{10.2}$$

where ρ denotes the signal-to-noise ratio (SNR).

We can see that essentially it is a fundamental tradeoff between error probability and transmission rate, and has been widely recognized and adopted in the literature.

Specifically, the following theorem is proved in [9]:

Theorem 10.1. *[9] In a single-cell network composed of N users and M relays, the above network coded cooperation scheme achieves a diversity-multiplexing tradeoff given as follows:*

$$D(R_{norm}) = 2\left(1 - \frac{N+1}{N}R_{norm}\right), R_{norm} \in \left(0, \frac{N}{N+1}\right). \tag{10.3}$$

As a comparison, conventional cooperation protocols for multi-user networks (distributed space-time coding or opportunistic relaying) achieve a diversity-multiplexing tradeoff of $D(R_{\mathrm{norm}}) = (M+1)\left(1 - 2R_{\mathrm{norm}}\right)$ for $R_{\mathrm{norm}} \in \left(0, \frac{1}{2}\right)$. Therefore, only a diversity gain of 2 can be achieved at high SNR, while the multiplexing gain can indeed be improved from $\frac{1}{2}$ to $\frac{N}{N+1}$. In other words, the proposed network coded cooperation scheme achieves a different tradeoff.

The intuition why XOR network coding does not help attaining the full diversity gain in this case is that, although the coded message can be potentially helpful for any user, it can only help at most one user provided that all the other $N-1$ users' data is decoded correctly, no matter what the number of relays is. The end-to-end performance of one user is bottlenecked by all the other users, attributing to the diversity order of 2[1].

Interestingly, as a special case, when $M = 1$, i.e., only one relay exists in the network, network coded cooperation achieves the full diversity order of 2 and has a strictly better diversity-multiplexing tradeoff than conventional scheme $2\left(1 - 2R_{\mathrm{norm}}\right)$ for $R_{\mathrm{norm}} \in \left(0, \frac{1}{2}\right)$. It entails less loss in spectral efficiency to achieve the same diversity gain, and offers larger diversity order at the same

[1] In the case of multiple unicast sessions, it can be shown that this network coded cooperation has a diversity-multiplexing tradeoff of $D(R_{\mathrm{norm}}) = (M+1)\left(1 - \frac{N+1}{N}R_{\mathrm{norm}}\right), R_{\mathrm{norm}} \in \left(0, \frac{N}{N+1}\right)$, assuming that each distinct destination d_j can reliably overhear messages from other sources $s_j, j \neq i$. This assumption, however, implicitly requires enormous power consumptions during s_j's transmission and are strongly weakened when the network scales. It may not be convincing to demonstrate the superiority of network coded cooperation.

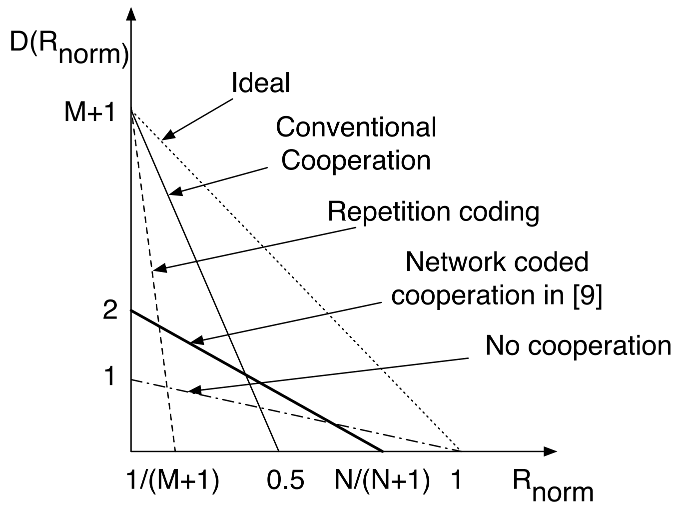


Figure 10.4: Diversity-multiplexing tradeoff comparison.

spectral efficiency in this case. Finally, Figure 10.4 graphically summarizes the above discussion and comparison of the diversity-multiplexing tradeoff for different cooperative diversity schemes.

Another user cooperation protocol named adaptive network coded cooperation (ANCC), that essentially achieves the same diversity-multiplexing tradeoff of conventional cooperation protocols in the multi-user setting, was later proposed in [11]. It also takes $2N$ time slots to complete one round of cooperation for N users. The basic idea is that in the second phase, each user, acting as a relay now, takes turns to select (randomly) a *subset* of messages it correctly overheard, and transmit the binary checksum of them. From a coding perspective, the network coded cooperation can be viewed as matching network-on-graph, i.e., instantaneous network topologies described in graphs, with the well-known class of code-on-graph, i.e., LDPC codes, on the fly. Therefore, the term *adaptive*. Analysis on the achievable rate and outage rate coupled with numerical evaluations show that ANCC outperforms repetition based cooperation and is on par with space-time coded cooperation. This scheme, however, does introduce some level of overhead in transmitting a bitmap so the base station knows how the checksums are formed. Further, decoding also requires an adaptive architecture, and may not be practical in reality.

From the above discussion, we can see that in a multi-user cellular network, the superiority of network coded cooperation mainly lies in its adaptivity to the lossy nature of the wireless media, and its operational simplicity. The per-user complexity, for both schemes in [9, 11], is invariant as the network grows, compared with space-time coded cooperation whose complexity and overhead increases linearly with the network size. We also note that, despite the fact that these proposals are available for combining network coding and cooperative relaying and show the benefits of network coding in some scenarios, further research is required to unleash its full potential and make it practical in more general scenarios.

10.2.2 Joint network and channel coding/decoding

The discussion so far has focused on using network coding upon packets produced by channel decoding. Network coding essentially works as an erasure-correction code that operates above the link layer and offers reliable communication through redundant transmission of packets. Channel coding, on the other hand, is an error-correction technique in wireless environment. It is used in the link layer to recover erroneous bits through appending redundant parity check bits to a packet. By treating them separately, a certain degree of performance loss is introduced. Erasure-correction decoding cannot utilize the redundant information in those packets that fail the channel decoding, and hence are discarded at the link layer, while error-correction decoding cannot take advantage of the additional redundancy provided in the erasure codes.

Indeed, in one early work of network coding [12] it is shown that in general, capacity can only be achieved by a joint treatment of channel and network coding. A number of research efforts have tried to unify the two techniques in order to obtain further performance improvement. We survey some important ones in the sequel.

Let us first revisit the example of Figure 10.3 where two users share one relay on the uplink to the base station, which resembles the multiple-access relay channel in coding and information theory as shown in Figure 10.5(a). Hausl *et al.* [13] first considered joint network-channel decoding in this model, assuming that the relay can reliably decode messages from both users. A distributed regular LDPC code is used for network-channel coding. Different from [7] that separately decodes the two messages transmitted from the users with the network coded message transmitted from the relay, they proposed to jointly decode the three messages on a single Tanner graph [14] with the iterative message-passing algorithm.

More precisely, let $\mathbf{u}_i$ denote the information bits of user i, $i \in \{1, 2\}$, $\mathbf{x}_i$ the network-channel code, and $\mathbf{G}_{ij}$ the generator matrix on the link from i to j. Then the network encoder of [14] at the relay produces[2]

$$\mathbf{x}_3 = \mathbf{u}_1\mathbf{G}_{31} + \mathbf{u}_2\mathbf{G}_{32}. \tag{10.4}$$

[2] Note that the scheme in [7] does not consider interleaving and produces $\mathbf{x}_3 = (\mathbf{u}_1 + \mathbf{u}_2)\mathbf{G}_3$, which is slightly inferior in terms of bit error probability, though a diversity order of 2 is achievable.

The encoding operations at users and the relay can be jointly described as

$$\mathbf{x} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3] = [\mathbf{u}_1 \quad \mathbf{u}_2] \begin{bmatrix} \mathbf{G}_{14} & \mathbf{0} & \mathbf{G}_{31} \\ \mathbf{0} & \mathbf{G}_{24} & \mathbf{G}_{32} \end{bmatrix} = \mathbf{u}\mathbf{G}. \tag{10.5}$$

Then the parity-check matrix $\mathbf{H}$ of the network-channel code has $2N + N_R$ columns and $2(N-K) + N_R$ rows and has to fulfill $\mathbf{G}\mathbf{H}^T = \mathbf{0}$, where N and N_R denote the channel code length and network code length, and K denotes the length of information bits $\mathbf{u}_1, \mathbf{u}_2$. The decoder at the base station decodes the LDPC code with parity-check matrix $\mathbf{H}$ on the Tanner graph and exploits the diversity provided by network coding.

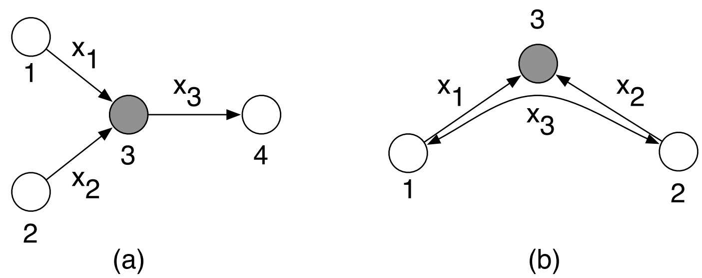


Figure 10.5: (a) A multiple-access relay channel. (b) A two-way relay channel.

[13] showed the diversity and code length gain of network-channel code by numerically comparing it with two references systems, one where the relay is shared without network coding, and another where no relay is employed. The diversity and code length gain is not difficult to intuitively explain. For the comparison to be fair, the number of code bits used in all systems, $2N + N_R$, has to be equal. Consider the setup with $K = 1500$ and $N = N_R = 2000$, which corresponds to a channel code rate of 0.75 in the first phase of cooperation. Network-channel code achieves the full diversity order of 2 since $\mathbf{x}_4$ contains both $\mathbf{u}_1$ and $\mathbf{u}_2$. For the system without network coding, the relay is shared and only transmits 1000 code bits for each user, while there are 1500 information bits. Thus it is impossible to achieve full diversity. Now consider another setup with $K = 1500$ and $N = N_R = 4000$. Without network coding, the relay transmits 2000 code bits for each user and full diversity order can be achieved, with a channel code rate of 0.75 to the base station in the relaying phase. The network-channel code, in this case, provides a more robust code rate of 0.375 to the base station for each user.

[13] did not compare the LDPC-based network-channel coding with separate network and channel coding. [15] followed the same idea and proposed turbo code based network-channel coding for the multiple-access relay channel. Moreover, it was shown that even though full diversity order can be obtained with separate network channel coding, joint network-channel coding based on turbo codes is

able to exploit the additional redundancy in the relay transmission and exhibits strictly better bit error rate (BER) and frame error rate (FER) performance.

Hausl further extended joint network-channel coding to a similar model, the two-way relay channel [16] as in Figure 10.5(b). It is essentially an abstraction of the "Alice and Bob" example shown in Figure 10.2. Note that simple network coded cooperation proposed for multiple-access relay channel as in [7] works without any change for two-way relay channel, and full diversity order can be exploited. However, joint network-channel coding does call for a different design. The subtle difference is that in the two-way relay channel, each user exchanges information through a relay in the middle and decodes each other's message respectively, whereas in the multiple-access relay channel a common destination decodes both messages from each user. An application of the two-way relay channel is the uplink and downlink transmissions between a mobile station and a base station with a relay. Thus the joint network-channel decoder at the mobile station takes as inputs the channel code from BS and the network-channel code from the relay, and jointly decodes them to obtain the data from BS. Its own data is utilized during the decoding process, since the network-channel encoder at the relay is a convolutional encoder coding the interleaved bits of both MS and BS together. The network-channel decoder at the BS works in a similar way. Again, joint network-channel decoding is able to obtain full diversity order of 2 here with better BER performance.

The proposals of [13, 15, 16] rely on the key assumption that the relay is able to reliably decode the messages from both ends, and utilize the simplest XOR network coding. Yang *et al.* and Kang *et al.* broke this assumption and designed iterative network and channel decoding when the relay cannot perfectly recover packets in [17, 18]. Guo *et al.* took it one step further and proposed a practical scheme, called non-binary joint network-channel decoding, that couples non-binary LDPC channel coding and random linear network coding in a high order Galois field in [19]. A joint network-channel coding scheme was also proposed in [20] for user cooperation that endows users with efficient use of resources by transmitting the algebraic superposition of their locally generated information and relayed information that originated at the other user.

Finally, we offer an information theoretic perspective regarding joint network-channel coding on the two-way relay channel in Figure 10.5(b) before we end the discussion. The second phase of transmission involves the relay broadcasting to two receivers at each end, and can be seen as the well-studied broadcast channel. Joint network-channel coding, as we discussed above, performs network coding across channel codewords, i.e.,

$$\mathbf{x}_3 = \pi_{32}(\mathbf{u}_1) + \pi_{31}(\mathbf{u}_2) \tag{10.6}$$

where $\pi_{ij}(\cdot)$ denotes the channel encoding function for link ij. In general, the capacity-achieving code rate varies for different links, and by coding channel codewords the capacity of each link involved in the broadcast is achieved as in the single-link transmission[3]. The only difference here is that for user i, certain bits in the channel code are flipped with a known pattern $\pi_{3j}(\mathbf{u}_i)$ at the relay as in Eq. (10.6), and they are flipped back after demodulation at i. If separate network channel coding is used, $\mathbf{x}_3 = \pi_3(\mathbf{u}_1 + \mathbf{u}_2)$. Readily we can see that the rate of π_3 has to be confined to the minimum rate of the two links 31 and 32 for both users to be able to decode, resulting in performance loss in terms of achieving the broadcast channel capacity. This is referred to as the *coding to the worst rate* problem. It is easy to verify that this problem does not exist in the multiple-access relay channel.

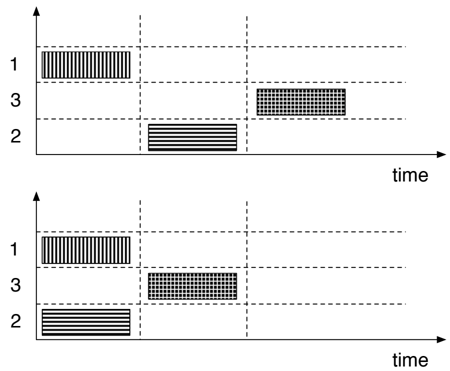
10.3 Physical-layer network coding

The main distinguishing characteristic of wireless communications is its broadcast nature. The radio signal transmitted from one node is often overheard by many neighboring nodes, and causes interference to their receptions. Conventional wisdom treats interference as a nuisance and strives to avoid it by making the transmissions orthogonal to each other in time, frequency or by code, which in effects disguises a wireless link into a wired one. In the preceding discussions of network coded cooperation, network coding takes advantage of the broadcast nature without changing this fundamental interference-avoiding structure of wireless transmissions.

Opposite to this line of thinking, a novel paradigm that embraces interference as a unique capacity-boosting advantage was recently developed in [21] and is gaining momentum. The idea is to encourage interference from concurrent transmissions, and by smart physical layer techniques transform the superposition of the electromagnetic waves as an equivalent network coding operation that mixes the radio signals in the air. An apparatus of network coding is then created at the physical layer, and works on the EM waves in the air rather than on the digital bits of data packets or channel codes. Therefore, it is termed *physical-layer network coding* (PNC), or *analog network coding*, while the term *digital network coding* refers to its conventional counterpart.

More specifically, Figure 10.6(a) illustrates the working and power of physical-layer network coding, on the familiar two-way relay channel. Using PNC, only 2 time slots are needed for each user to exchange information, as opposed to 3 using digital network coding and 4 using direct transmission. In the first time slot each user *concurrently* transmits to the relay. For simplicity, assume that BPSK modulation is used, and symbol-level and carrier phase synchronization

[3] Note that all works we described, including [13, 15, 17, 18, 19, 20] that are not designed for the two-way relay channel, assume that the code rate is the same for all channel codes involved on all links to ease the decoding design. These proposals will work for the more general case where code rate is different, with proper modification, mostly on the decoding algorithm.



(a)

Modulation mapping at sources		Demodulation mapping at relay	
Data	Signal	Signal	Data
0	-1	0	1
1	1	2,-2	0

(b)

Figure 10.6: (a) The intuitive benefits of physical-layer network coding. The graph above shows the transmission schedule using digital network coding, and graph below shows that of physical-layer network coding. (b) BPSK modulation/demodulation mapping for physical-layer network coding.

and channel pre-equalization is perfectly done, so that the frames from users arrive at relay with same amplitude and phase. Then the baseband signal received by relay is

$$r(t) = S_1(t) + S_2(t) = (a_1 + a_2)\cos\omega t \tag{10.7}$$

where $S_i(t)$ is the baseband signal from user i. The relay does not decode both BPSK signals a_1 and a_2 from $r(t)$. Instead it tries to decode and transmit the signal $a_1 + a_2$. The original BPSK scheme has only two signals, 1 and -1 corresponding to data bits of 1 and 0, respectively. However $r(t)$ has three signals, -2, 0, and 2. A modulation/demodulation mapping shown in Figure 10.6(b) is developed for BPSK in [21] so that superposition of the analog signals $a_1 + a_2$ can be mapped to arithmetic addition of digital bits they represent in $GF(2)$.

In theory, physical-layer network coding has the potential to greatly improve throughput performance (by 100% compared to direct transmissions and by 33.3% compared to digital network coding as in Figure 10.6(a)). This, however, does come with a non-negligible price — the loss of diversity gain. Recall that the full diversity order of 2 is achievable in two-way relay channel for conventional and network coded cooperation as discussed in Section 10.2.2. Since PNC entails concurrent transmissions in the first time slot, both users only receive one transmission in the second slot when the relay broadcasts[4], whereas they receive from a direct transmission and a relay transmission in conventional and network coded cooperation. Hence, cooperative diversity cannot be exploited and only a diversity order of 1 can be obtained with PNC, and relaying here boils down to simple multi-hopping. For this reason, to our knowledge almost all

[4] The standard assumption of half-duplex radio hardware is assumed.

of the follow-up works compare PNC with multi-hopping and draw conclusions without considering cooperative diversity, even if possible. Hence unless otherwise specified, our discussion of PNC hereafter inherits this assumption from the literature, the validity and fairness of which, however, remains disputable and are left to the judgment of readers.

Given its promising potential, the idea of physical-layer network coding has generated considerable amount of interests in the community. Hao *et al.* first analyzed the achievable rates of PNC, assuming the additive white Gaussian noise (AWGN) channel model in [22]. Through numerical analysis it was shown that PNC outperforms digital network coding and direct transmission significantly, and approaches the capacity limit of the two-way relay channel with appropriate modulation schemes. This is not unexpected as illustrated in the aforementioned example in Figure 10.6(a). Several estimation techniques were developed for PNC to deal with noise in the decoding process in [23]. The rigid and unpractical requirements of perfect synchronization and channel pre-equalization then becomes the focal point of critics about PNC, as the ill effect of imperfect synchronization can be substantial (6 dB loss as shown in [22]).

The topic of general PNC schemes that relax the synchronization requirement while preserving the performance superiority is actively pursued. An amplify-and-forward scheme, in which the relay directly amplifies and forwards the interfered signal to both users as opposed to the decode-and-forward strategy in [21], was proposed and studied in several works [22, 24, 25, 26, 27]. It is generally found that amplify-and-forward PNC is more robust and offers better performance when synchronization is absent compared to decode-and-forward PNC [22], whereas when perfect synchronization is provided amplify-and-forward PNC suffers from a loss of optimality in terms of achievable rate [25].

The robustness of amplify-and-forward PNC is more pronounced for general scenarios because no channel pre-equalization is required. [26, 27] considered the case when there are M relays available for the two-way relay channel. To effectively utilize multi-user diversity offered by relays, a distributed relay selection strategy was proposed with a selection criteria specifically designed for amplify-and-forward PNC. Two information theoretic metrics, outage and ergodic capacities, closely related to the diversity and multiplexing gains of the system, were analytically and numerically evaluated to confirm its advantage. Indeed, the following can be proved as in [27]:

Theorem 10.2. *[27] The amplify-and-forward PNC scheme in [27] achieves a diversity-multiplexing tradeoff of the following:*

$$D(R_{norm}) = M(1 - R_{norm}), R_{norm} \in (0, 1]. \tag{10.8}$$

Clearly, multi-user diversity is capitalized as the diversity order is M. Moreover, the same multiplexing gain of 1 as direct transmission can also be attained, echoing our remark on the benefit of network coding in suppressing bandwidth for relaying in Section 10.2.1. On the other hand, decode-and-forward PNC is not able to harvest multi-user diversity gain, due to the operational requirement of channel pre-equalization before decoding in order to make sure both signals are received with equal amplitude. One immediately realizes that, despite the technical difficulty, channel pre-equalization effectively inverses the fading effect of the channel, and results in loss of diversity gain.

Joint network-channel coding/decoding is also explored in the area of PNC. The main objective is to design a good coding/decoding scheme that maximizes the amount of information that can be reliably exchanged through the two-way relay channel. More advanced coding that applies lattice code on the relay was discussed in [28, 29].

We make a few comments about open issues and future directions of physical-layer network coding. First, several theoretic issues remain unsolved or unexplored, as the research in this area is still in its nascent stage. Most importantly, only the simplest XOR network coding on $GF(2)$ is realized with signal superposition in the air, while linear network coding can work on a much larger Galois field. Note that in wireless communications, path-loss and channel fading effectively couple to the transmitted signal a complex coefficient, which, to some extent, resembles the multiplication operation in a finite field of complex numbers, while signal superposition resembles the addition operation. It is therefore interesting, but certainly non-trivial, to investigate whether the complex channel coefficient can be "transformed" to an equivalent network coding coefficient to realize even larger performance improvement.

Besides, all the developed schemes remain fairly theoretical, while very few attempted to evaluate the true performance of PNC in a real hardware implementation. To the best of our knowledge, [24] is the only work that sketches the practical coding and decoding algorithms for MSK modulation based on amplify-and-forward PNC, and implements the system using software radios. It is asserted that PNC is indeed practical, and achieves significantly higher throughput than digital network coding. While this conveys an optimistic message, one realizes that physical-layer network coding marks a significant departure from the conventional wisdom, and tremendous efforts have to be made across layers of the protocol stack which are designed to avoid interference before it can be implemented in general scenarios.

10.4 Scheduling and resource allocation: Cross-layer issues

The discussion so far has largely focused on the lower layers of the protocol stack. From communications and information theory point of view, numerous schemes are developed based on network coding to utilize a given budget of resources to transmit information as efficiently as possible. In other words, they answer the question of how to transmit in a given scenario for individual nodes. From a networking point of view, provided with the lower-layer transmission protocol,

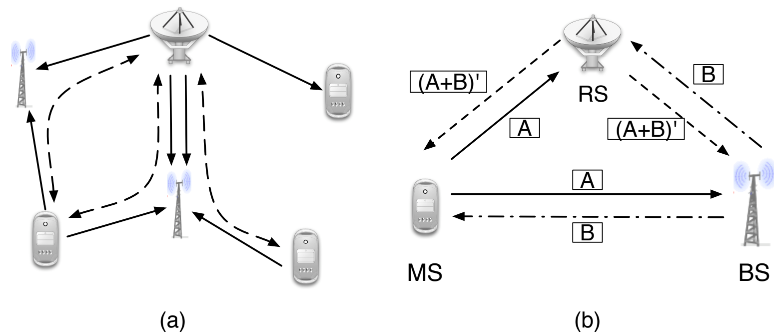


Figure 10.7: (a) A cellular network with multiple relays. Solid lines denote direct transmissions while dotted ones denote relay transmissions. All transmissions happen on orthogonal OFDM subchannels. Note that one MS can be paired up with multiple relays, while one relay could help multiple MS, complicating the scheduling and resource allocation. (b) XOR-assisted cooperative diversity as adapted from [31]. Here different line types denote different subchannels.

upper-layer protocols have to coordinate the transmissions between nodes in the network, and allocate network-wise (such as channels) and node-wise (such as power) resources adequately to these competing sessions so as to optimize some network-wise metrics. In other words, we need to create scenarios amenable for network coding aided relay transmission protocols such that performance is maximized from a network perspective.

The use of network coding certainly calls for new designs of cross-layer scheduling and resource allocation protocols on existing cellular network architectures. To this end, there exist only a few publications in the literature. Zhang *et al.* [30] are arguably the first to venture into this area. They considered the use of simple network coded cooperation on the two-way relay channel, with the base station serving as the relay to XOR the incoming packets and broadcast to two users. They assumed an OFDMA-based cellular network model, and developed a coding aware dynamic subcarrier assignment heuristic in a frequency-selective multipath fading environment. The intuition is that in such an environment, different OFDM subcarriers have independent channel gains to the same mobile station (multi-channel diversity), and even the same subcarrier fades differently on different mobile stations (multi-user diversity). By dynamically matching the subcarriers to the best mobile stations for network coded cooperation while taking fairness into account, a substantial throughput improvement was reported in [30].

Another work of Xu *et al.* [31] represents an in-depth investigation in this area, which is discussed in details here. Consider a more general scenario in Figure 10.7(a). Assume multiple relays are available for the OFDMA based network, and all transmissions happen on orthogonal OFDM subchannels. An XOR-assisted cooperative diversity scheme, named *XOR-CD*, is employed by replicating the two-way relay channel using bi-directional traffic of a given mobile station as in Figure. 10.7(b). Specifically, it works as follows. In the first phase of cooperation, on the uplink MS sends packet A to BS and on the downlink BS sends packet B to MS, simultaneously on orthogonal subchannels. Transmissions are overheard by the relay. At the second phase, relay broadcasts $A + B$ using another orthogonal subchannel. Readily we can see that cooperative diversity can still be exploited as in conventional schemes, which requires 4 subchannels and more power to complete the same job.

Extending to the network scale, while XOR-CD has promising potentials, it seems prohibitive to design good scheduling and resource allocation algorithms for the following reasons. First, one mobile station can be paired up with multiple relays and one relay could help multiple MS, as can be illustrated in Figure 10.7(a). Second, for one mobile station, it can utilize direct transmission, conventional cooperative diversity, and XOR-CD at the same time on different subchannels, depending on the dynamic channel conditions and resource availability. Third, the three dimensions of the problem, namely relay assignment, relay-strategy selection, and subchannel assignment, intricately interplay with each other, further aggravating the problem. Towards this end, the contributions of [31] is two-fold. First, a unifying optimization framework is developed that jointly considers relay assignment, relay-strategy selection, and subchannel assignment for both MS and relays. Second, the joint optimization problem, referred to as RSS-XOR problem, is shown to be NP-hard, and an efficient approximation algorithm is proposed based on set packing with a provable approximation ratio.

Specifically, the following is shown to hold:

Theorem 10.3. *[31] The RSS-XOR problem is equivalent to a maximum weighted 3-set packing problem, and is NP-hard.*

Proof. Construct a collection of sets C from a base set $\zeta \cup \psi$ as shown in Figure 10.8, where ζ is the set of data subchannels and ψ the set of relay subchannels. As we can see, there are three kinds of sets, representing three possible transmission modes, respectively. $(c_i), c_i \in \zeta$ represents the direct transmission with data subchannel c_i. $(c_i, c_r), c_i \in \zeta, c_r \in \psi$ corresponds to the conventional cooperative diversity with data subchannel c_i and relay subchannel c_r. Finally, $(c_i, c_j, c_r), c_i, c_j \in \zeta, c_r \in \psi$ corresponds to XOR-CD with data subchannel pair (c_i, c_j) and relay subchannel c_r. Sets intersect if they share at least one common channel, and are otherwise said to be disjoint. Each set has a corresponding weight, denoting the maximum marginal utility found across all possible assignments of this set to different combinations of relays and links[5]. The utility function is defined such that proportional fairness is taken into account. For

[5] One MS has two links, namely the uplink and downlink.

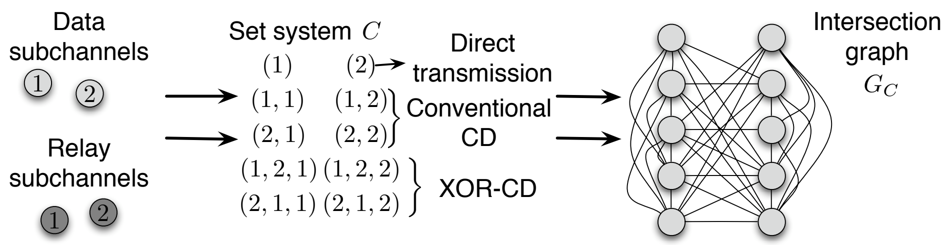


Figure 10.8: Set construction and transformation into an intersection graph with 2 data subchannels and 2 relay subchannels. Vertices in G_C correspond to sets in C. Edges are added between vertices whose corresponding sets intersect.

(c_i, c_j, c_r), its weight is found over all possible assignments of this set to combinations of relays and both links of a MS.

The RSS-XOR problem is to find the optimal strategy to choose the transmission mode and assign relays and channels to each link in order to maximize the aggregated throughput. The maximization is done across all links. Equivalently, we can also interpret it as to find the optimal strategy to select disjoint channel combinations, and assign relays and links to them so as to maximize the objective. In this alternative interpretation, the maximization is done over all possible channel sets by matching them to the best possible links and relays, without violating the obvious constraint that each channel can only be used once. The number of elements in a set is at most 3; therefore, this problem is essentially a weighted 3-set packing problem [32], which is NP-complete. □

To propose an approximation algorithm, we first construct an intersection graph G_C, such that the a vertex in the vertex set V_C corresponds to a set in the set system C, and there is an edge between two vertices in G_C if the two corresponding sets intersect, as shown in Figure 10.8. Weighted set packing can then be generalized as a weighted independent set problem, the objective of which is to find a maximum weighted subset of mutually non-adjacent vertices in G_C [33]. The size of sets is at most 3, therefore G_C is 3-claw free. Here, a d-claw c is an induced subgraph that consists of an independent set T_c of d nodes. The best known approximation algorithm for the weighted independent set problem in claw-free graph is proposed in [33] and then acknowledged in [32], which is then adopted in [31] as the solution algorithm with an approximation ratio of $\frac{2}{3}$ as proved in [33].

As a side note, it was also shown in their work [31] that the joint optimization problem that involve conventional cooperative diversity and direction transmission only can be casted into a weighted bipartite matching problem, which admits polynomial-time algorithms to obtain the optimal solution. This demonstrates an interesting point that, although network coding provides significant performance gains, it may render the scheduling and resource allocation problems even more involved at least in some cases. Moreover, this also reflects the importance of developing cross-layer protocols to fully realize the benefits of network coding in relay-based networks.

While [30, 31] attempted to address these upper-layer issues, the algorithms developed are still impractical to be implemented in real-world cellular networks. Full channel-side information (CSI) is assumed to be available at the base station, which may not be the case for fast fading environments. The complexity of the algorithms is high for real-time scheduling at the time scale of 5-10 ms, which is the common frame length. Therefore, substantial efforts in design, implementation and evaluation of practical protocols are imperative to further understand and conquer the challenges network coding brings to relay-based networks.

10.5 Conclusion

Network coding, by allowing network nodes to mix information flows, represents a paradigm shift for communication networks. With the broadcast nature of wireless medium, it can be naturally applied to wireless communications and bring promising performance improvement in relay-based cellular networks. The question is, what is the most efficient way to utilize network coding in relay-based networks, and how practical is it to be adopted in the real world?

In this chapter, we have started with network coded cooperation that applies network coding on digital bits of information, both without and with a joint consideration of channel codes. We noted the operational simplicity and multiplexing gain of network coded cooperation, and showed that in general a joint design of network and channel coding/decoding is required to achieve the information theoretic capacity, at a cost of increased decoding complexity. We then extended the discussion to physical-layer network coding, a radical yet interesting idea that treats the signal superposition in the air as the exclusive-or network coding operation. By embracing interference, physical-layer network coding has the potential to dramatically improve the throughput performance of relay-based networks.

From communications and information theory perspective, the overall message appears to be very optimistic. From a networking perspective, although network coding is mostly applied in the lower layers of the protocol stack, upper layers have to be *coding-aware* and perform scheduling and resource allocation accordingly. In other words, cross-layer efforts are needed to materialize network coding in a practical network setting. Towards this end, we have presented some initial cross-layer studies on cellular relay networks. An important lesson is that, network coding may render the scheduling and resource allocation more complicated, at least in some cases, because it mixes information flows from different sessions.

Despite the significant amount of research, the use of network coding in relay-based cellular networks remains largely theoretical. Substantial efforts on implementation and evaluation of network coding aided transmissions in large-scale relay networks is imperative at this stage. With the proliferation of wireless technologies and devices, and the ever-growing bandwidth demand from data-intensive mobile applications, we envision that network coding could become a practical and important component in the forth-coming wireless technologies, and play an important role in the quest of high throughput and ubiquitous connectivity in wireless communications.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, **46**:4 (2000), pp. 1204–1216.

[2] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, **49**:2 (2003), pp. 371–381.

[3] T. Ho, M. Médard, J. Shi, M. Effros, and D. R. Karger, "On randomized network coding," in *Proc. 41st Allerton Conference on Communications, Control, and Computing*, 2003.

[4] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inform. Theory*, **52**:10 (2006), pp. 4413–4430.

[5] Y. Wu, P. A. Chou, and S.-Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *Proc. 39th Annual Conference on Information Sciences and Systems (CISS)*, 2005.

[6] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, **50**:12 (2004), pp. 3062–3080.

[7] Y. Chen, S. Kishore, and J. Li, "Wireless diversity through network coding," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2006.

[8] C. Peng, Q. Zhang, M. Zhao, and Y. Yao, "On the performance analysis of network-coded cooperation in wireless networks," in *Proc. IEEE INFOCOM*, 2007.

[9] C. Peng, Q. Zhang, M. Zhao, Y. Yao, and W. Jia, "On the performance analysis of network-coded cooperation in wireless networks," *IEEE Trans. Wireless Commun.*, **7**:8 (2008), pp. 3090–3097.

[10] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, **24**:3 (2006), pp. 659–672.

[11] X. Bao and J. Li, "Adaptive network coded cooperation (ANCC) for wireless relay networks: Matching code-on-graph with network-on-graph," *IEEE Trans. Wireless Commun.*, **7**:2 (2008), pp. 574–583.

[12] M. Effros, M. Médard, T. Ho, S. Ray, D. Karger, and R. Koetter, "Linear network codes: A unified framework for source, channel and network coding," in *Proc. DIMACS Workshop on Network Information Theory*, 2003.

[13] C. Hausl, F. Schreckenbach, I. Oikonomidis, and G. Bauch, "Iterative channel and network decoding on a Tanner graph," in *Proc. 43rd Allerton Conference on Communications, Control, and Computing*, 2005.

[14] R. M. Tanner, "A recursive approach to low-complexity codes," *IEEE Trans. Inform. Theory*, **27**:5 (1981), pp. 533–547.

[15] C. Hausl and P. Dupraz, "Joint network-channel coding for the multiple-access relay channel," in *Proc. International Workshop on Wireless Ad-hoc and Sensor Networks (IWWAN)*, 2006.

[16] C. Hausl, and J. Hagenauer, "Iterative network and channel decoding for the two-way relay channel," in *Proc. IEEE ICC*, 2006.

[17] S. Yang and R. Koetter, "Network coding over a noisy relay: A belief propagation approach," in *Proc. IEEE Intl. Symposium on Information Theory (ISIT)*, 2007.

[18] J. Kang, B. Zhou, Z. Ding, and S. Lin, "LDPC coding schemes for error control in a multicast network," in *Proc. IEEE Intl. Symposium on Information Theory (ISIT)*, 2008.

[19] Z. Guo, J. Huang, B. Wang, J.-H. Cui, S. Zhou, and P. Willett, "A practical joint network-channel coding scheme for reliable communication in wireless networks," in *Proc. ACM MobiHoc*, 2009.

[20] L. Xiao, T. E. Fuja, J. Kliewer, and D. J. Costello, Jr., "A network coding approach to cooperative diversity," *IEEE Trans. Inform. Theory*, **53**:10 (2007), pp. 3714–3722.

[21] S. Zhang, S. C. Liew, and P. Lam, "Hot topic: Physical-layer network coding," in *Proc. ACM Mobicom*, 2006.

[22] Y. Hao, D. L. Goecket, Z. Ding, D. Towsley, and K. K. Leung, "Achievable rates of physical layer network coding schemes on the exchange channel," in *Proc. IEEE Military Communications Conference (MILCOM)*, 2007.

[23] S. Zhang, S. C. Liew, and L. Lu, "Physical layer network coding schemes over finite and infinite fields," in *Proc. IEEE Globe Communications Conference (Globecom)*, 2008.

[24] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proc. ACM SIGCOMM*, 2007.

[25] P. Popovski and H. Yomo, "Physical network coding in two-way relay channels," in *Proc. IEEE ICC*, 2007.

[26] Z. Ding, K. K. Leung, D. Goeckel, and D. Towsley, "On the capacity of network coding with diversity," in *Proc. First Annual Conference of the International Technology Alliance*, 2008.

[27] ——, "On the study of network coding with diversity," *IEEE Trans. Wireless Commun.*, **8**:3 (2009), pp. 1247–1259.

[28] K. R. Narayanan, M. P. Wilson, and A. Sprintson, "Joint physical layer coding and network coding for bi-directional relaying," in *Proc. 45th Annual Allerton Conference on Communications, Control and Computing*, 2007.

[29] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference with structured codes," in *Proc. IEEE Intl. Symposium on Information Theory (ISIT)*, 2008.

[30] X. Zhang and B. Li, "Joint network coding and subcarrier assignment in OFDMA-based wireless networks," in *Proc. 4th Intl. Workshop on Network Coding, Theory and Applications (NetCod)*, 2008.

[31] H. Xu and B. Li, "XOR-assisted cooperative diversity in OFDMA wireless networks: Optimization framework and approximation algorithms," in *Proc. IEEE INFOCOM*, 2009.

[32] B. Chandra and M. M. Halldórsson, "Greedy local improvement and weighted set packing approximation," *Journal of Algorithms*, **29**:2 (2001), pp. 223–240.

[33] P. Berman, "A d/2 approximation for maximum weight independent set in d-claw free graphs," *Nordic Journal of Computing*, **7**:3 (2000), pp. 178–184.