

Resource Allocation and QoS Provisioning for Wireless Relay Networks

Long Bao Le¹, Sergiy A. Vorobyov², Khoa T. Phan³, and Tho Le-Ngoc⁴

¹ Department of Aeronautics and Astronautics, Massachusetts Institute of Technology

² Department of Electrical and Computer Engineering, University of Alberta

³ Department of Electrical Engineering, California Institute of Technology

⁴ Department of Electrical and Computer Engineering, McGill University

1.1 Introduction

Emerging broadband wireless applications in most wireless networks require increasingly high throughput and more stringent quality-of-service (QoS) requirements. In this respect, multiple-antenna technologies have been recognized as important solutions for future high-speed wireless networks [57], [58], [59]. Particularly, employment of multiple antennas at transmitter and/or receiver sides can provide significant multiplexing and/or diversity gains [66]. The net effects of these gains are the improvements in terms of wireless link robustness (i.e., lower bit error rate (BER)) and network capacity. Unfortunately, the implementation of multiple antennas in most modern mobile devices may be challenging due to their small sizes.

Cooperative diversity has been proposed as an alternative solution where a virtual antenna array is formed by distributed wireless nodes each with one antenna. Cooperative transmission between a source node and a destination node is performed with assistance of a number of relay nodes. In particular, the source and relay nodes collaboratively transmit information to the destination node [22], [23], [24], [39], [42], [51], [52], [67]. It is intuitive that in order to make cooperative transmission efficient or even possible, the source node has to carefully choose one or several “good” relays and first forward its data to those relays. Then, the source and relays can coordinate their transmissions in such a way that maximum multiplexing/diversity gains can be achieved at the destination node.

Although cooperative diversity is simple in concept, there are many technical issues to be resolved for practical implementation. First, protocol design for cooperative diversity is one of the important research focuses [1], [22], [23], [51], [52]. Second, it is worth noting that most practical cooperative diversity protocols have two phases: in the first phase, the source node broadcasts its message to assisting relays; in the second phase, the relays collaboratively

transmit the received information to the destination. Therefore, cooperative transmission may not be always beneficial or even necessary because direct transmission from the source to the destination node may already be successful. Adaptive cooperative protocols, where nodes cooperate only when necessary and/or they cooperate using incremental transmissions, usually have significantly better performance than “straight-forward” protocols [1], [10], [25], [26], [64]. In addition, emerging technology such as network coding can be employed to design even smarter cooperative protocols [21], [28], [61]. Finally, other important issues such as relay selection, synchronization among relays’ transmissions need to be considered for practical implementation [2], [4], [27], [31], [43], [56], [65].

While most existing works on cooperative diversity in the literature focus on design and performance analysis of cooperative protocols, resource allocation for wireless relay networks receives less attention. However, resource allocation also has significant impacts on system performance [11], [29], [30], [32], [36], [63]. In fact, assisting relays usually have limited radio resources (e.g., bandwidth and power) and they are shared by several source-destination pairs. Therefore, a smart radio resource allocation for wireless relay networks guarantees both fair access to available relays and good overall network throughput performance. In addition, by using a proper relay selection strategy where each source-destination pair only selects one or a small number of good relays, efficient resource utilization can be achieved with low implementation complexity. Finally, distributed resource allocation algorithms are usually required in wireless relay ad hoc networks because there is no central controller in such applications.

In this chapter, we attempt to provide a survey on cooperative protocols, key design issues and resource allocation problems in wireless relay networks. In particular, we describe popular cooperative protocols and their possible extensions and enhancements. We also briefly present typical applications of cooperative communications, namely for multihop cellular and ad hoc networks, and broadcasting applications in ad hoc networks. In addition, we review some existing works on performance analysis and QoS provisioning issues for wireless relay networks. Finally, we introduce a resource allocation framework for single-user and multi-user relay networks including both centralized and decentralized power allocation algorithms.

This chapter is organized as follows. In section 1.2, we describe some fundamentals of cooperative protocols. Section 1.3 overviews some typical applications and research issues. In section 1.4, we discuss performance analysis and QoS provisioning issues in wireless relay networks. An overview of resource allocation via power allocation for both single- and multi-user relay networks is presented in section 1.5. Conclusions are stated in section 1.6.

1.2 Overview of Cooperative Diversity

Cooperative diversity protocols allow a number of users to relay signals for one another in such a way that a diversity gain can be achieved. In fact, information theoretic capacity of such a network setting, named a relay channel, has been investigated a few decades ago [9]. Deep understanding of MIMO systems from both information theoretic and practical system design viewpoints over the past decade has stimulated and attracted significant research efforts in cooperative diversity. In this section, we provide a survey on fundamentals of cooperative diversity.

Consider a source node s communicating to a destination d with the help of m relays, r_1, r_2, \dots, r_m . Let a_{ij} be the channel gain between nodes i and j , and P_i be the transmission power of node i . The signal is corrupted by white Gaussian noise. For simplicity, throughout this section we assume that N is the white Gaussian noise power measured in the signal bandwidth at all nodes. We assume that cooperation among users is performed in phases (i.e. time slots) and users can be synchronized by a common system clock. Fig. 1.1 illustrates a general cooperative diversity protocol where the source broadcasts its message in the first phase and the relays retransmit the message in the second phase. In the following, we describe some popular cooperative diversity protocols and their corresponding performances.

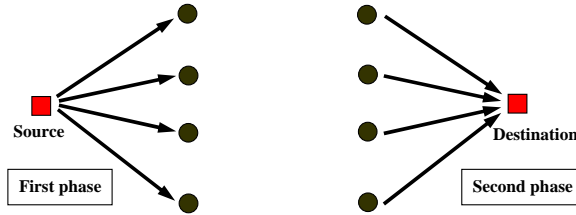


Fig. 1.1. Cooperative protocols

1.2.1 Amplify-and-Forward

In this cooperative protocol, the source broadcasts message x_s in the first phase. The message is received by the destination and relays. Each relay r_i amplifies the received signal in the first phase and transmits to the destination in the second phase. The destination combines the signals received in both phases to decode the message. Specifically, the signal received by relay r_i in the first phase (denoted as y_{r_i}) can be written as

$$y_{r_i} = a_{sr_i}x_s + z_{r_i} \quad (1.1)$$

where a_{sr_i} is the channel gain for link s - r_i and z_{r_i} denotes Gaussian noise at relay r_i . Suppose each relay normalizes the received signal before transmitting to the destination. Then, the transmitted signal can be written as

$$x_{r_i} = g_{r_i} y_{r_i} \quad (1.2)$$

where g_{r_i} is the amplifying gain which is given by

$$g_{r_i} = \sqrt{\frac{P_{r_i}}{|a_{sr_i}|^2 P_s + N}}. \quad (1.3)$$

Assuming that a maximum-ratio-combiner (MRC) is used at the destination, the source-destination capacity of this protocol is given as [65]

$$C_{\text{AF}} = \frac{1}{m+1} \log \left(1 + \text{SNR}_s |a_{sd}|^2 + \sum_{i=1}^m \frac{\text{SNR}_s |a_{sr_i}|^2 \text{SNR}_{r_i} |a_{r_i d}|^2}{\text{SNR}_s |a_{sr_i}|^2 + \text{SNR}_{r_i} |a_{r_i d}|^2 + 1} \right) \quad (1.4)$$

where $\text{SNR}_j = P_j/N$ is the signal-to-noise ratio (SNR) at node $j \in \{s, r_i | i = 1, \dots, m\}$, P_j denotes the power at the source or relay node, N is the noise power, and a_{sd} is the channel gain for link s - d .

Another important performance measure which is extensively used for investigating the performance of different cooperative diversity protocols is the outage probability. In Rayleigh fading channels, the outage probability of the amplify-and-forward (AF) cooperative protocol can be approximated as [65]

$$p_{\text{AF}}^{\text{out}}(\text{SNR}, R) \triangleq \Pr [C_{\text{AF}} < R] \propto \left(\frac{2^{2R} - 1}{\text{SNR}} \right)^{m+1}.$$

where $\text{SNR} = P/N$ and it is assumed that all nodes transmit at power level P . This outage probability shows that AF cooperative protocol achieves diversity order of $m+1$ with m relays.

1.2.2 Decode-and-Forward

For the decode-and-forward (DF) cooperative protocol, relay nodes apply some forms of detection and/or decoding before encoding the information and forwarding it to the destination. Such a cooperative protocol also has two phases (i.e., time slots). In the first phase, the source broadcasts the signal to the relays which subsequently detect and/or decode it. In the second phase, the relays transmit re-encoded signals to the destination using repetition or space-time codes.

For protocols that require relays to fully decode the received signal in the first phase, the set of relays which successfully decode the signal at the end of the first phase is only a subset of all available relays. Let $D(s)$ denote the set of successfully-decoding relays which will be called a decoding set in

the following. For repetition-based coding, the destination receives separate retransmission from each relay $r_i \in D(s)$. Hence, we can write the signal from relay r_i received at the destination d as

$$y_d = a_{r_i d} x_{r_i} + z_d \quad (1.5)$$

where x_{r_i} denotes the signal transmitted by relay node r_i , $a_{r_i d}$ stands for the channel gain for link r_i - d , and z_d denotes the Gaussian noise at the destination. If space-time coding is used, the destination will receive the superimposed signals from all relays $r_i \in D(s)$ simultaneously. Hence, the received signal at the destination in the second phase can be expressed as

$$y_d = \sum_{r_i \in D(s)} a_{r_i d} x_{r_i} + z_d. \quad (1.6)$$

It has been shown in [23] that both repetition-based or space-time-coding-based DF protocols achieve full diversity order of $m+1$ in the low rate regime. This diversity gain has been shown to be achievable by a distributed linear dispersion codes [15] and a randomized space-time codes [55]. Although both AF and repetition-coding-based DF protocols achieve a full diversity gain, their throughput may degrade because each transmitting relay takes one time slot to transmit to the destination. This limitation can be overcome by enhancing cooperative protocols, namely by using selection/opportunistic or incremental relaying protocol which will be described in the following.

1.2.3 Selection/Opportunistic Relaying

Consider m relays available to assist transmission from the source to the destination. Instead of allowing all relays in the AF protocol or all the relays in the decoding set in the DF protocol to transmit in the second phase, selection/opportunistic relay protocols choose one “best” relay to transmit in the second phase [2], [4], [27], [31], [43], [56], [16] [65]. Surprisingly, cooperative protocols based on using smart relay selection strategies usually achieve full diversity order while providing higher throughput than the standard protocols. In fact, the superior throughput performance of selection relaying protocols stems from the fact that they use radio resources (i.e., power and bandwidth) more efficiently than basic cooperative protocols presented in the previous sections.

Some typical relay selection strategies for both AF and DF based protocols are presented next. Consider an AF protocol with one selected relay, say r_i . From (1.4), the capacity of the source-destination channel with one relay is

$$C_{\text{SAF}} = \frac{1}{2} \log \left(1 + \text{SNR}_s |a_{sd}|^2 + \frac{\text{SNR}_s |a_{sr_i}|^2 \text{SNR}_{r_i} |a_{r_i d}|^2}{\text{SNR}_s |a_{sr_i}|^2 + \text{SNR}_{r_i} |a_{r_i d}|^2 + 1} \right). \quad (1.7)$$

Therefore, to maximize the capacity, a relay selection strategy would choose a relay that maximizes [65]

$$\frac{\text{SNR}_s |a_{sr_i}|^2 \text{SNR}_{r_i} |a_{r_i d}|^2}{\text{SNR}_s |a_{sr_i}|^2 + \text{SNR}_{r_i} |a_{r_i d}|^2 + 1}. \quad (1.8)$$

For the DF protocol, there is a set of relays which successfully decode the signal in the first phase (i.e., in the decoding set $D(s)$). If relay $r_i \in D(s)$ is chosen for transmission in the second phase, the capacity of the source-destination channel is

$$C_{\text{SDF}} = \frac{1}{2} \log \left(1 + \text{SNR}_s |a_{sd}|^2 + \text{SNR}_{r_i} |a_{r_i d}|^2 \right). \quad (1.9)$$

Therefore, to maximize the source-destination capacity, an opportunistic relay selection strategy would choose a relay in the decoding set that maximizes

$$\text{SNR}_s |a_{sd}|^2 + \text{SNR}_{r_i} |a_{r_i d}|^2. \quad (1.10)$$

In [2] and [65], it has been shown that relay selection strategies in (1.8) and (1.10) achieve the full diversity order. Note that these selection metrics require the estimates of $\text{SNR}_s |a_{sd}|^2$ and $\text{SNR}_{r_i} |a_{r_i d}|^2$. In [4], two simpler relay selection metrics which require only channel gains a_{sr_i} and $a_{r_i d}$ have been proposed. Specifically, relay selection strategies that choose a relay such that

$$r_i^* = \arg \max_{r_i} \left\{ \min \left\{ |a_{sr_i}|^2, |a_{r_i d}|^2 \right\} \right\} \quad (1.11)$$

$$r_i^* = \arg \max_{r_i} \left\{ \frac{2}{\frac{1}{|a_{sr_i}|^2} + \frac{1}{|a_{r_i d}|^2}} = \frac{2 |a_{r_i d}|^2 |a_{sr_i}|^2}{|a_{r_i d}|^2 + |a_{sr_i}|^2} \right\}. \quad (1.12)$$

have been developed.

Note that the relay selection criterion in (1.11) chooses a relay with largest channel gains in both source-relay and relay-destination links. On the other hand, the relay selection rule in (1.12) maximizes the harmonic mean of channel gains for the source-relay and relay-destination links. It has been shown in [4] that these relay selection criteria provide the optimum diversity-multiplexing tradeoff achieved by the distributed space-time cooperative protocol [23]. Other relay selection strategies for orthogonal frequency-division multiple access (OFDMA)-based wireless cellular relay networks and ad hoc networks can be found in [27], [43], [56].

1.2.4 Incremental Relaying

Although selection relaying uses radio resources more efficiently than fixed relaying, both fixed and selection relaying protocols have to always repeat transmission. In fact, direct transmission from the source to the destination may be successful if the corresponding channel condition is not too “bad”. Therefore, it can be more efficient if relay transmission is invoked only when direct transmission from the source to the destination in the first phase fails.

One simple incremental relaying protocol based on using AF principle which exploits the aforementioned aspect works as follows [22]. The destination upon decoding its received signal at the end of the first phase broadcasts the decoding outcome to the source and relays. If the destination succeeds in decoding the message in the first phase, the source and relays do nothing. Otherwise, all or selected relays amplify their received signals and transmit to the destination. The destination combines all the signals and decodes again.

In fact, incremental relaying protocol can be implemented as an extension of hybrid automatic repeat request (ARQ) protocol [1], [10], [25], [26], [64]. One possible implementation of ARQ-based relaying can be described as follows [64]. Initially, the source node encodes b bits of information into a code-word with length n symbols. The code-word is broken into M blocks each of which has length n/M . The code can be a simple repetition code where all blocks are identical or the blocks can be obtained by puncturing a mother code. The protocol starts by transmitting the first block from the source node. The destination upon decoding the message broadcasts the decoding outcome to all other nodes. If the decoding at the destination is successful, the source proceeds to transmit a new message. Otherwise, either all or one selected relay in the decoding set (i.e., relays that successfully decode the message) re-encode the message and transmit the second block to the destination. The destination combines all the received blocks and attempts to decode again. This procedure continues until the destination is successful in decoding the message or all M blocks are transmitted and the message is discarded.

Incremental relaying has both diversity and throughput advantages because relaying is invoked only when necessary. In [22], the authors have shown that incremental relaying using AF principle as presented above achieves the full diversity order. In addition, it can be seen that ARQ-based incremental relaying allows many different code designs where well-investigated hybrid ARQ protocol can be adapted to the relaying network setting. Also, a combination of incremental relaying, hybrid ARQ and relay selection achieves throughput and energy improvement compared to the standard protocols while still having a full diversity gain.

1.2.5 Other Protocol Enhancements

There are some other possible enhancements of the aforementioned cooperative protocols available in the literature. In particular, network coding can be combined with standard cooperative protocols to improve throughput performance [21], [28], [61]. The network coding is based on the idea that the users involved in cooperative transmissions can combine their own information with other users's information (e.g., by using linear coding [28]) and transmit the combined information in an appropriate manner. This is because through cooperation, users know the messages of their assisted users. This would enhance throughput performance for each user because a single transmission transmits

both the user's own message and the message of an assisted user in the combined signal.

Other possible enhancements include combination of adaptive modulation and coding into cooperative protocols [41], employing coding in cooperative protocols [14], adding power and scheduling considerations for selection of a group of active retransmitting nodes [19], [20]. In [60], a detection technique for wireless networks where synchronization of users are not possible has been proposed. This technique mimics an equalization technique employed in a frequency-selective fading channel.

Finally, relaying transmission concepts can be combined with a medium-access-control (MAC) protocol to improve its throughput performance [68]. Specifically, through exchanging control information (e.g., RTS/CTS handshake signals), each user can find the optimal transmission strategy between direct transmission and relaying transmission through other relays (i.e., neighboring nodes). By choosing a transmission strategy with higher throughput, the MAC protocol can achieve better overall throughput performance.

1.2.6 Further Discussions

Summarizing the aforementioned cooperative protocols, it is also worth pointing out some important design issues. First, in principle a source-destination pair can be assisted by a large number of relays; however, a small number of "good" relays would be selected for cooperative transmission in most practical applications. Selecting a small number of relays for cooperation would be preferred taking into account both design complexity and overall network performance. Second, cooperative transmission may not be always beneficial especially if the source-destination link is very strong. Therefore, an adaptive cooperative protocols based on using a right amount of cooperation such as incremental relaying protocols would perform better than non-adaptive protocols (e.g., AF and DF protocols).

Due to the distributed nature of cooperative diversity protocols, their employment raises several practical implementation issues. First, synchronization among wireless nodes for implementing the MRC or distributed beamforming may be difficult. In order to resolve this challenge, a receiver detector at the destination node must be able to operate under asynchronous transmissions from a source and relay nodes [60]. In scenarios where space-time coding is employed, the underlying coding strategy should be designed to operate in a decentralized manner [15], [55]. Second, optimal resource allocation should be performed so that optimum network performance is achieved. Therefore, optimal resource allocation guarantees that radio resources is allocated efficiently to source and relay nodes in the network. We discuss resource allocation issues in more details in the following sections.

1.3 Applications and Implementation of Cooperative Diversity

1.3.1 Cellular Relay Networks

Cooperative diversity can be employed to enhance throughput and/or improve BER performance of a multi-hop cellular network [24]. In particular, users can take turn to serve as relays for one another. Alternatively, a set of fixed relays can be implemented to assist all the users in each cell. In [51], [52], a cooperation strategy for a two-user code division multiple access (CDMA) cellular wireless network has been proposed. According to this strategy, each user has two transmission periods where it transmits directly to the base station (BS) in the first period and cooperates with the other user to transmit in the second period. It has been shown that user cooperation indeed increases network throughput and decreases network sensitivity to channel variations.

For multihop cellular networks with fixed relays, transmissions from/to the BS of different users with the help of deployed relays can enhance throughput and BER performances. Since a small number of deployed relays is shared by a large number of users, a relay selection strategy should be employed for cooperative transmissions between users and the BS. In addition, if each fixed relay has several transceivers which can assist several users simultaneously, power and bandwidth allocation should be performed at these relays to optimize the overall network performance. In general, cooperative transmissions between users and the BS can occur in a multihop fashion [24]. In this case, a joint cluster-based routing and cooperative transmission can be employed as for wireless ad hoc networks. This is presented in the next subsection.

1.3.2 Cluster-Based Wireless Ad Hoc Networks

In wireless ad hoc networks, a source may want to communicate with a destination that is far away. Hence, a routing protocol is needed to deliver data in a multihop fashion. A traditional routing protocol typically finds a set of wireless links from the source to the destination to establish a multihop route for end-to-end data delivery. Using cooperative diversity, the multihop route can be formed by a set of cooperative and robust abstract “links” instead of simple wireless links [49]. In fact, cooperative diversity can be jointly used with a hierarchical routing to enhance end-to-end performance [12].

For hierarchical routing in wireless ad hoc networks, wireless nodes in the network form clusters each of which is a set of wireless nodes in a neighborhood [38]. Each cluster has one cluster head. A cluster mimics a cell in wireless cellular network where the cluster head functions similarly to a BS. A hierarchical routing protocol typically finds a set of clusters between the source and the destination. Then, end-to-end routing of information is performed within and between clusters independently. Cooperative diversity can be used for inter-cluster routing as being shown in Fig. 1.2.

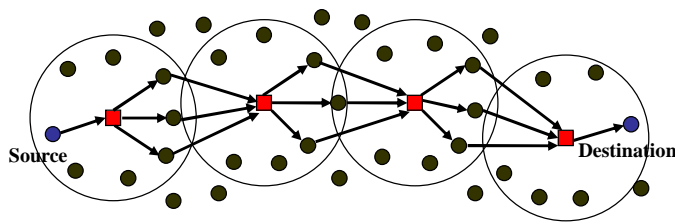


Fig. 1.2. Cluster-based cooperative transmission

In this cluster-based cooperative routing, a set of wireless nodes between any two neighboring clusters is chosen by cluster heads to serve as gateway nodes. The gateway nodes are the relay nodes that assist transmission between two clusters. Therefore, any cooperative protocols presented in the previous sections can be used for inter-cluster transmission. Note that this network architecture can also be used in infrastructure-based wireless mesh networks where mesh routers serving a number of mesh clients can serve as cluster heads. In [26], an analytical model has been developed to quantify performance of the aforementioned cluster-based cooperative routing where incremental relaying is employed for inter-cluster transmission.

1.3.3 Cooperative Broadcast in Wireless Ad Hoc Networks

Cooperative diversity can be exploited to enhance broadcast performance in wireless ad hoc networks [37], [50], [53], [54], [55]. In broadcast applications, a message is required to be transmitted from a source to all other nodes in the network. By using cooperative diversity, performance improvement in terms of energy consumption or message delivery probability can be achieved by exploiting the fact that each node in the network can collect signals from several simultaneously transmitting nodes. As a special case, cooperative broadcast can be performed in different levels as follows [53]. Each node in the network accumulates signals transmitted from other nodes until it achieves high enough SNR to decode the message. After successfully decoding the message, a node broadcasts it into the network. Therefore, by using a smart detection technique, each node can combine signals transmitted from different nodes to enhance the broadcast performance.

1.4 Performance Analysis and QoS Provisioning for Wireless Relay Networks

There is a large body of literature on performance analysis and QoS provisioning for wireless relay networks. In this section, we attempt to review some important research problems and issues along these lines. In fact, there

are two important research directions pursued in the literature. The first direction focuses on analyzing performances of cooperative diversity protocols. Performance measures under consideration include ergodic, outage capacity, bit/symbol error rate (B/SER), throughput and packet/frame delay. The second direction concentrates on QoS provisioning, resource allocation and protocol engineering for particular cooperative protocols and applications. In fact, solution approaches for the underlying problems in this direction usually rely on some results in the first direction.

Here, we review some research issues and results for the aforementioned directions. Note that we have discussed ergodic and outage capacity for several important cooperative diversity protocols in section 1.2. In addition, resource allocation for wireless relay networks plays an important role in improving the network performance; therefore, we will treat this topic in more detail in the next section. Regarding B/SER performance analysis of cooperative diversity protocols, there exists many publications which consider different protocols, network settings (e.g., multi-branch, multihop relay networks) [6], [48]. In general, the BER of a particular cooperative diversity protocol is lower-bounded by the corresponding outage probability.

Exact analysis of B/SER for cooperative diversity protocols (e.g., AF and DF protocols) is usually cumbersome. However, there are some existing works which consider approximated B/SER analysis for these protocols (e.g., [48]). In particular, the approximated analysis in [48] gives closed-form expressions of SER for the AF cooperative protocol, which is quite accurate in the high SNR regime. Specifically, consider a scenario where there are m relays helping a source-destination pair. Let γ_{sd} , γ_{sr_i} , γ_{r_id} be the average received SNR for the source-destination, source-relay i , relay i -destination links, respectively. The SER of the AF cooperative protocol can be approximated as [48]

$$\bar{P}_e \approx C(m, K) \frac{1}{\gamma_{sd}} \prod_{i=1}^m \left(\frac{1}{\gamma_{sr_i}} + \frac{1}{\gamma_{r_id}} \right) \quad (1.13)$$

where $C(m, K)$ is a constant depending on the number of relays m , modulation scheme, and specular factor K of the Ricean fading channel. The SER in (1.13) shows that the AF cooperative protocol achieves the full diversity order. Derivation of SER for the DF protocol can be found in [6].

Regarding QoS provisioning issues, many wireless applications have delay constraints to guarantee minimum QoS performance besides a common minimum B/SER requirement. In addition, data traffic may be bursty which is usually queued in data buffers upon arriving from the higher layers. Therefore, the total packet delay may consist of queueing and transmission delay components [7]. Of course, when data is not buffered, the total packet delay is simply the transmission delay [40]. For cooperative diversity protocols which involve several block transmissions for each data packet such as incremental relay protocols, the total packet delay can be controlled by smartly regulating the average number of transmissions. This is similar to controlling the num-

ber of transmission attempts in a classical truncated ARQ protocol [25]. In general, a cross-layer model should be developed to harmonize and optimize the network performance while meeting delay constraints [26].

For emerging applications in multihop wireless networks (e.g., wireless mesh and sensor networks), network/protocol design should be performed to optimize network or QoS performance measures of interest. In [27], optimal cross-layer algorithms have been developed to perform joint relay selection, power allocation, and routing to optimize different performance measures including power minimization and rate utility maximization in a general multihop wireless network. In [18], [35], centralized and distributed cooperative routing protocols have been proposed to minimize energy consumption. Finally, relay-selection and power allocation strategies have been proposed to maximize lifetime of a wireless sensor network using the AF cooperative diversity protocol in [13]. These are just few examples where network protocol design and QoS provisioning problems for the corresponding applications are considered. In general, these design problems depend on the specifics of underlying applications which may require very diverse solution approaches to resolve.

1.5 Resource Allocation for Cooperative Wireless Networks

1.5.1 Single-User Resource Allocation

There are quite a few existing works considering resource allocation for single-user cooperative wireless networks [11], [29], [30], [32], [36], [63]. For the single-user setting, there is only one source communicating to only one destination with the help of one or several relays. Since the capacity of a general relay channel is still an open problem, only some upper and lower capacity bounds are derived in the literature. In [36], lower and upper capacity bounds for different cooperation strategies including time-division relaying and compress-and-forward have been derived. Optimal power allocation schemes which aimed at maximizing these capacity bounds have also been adopted. In [30], the capacity bounds for parallel relay channels with degraded sub-channels have been derived and optimized through power allocation.

For the practical AF and DF protocols, optimal power allocation methods aiming at maximizing the SNR have been developed in [29], [65]. Using the SNR expression of the AF protocol with m relays (1.4), the problem of SNR maximization under total and individual relay power constraints can be written as

$$\max_{\{P_{r_i}\}} \text{SNR}_s |a_{sd}|^2 + \sum_{i=1}^m \frac{\text{SNR}_s |a_{sr_i}|^2 \text{SNR}_{r_i} |a_{r_i d}|^2}{\text{SNR}_s |a_{sr_i}|^2 + \text{SNR}_{r_i} |a_{r_i d}|^2 + 1} \quad (1.14a)$$

$$\text{subject to: } \sum_{i=1}^m P_{r_i} \leq P_T \quad (1.14b)$$

$$0 \leq P_{r_i} \leq P_i^{\max}. \quad (1.14c)$$

Assuming that white Gaussian noise powers measured in the signal bandwidth at all relays and the destination are equal to N and the source transmission power P_s is fixed. The optimal relay power in the high SNR regime can be found as [65]

$$P_{r_i} = \left[\frac{P_s |a_{sr_i}|^2}{|a_{r_i d}| \sqrt{N}} \lambda - \frac{P_s |a_{sr_i}|^2}{|a_{r_i d}|^2} \right]_0^{P_i^{\max}} \quad (1.15)$$

where $[\cdot]_0^{P_i^{\max}}$ denotes the projection operation on the interval $[0, P_i^{\max}]$, and λ is chosen such that the total relay power is satisfied. Optimal relay power allocation for DF protocol is more involving and depends on the decoding strategy employed at the relays. In [29], optimal relay solutions have been derived for some special cases.

1.5.2 Multi-User Resource Allocation

In this section, we present a resource allocation framework for a multi-user wireless relay network. More details can be found in [44], [45], [46], [47].

System Models

Consider a multi-user relay network in which M source nodes s_i transmit data to their corresponding destination nodes d_i , $i \in \{1, \dots, M\}$. There are also L relay nodes r_j , $j \in \{1, \dots, L\}$ which are employed to assist transmissions from source to destination nodes. The set of relay nodes assisting the transmission of the source node s_i is denoted by $\mathcal{R}(s_i)$. The set of source nodes using the relay node r_j is denoted by $\mathcal{S}(r_j)$, i.e., $\mathcal{S}(r_j) = \{s_i | r_j \in \mathcal{R}(s_i)\}$. Therefore, one particular relay node can forward data for several users.⁵ We assume that the AF cooperative scheme is used for re-transmission. Moreover, *orthogonal transmissions* are assumed for simultaneous transmissions among different users by using different channels, (e.g., different frequency bands), and time division multiplexing is employed by the AF cooperative scheme for each user. Then, the transmission from a source to a destination node can be described as follows. In the first phase, each source node s_i transmits data to its chosen relays in the set $\mathcal{R}(s_i)$. Then, in the second phase, each relay node amplifies and forwards its received signal to d_i . The corresponding system model is shown in Fig. 1.3.

⁵ The term ‘user’ refers to a source-destination pair in this context.

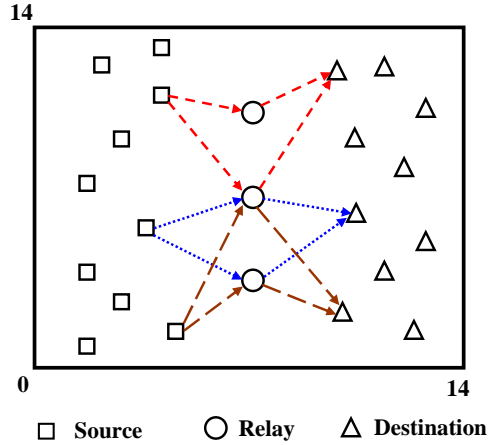


Fig. 1.3. Multi-user wireless relay network

The investigated system model is quite general and it covers a large number of applications in different network settings. For example, the model can be applied to cellular wireless networks using relays for uplink with one destination (BS) or downlink with one source (BS) and many destinations. It can also be directly applied to multi-hop wireless networks such as sensor/ad hoc or wireless mesh networks. Moreover, in our model, each source can be assisted by one, several, or all available relays. The presented model, therefore, captures most relay models considered in the literature.

Let P_{s_i} denote the power transmitted by source node s_i , $P_{r_j}^{s_i}$ denotes the power transmitted by relay node $r_j \in \mathcal{R}(s_i)$ for assisting the source node s_i , and $a_{s_i r_j}$ and $a_{r_j d_i}$ denote the channel gains for links s_i-r_j and r_j-d_i , respectively. The channel gains could include the effects of path loss, shadowing and fading. To keep the model in this section general, we assume that the variances of additive circularly symmetric white Gaussian noise (AWGN) at the relay r_j and at the destination node d_i are N_{r_j} , N_{d_i} , respectively. Assuming that MRC is employed at the destination node d_i , the SNR of the combined signal at the destination node d_i can be written as ⁶

$$\gamma_i = \sum_{r_j \in \mathcal{R}(s_i)} \frac{P_{r_j}^{s_i}}{\alpha_{r_j}^{s_i} P_{r_j}^{s_i} + \beta_{r_j}^{s_i}} \quad (1.16)$$

where

$$\alpha_{r_j}^{s_i} = \frac{N_{r_j}}{|a_{s_i r_j}|^2 P_{s_i}}, \quad \beta_{r_j}^{s_i} = \frac{N_{d_i} N_{r_j}}{|a_{s_i r_j}|^2 |a_{r_j d_i}|^2 P_{s_i}} + \frac{N_{d_i}}{|a_{r_j d_i}|^2}.$$

⁶ We consider the case when the source-to-relay link is (much) better than the source-to-destination link, which would be an outcome of a typical relay selection strategy employed by each source node.

It can be verified that the SNR γ_i for user s_i is concave increasing with respect to $P_{r_j}^{s_i}$, $r_j \in \mathcal{R}(s_i)$. Moreover, the rate of user s_i which is defined as $R_i = \log(1 + \gamma_i)$ is also concave increasing.

Formulations of Power Allocation Problem

In general, resource allocation in wireless networks should take into account fairness among users. An attempt to maximize sum rate of all the users would degrade performance of the worst user(s) significantly. To balance fairness and throughput performance for all the users, we consider two different optimization criteria for power allocation. The first criterion aims at maximizing the minimum rate among all users. In essence, this criteria tries to make rates of all users as equal as possible. For the second criterion, users are given different weights and power allocation is performed to maximize the weighted sum of rates for all users. In the latter case, user(s) in unfavorable conditions could be allocated large weights to prevent severe degradation of their performance. Another possible application for this optimization criterion is to perform QoS differentiation when users of higher service priority can be allocated larger weights. In both optimization criteria, we impose constraints on the total maximum power that each relay can use to assist the corresponding users.

A. Max-Min Rate Fairness Based Power Allocation

We first consider the power allocation problem under max-min rate fairness for the users. Mathematically, it can be formulated as [47]

$$\max_{\{P_{r_j}^{s_i} \geq 0\}} \min_{s_i} R_i \quad (1.17a)$$

$$\text{subject to: } \sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} \leq P_{r_j}^{\max}, \quad j = 1, \dots, L \quad (1.17b)$$

where R_i is the rate of user s_i and $P_{r_j}^{\max}$ is the maximum power of relay r_j . The left-hand side of (1.17b) is the total power that relay r_j allocates to its assisted users which is constrained to be less than its maximum power budget. This constraint is required to avoid overloading the relays in the network.

In general, the power allocation obtained according to the problem (1.17a)–(1.17b) can result in a loss in network throughput because the objective function (1.17a) requires to improve the performance of the worst user(s) that can decrease the overall system throughput. Therefore, this criterion is applicable for networks in which all users are of (almost) equal importance. This is the case, for example, when wireless users pay the same subscription fees, and thus, demand similar level of QoS. It can be seen that the set of linear inequality constraints with positive variables in the optimization problem (1.17a)–(1.17b) is compact and nonempty. Hence, the problem (1.17a)–(1.17b) is always feasible. Moreover, since the objective function $\min_{s_i} R_i$ is an increasing function of allocated powers, the inequality constraints (1.17b) should be

met with equality at optimality. Introducing a new variable T , we can equivalently rewrite the optimization problem (1.17a)–(1.17b) in a standard form as

$$\max_{\{P_{r_j}^{s_i} \geq 0, T \geq 0\}} T \quad (1.18a)$$

$$\text{subject to: } T - R_i \leq 0, \quad i = 1, \dots, M \quad (1.18b)$$

$$\sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} = P_{r_j}^{\max}, \quad j = 1, \dots, L. \quad (1.18c)$$

It can be verified that the optimization problem (1.18a)–(1.18c) is convex. Thus, its optimal solution can be obtained by standard convex optimization algorithms [5]. In the following, we describe a different formulation for the power allocation problem, which achieves better throughput performance.

B. Weighted-Sum of Rates Fairness Based Power Allocation

As discussed before, max-min rate fairness based power allocation tends to improve performance of the worst user at the cost of overall network throughput degradation. Maximization of the weighted-sum of rates can potentially achieve certain fairness for different users by allocating large weights to users in unfavorable channel conditions while maintaining good network performance in general. Let w_i denote the weight allocated to user s_i . Then, the weighted-sum of rates fairness based power allocation problem can be mathematically posed as [47]

$$\max_{\{P_{r_j}^{s_i} \geq 0\}} \sum_{i=1}^M w_i R_i \quad (1.19a)$$

$$\text{subject to: } \sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} \leq P_{r_j}^{\max}, \quad j = 1, \dots, L. \quad (1.19b)$$

As in the optimization problem (1.17a)–(1.17b), it can be seen that the constraints (1.19b) in the problem (1.19a)–(1.19b) must be met with equality at optimality. Otherwise, the allocated powers can be increased to improve the objective function, and thus, it contradicts with the optimality assumption. In addition, it can be verified that this optimization problem is convex; therefore, its optimal solution can be obtained by any standard convex optimization algorithms.

We would like to note that power allocation schemes based on other fairness criteria can also be considered. For instance, the proportional fairness criterion can be adopted. In terms of system-wide performance metric such as the network throughput, the latter criterion can ensure more fairness than the weighted-sum of rates, while achieving better performance than the max-min fairness [17]. It can be shown that the objective function to be maximized in the proportional fairness based power allocation scheme is $\prod_{i=1}^M R_i$. Conse-

quently, this objective function can be re-formulated as convex function using the log function.

Distributed Implementation for Power Allocation

To reduce communication overhead and to implement online power allocation for the multi-user relay network, we now develop a distributed algorithm for solving the optimization problem (1.19a)–(1.19b) and show that such a solution converges to the optimal solution. The algorithm is developed based on the dual decomposition approach in convex optimization [3]. Applications of this optimization technique for distributed routing, reverse engineering of MAC, and transmission control protocols can be found in [8], [17], [62].

In dual decomposition method, the original problem is separated into independent subproblems that are coordinated by a higher-level master dual problem. Now, we first write the Lagrangian function by relaxing the total power constraints for the relays as follows

$$\mathcal{L}(\boldsymbol{\mu}, P_{r_j}^{s_i}) = \sum_{i=1}^M w_i R_i - \sum_{j=1}^L \mu_j \left(\sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} - P_{r_j}^{\max} \right) \quad (1.20)$$

where $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_L]$, $\mu_j \geq 0$, $j = 1, \dots, L$ are the Lagrange multipliers corresponding to the L linear constraints on the relay powers.

Using the fact that

$$\sum_{j=1}^L \mu_j \sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} = \sum_{i=1}^M \sum_{r_j \in \mathcal{R}(s_i)} \mu_j P_{r_j}^{s_i}$$

the Lagrangian in (1.20) can be rewritten as

$$\mathcal{L}(\boldsymbol{\mu}, P_{r_j}^{s_i}) = \sum_{i=1}^M \left[w_i R_i - \sum_{r_j \in \mathcal{R}(s_i)} \mu_j P_{r_j}^{s_i} \right] + \sum_{j=1}^L \mu_j P_{r_j}^{\max}.$$

The corresponding dual function of the Lagrangian can be written as

$$g(\boldsymbol{\mu}) = \max_{\{P_{r_j}^{s_i} \geq 0\}} \mathcal{L}(\boldsymbol{\mu}, P_{r_j}^{s_i}). \quad (1.21)$$

Since the original optimization is convex, strong duality holds, and the solution of the underlying optimization problem can be obtained from that of the corresponding dual problem as follows

$$\min_{\{\mu_j \geq 0\}} g(\boldsymbol{\mu}) \quad (1.22)$$

It can be seen that the dual function in (1.21) can be found by solving M separate subproblems corresponding to M different users as follows

$$\max_{\{P_{r_j}^{s_i} \geq 0\}} \mathcal{L}_i(\boldsymbol{\mu}, P_{r_j}^{s_i}) = w_i R_i - \sum_{r_j \in \mathcal{R}(s_i)} \mu_j P_{r_j}^{s_i} \quad (1.23)$$

where $\mathcal{L}_i(\boldsymbol{\mu}, P_{r_j}^{s_i})$ corresponds to the i th component of the Lagrangian. Let $\mathcal{L}_i^*(\boldsymbol{\mu})$ be the optimal value of $\mathcal{L}_i(\boldsymbol{\mu}, P_{r_j}^{s_i})$ obtained by solving the problem (1.23), then the dual problem in (1.22) can be rewritten as

$$\min_{\{\mu_j \geq 0\}} g(\boldsymbol{\mu}) = \sum_{i=1}^M \mathcal{L}_i^*(\boldsymbol{\mu}) + \sum_{j=1}^L \mu_j P_{r_j}^{\max} \quad (1.24)$$

A distributed power allocation algorithm can be developed by iteratively and sequentially solving the problems (1.23) and (1.24). This algorithm is also known in optimization theory as a primal-dual algorithm. The Lagrange multiplier $\mu_j \geq 0$ represents the pricing coefficient for each unit power at relay j . Therefore, $\mu_j P_{r_j}^{s_i}$ can be seen as the price that user s_i must pay for using $P_{r_j}^{s_i}$ at each relay $r_j \in \mathcal{R}(s_i)$. In particular, the optimization problem (1.23) can be interpreted as follows. The user s_i tries to maximize its rate minus the total price that it has to pay given the price coefficients at relays. The weight w_i can be seen as a ‘gain’ coefficient for each unit rate for user s_i .

The details of the distributed algorithm which implements the power allocation are as follows. The master dual problem is solved in a distributed fashion at each relay. Specifically, each relay r_j first broadcasts its initial ‘price’ value, i.e., Lagrange multiplier μ_j . These price values are used by the receivers to compute the optimal power levels that the relays should allocate to that particular user. The optimal powers are fed back to the relays, which then updates the next values of the μ_j , $j = 1, \dots, L$. This procedure is repeated until the so-obtained solution converges to the optimal one.

Note that the dual function $g(\boldsymbol{\mu})$ is differentiable. Therefore, the master dual problem (1.22) can be solved by using the gradient descent method. The dual decomposition presented in (1.23) allows each user s_i , for the given μ_j , to find the optimal allocated power $r_j \in \mathcal{R}(s_i)$ as follows:

$$P_{r_j}^{s_i}(\boldsymbol{\mu}) \Big|_{\text{opt}} = \arg \max \left\{ w_i R_i - \sum_{r_j \in \mathcal{R}(s_i)} \mu_j P_{r_j}^{s_i} \right\} \quad (1.25)$$

which is unique due to the strict concavity.

Due to the fact that the solution of the problem (1.25) is unique, the dual function $g(\boldsymbol{\mu})$ in the master problem (1.22) is differentiable, which allows us to use the following iterative gradient method to update the dual variables

$$\mu_j(t+1) = \left[\mu_j(t) - \zeta \left(P_{r_j}^{\max} - \sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i}(\boldsymbol{\mu}(t)) \Big|_{\text{opt}} \right) \right]^+ \quad (1.26)$$

where t is the iteration index, $[\cdot]^+$ denotes projection onto the feasible set of non-negative numbers, and ζ is the sufficiently small positive step size. The

dual variables $\boldsymbol{\mu}(t)$ will converge to the dual optimal $\boldsymbol{\mu}_{\text{opt}}$ as $t \rightarrow \infty$, and the primal variable $P_{r_j}^{s_i}(\boldsymbol{\mu}(t))\Big|_{\text{opt}}$ will also converge to the primal optimal variable $P_{r_j}^{s_i}(\boldsymbol{\mu}_{\text{opt}})\Big|_{\text{opt}}$. Updating $\mu_j(t)$ via (1.26) can be interpreted as follows. The relay r_j updates its price depending on the requested power levels from its users. The price is increased when the total requested power from users is larger than its maximum limit. Otherwise, the price is decreased. Finally, we summarize the distributed power allocation algorithm as follows.

Distributed Power Allocation Algorithm

- **Parameters:** the receiver of each user estimates/collects its weight coefficient w_i and channel gains of its transmitter-relay and relay-receiver links.
- **Initialization:** set $t=0$, each relay j initializes $\mu_j(0)$ equal to some non-negative value and broadcasts this value.
- **Iterations:**
 1. The receiver of user s_i solves its problem (1.25) and then broadcasts the solution $P_{r_j}^{s_i}(\boldsymbol{\mu}(t))\Big|_{\text{opt}}$ to its relays.
 2. Each relay r_j receives the requested power levels and updates its prices with the gradient iteration (1.26) using the information received from its assisted users. Then, it broadcasts the new value $\mu_j(t+1)$.
 3. Set $t = t + 1$ and go to step 1 until satisfying a predetermined stopping criterion.

The convergence proof of the general primal-dual algorithm can be found in [3]. This algorithm only requires message exchange between relays and their assisted receivers. Therefore, it can be easily implemented in a distributed manner with low overhead.

Joint Admission Control and Power Allocation

Here, we consider a scenario in which users have minimum rate requirements. This scenario is important for real-time/multimedia applications which require certain minimum rates to maintain QoS performance. Because network radio resources may be limited (e.g., limited source and/or relay power), supporting all users with their minimum required rates may not be feasible. Therefore, an admission control mechanism should be employed to determine which users to be admitted into the network. Then, power can be allocated to admitted users in order to ensure that each admitted user achieves the required QoS performance.

Specifically, consider a resource allocation problem that aims at minimizing the total relay power. In addition, each user has a minimum rate requirement. For the above described wireless systems with multiple users and multiple

relays, the problem of minimizing the total relay power given a minimum rate constraint for each user can be posed as

$$\min_{\{P_{r_j}^{s_i} \geq 0\}} \sum_{j=1}^L \sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} \quad (1.27a)$$

$$\text{subject to: } R_i \geq R_i^{\min}, \quad i = 1, \dots, M \quad (1.27b)$$

$$\sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} \leq P_{r_j}^{\max}, \quad j = 1, \dots, L \quad (1.27c)$$

where R_i^{\min} denotes the minimum rate requirement for user s_i .

Mathematically, there are instances in which the optimization problem (1.27a)–(1.27c) becomes infeasible. A practical implication of the infeasibility is that it is impossible to serve all M users at their desired QoS requirements. In QoS-supported systems, some users can be dropped or the rate targets can be relaxed as a consequence. We investigate the former scenario and try to maximize the number of users that can be admitted at their minimum rate requirements.

The joint admission control and power allocation problem can be mathematically posed as a two-stage optimization problem [33], [34]. All possible sets of admitted users S_0, S_1, \dots with possibly maximal cardinality (can be only one or several sets) are found in the first admission control stage, while the optimal set of admitted users S_k is the one among the sets S_0, S_1, \dots , which requires minimum transmit power in the second power allocation stage. Once the candidate set of admitted users has been determined, the power allocation problem can be shown to be a convex programming problem. However, the admission control problem is combinatorially hard, which introduces high complexity for practical implementation. Therefore, a low-complexity solution approach for the joint admission control and power allocation problem is highly desirable.

A. Reformulation of Admission Control and Power Allocation Problem

The joint admission control and power allocation problem can be equivalently written as a one-stage optimization problem that enables us to develop a low-complexity algorithm to solve the underlying problem. Toward this end, let x_i , $i = 1, \dots, M$ denote an indicator variable for user s_i where $x_i = 1$ if user i is admitted and $x_i = 0$, otherwise. Given these variables, the underlying problem can be rewritten as [47]

$$\max_{\{s_i \in \{0,1\}, P_{r_j}^{s_i} \geq 0\}} \sum_{i=1}^M x_i \quad (1.28a)$$

$$\text{subject to: } R_i \geq R_i^{\min} x_i, \quad i = 1, \dots, M \quad (1.28b)$$

$$\sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} \leq P_{r_j}^{\max}, \quad j = 1, \dots, L \quad (1.28c)$$

Note that the constraints (1.28b) are automatically satisfied for the users that are not admitted. The indicator variables help to represent the admission control problem in a more compact form. However, the combinatorial nature of the admission control problem still remains due to the binary variables x_i .

Following the conversion steps similar to those used in [34], the joint admission control and power allocation problem can be converted to the following one-stage optimization problem

$$\max_{\{x_i \in \{0,1\}, P_{r_j}^{s_i} \geq 0\}} \epsilon \sum_{i=1}^M x_i - (1 - \epsilon) \sum_{j=1}^L \sum_{s_i \in \mathcal{S}(r_j)} P_{r_j}^{s_i} \quad (1.29a)$$

$$\text{subject to:} \quad \text{constraints (1.28b), (1.28c)} \quad (1.29b)$$

where ϵ is some constant which is chosen such that

$$\frac{\sum_j P_{r_j}^{\max}}{\sum_j P_{r_j}^{\max} + 1} < \epsilon < 1. \quad (1.30)$$

The problem (1.29a)–(1.29b) is a compact mathematical formulation of the joint optimal admission control and power allocation problem. The proof of the equivalence of the one-stage optimization problem and the original two-stage optimization problem can be found in [34]. Moreover, the one-stage optimization problem is always feasible since in the worst case no users are admitted, i.e., $x_i = 0, \forall i = 1, \dots, M$.

B. Low-Complexity Algorithm

Although the original optimization problem (1.29a)–(1.29b) is NP-hard, its relaxation for which $x_i, i = 1, \dots, M$ are relaxed to be continuous can be shown to be a convex programming problem. In the following, we propose a reduced-complexity heuristic algorithm to perform joint admission control and power allocation. The following heuristic algorithm can be used to solve (1.29a)–(1.29b).

Joint Admission Control and Power Allocation Algorithm

1. Set $S := \{s_i \mid i = 1, \dots, M\}$.
2. Solve convex problem (1.29a)–(1.29b) for the sources in S with x_i being relaxed to be continuous in the interval $[0,1]$. Denote the resulting power allocation values as $P_{r_j}^{s_i^*}, j = 1, \dots, M$.
3. For each $s_i \in S$, verify whether

$$R_i^* \geq R_i^{\min}, \forall s_i \in S.$$

If this is the case, then stop and $P_{r_j}^{s_i^*}$ are power allocation solutions. Otherwise, remove the user s_i with largest gap to its target R_i^{\min} , i.e.,

$$s_i = \arg \min_{s_i \in S} \left\{ R_i^* - R_i^{\min} < 0 \right\}$$

from set S and go to Step 2.

It can be seen that after each iteration, either the set of admitted users and the corresponding power allocation levels are determined or one user is removed from the list of the most likely admitted users. Since there are M initial users, the complexity is bounded above by that of solving M convex optimization problems with different dimensions, where the dimension of the problem depends on the iteration. It is worth mentioning that the proposed reduced complexity algorithm always returns one solution.

Note that the objective function for the considered above joint admission control and power allocation problem was the minimization of the total relay power. However, the principle used to construct the above algorithm can be employed to develop similar algorithms for other object functions as well (e.g., max-min, weight-sum-rate functions). Due to space constraints, we do not consider these problems here.

1.5.3 Numerical Results for Multi-User Resource Allocation

Consider a wireless relay network as in Fig. 1.3 with ten users and three relays distributed in a two-dimensional region 14×14 where network sizes are measured with respect to some reference distance. The relays are fixed at coordinates $(10,7)$, $(10,10)$, and $(10,12)$. The source and destination nodes are deployed randomly in the area inside the box area $[(0,0), (7,14)]$ and $[(12,0), (14,14)]$, respectively. In our simulations, each user is assisted by two relays. The noise power is taken to be equal to $N_0 = 10^{-5}$. All users and relays are assumed to have the same minimum rate R^{\min} and maximum transmit power $P_{r_j}^{\max}$.

Numerical Results for Power Allocation

We show that by proper weight setting, the weighted-sum of rates maximization based power allocation scheme provides the flexibility required to support users with differentiated services. Particularly, we suppose that users 1 and 2 have higher priority than the others, and set the corresponding weights as $w_1 = w_2 = 5$, $w_3 = \dots = w_{10} = 1$ in the optimization problem (1.19a)–(1.19b). Fig. 1.4 displays the resulting rate of the high-priority users.⁷ For reference, we also include in Fig. 1.4 the corresponding results obtained by equal power allocation (EPA) and by weighted-sum of rates maximization with equal weight coefficients. It can be seen that over the wide range of the relay power limits, the weighted-sum of rates maximization scheme outperforms the EPA. Without much surprise, the performance of the EPA scheme is quite close to that of the weighted-sum of rates maximization with *equal*

⁷ We observe that users 1 and 2 have indistinguishable performance, so only one curve for each scheme is plotted.

weight coefficients. On the other hand, the weighted-sum of rates maximization with *unequal* weight coefficients provides noticeable rate enhancement to the high-priority users as compared to the other schemes, especially when the relays have severe power limitation, e.g., a rate gain of about 0.2 b/s/Hz when $P_{r_j}^{\max} = 10$. This figure indicates that the performance difference between different algorithms gets smaller for larger relay power limits. In other words, this reveals an interesting property that when the relays have more (or unlimited) available power, different (relay) power allocation strategies have much less impact on the user rate performance, which is limited by the source transmit power in this case.

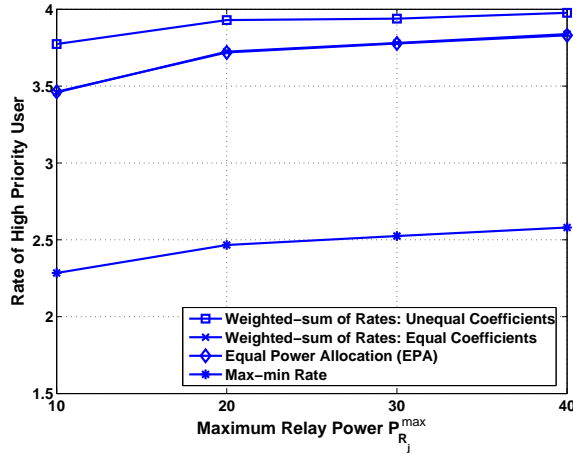


Fig. 1.4. Rate of high priority users versus $P_{r_j}^{\max}$

Fig. 1.5 shows the network throughput for the aforementioned power allocation schemes. In the max-min rate fairness based power allocation scheme, there is a significant loss in the network throughput since the objective is to improve the performance of the worst users. This confirms that achieving max-min fairness among users results in a performance loss for the whole system. The weighted-sum of rates fairness based scheme results in maximum throughput. Moreover, the rate gain of the weighted-sum of rates scheme over the EPA scheme is about 1.8 b/s/Hz over the range of the relay power limits. This gain comes at the cost of more complexity in system implementation to optimize the power levels. The weighted-sum of rates based scheme with unequal weights achieves slightly worse performance as compared to its counterpart with equal weights while providing better performance for the high priority users, i.e., users 1 and 2 in Fig. 1.4.

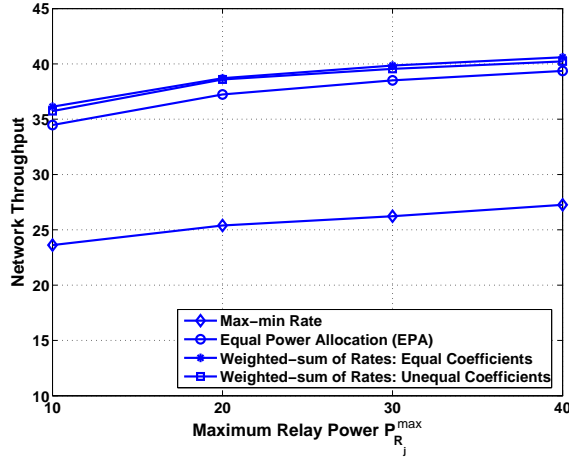


Fig. 1.5. Network throughput versus $P_{r_j}^{\max}$

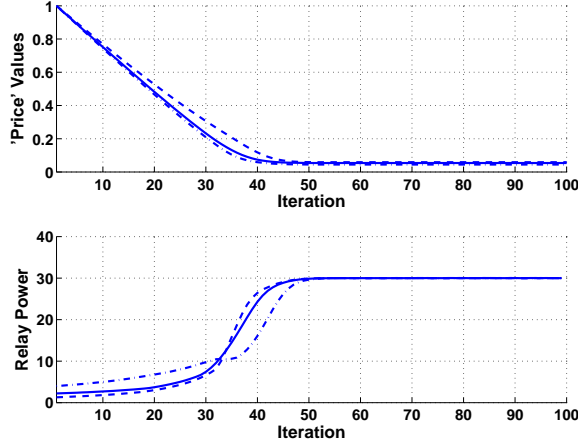


Fig. 1.6. Evolution of 'price' values and powers allocated at each relay

Figs. 1.6 and 1.7 show the evolution of different parameters in the proposed distributed implementation of the power allocation scheme for one particular channel realization. Specifically, Fig. 1.6 shows the evolution of the price values μ_j , $j = 1, 2, 3$ and the powers at the relays, while Fig. 1.7 displays the rate for each of the ten users and sum rates of all users. The update parameter ζ was set to 0.001 in this example. With such a choice of the update parameter, we can see that after about 50 updates, the algorithm converges to the op-

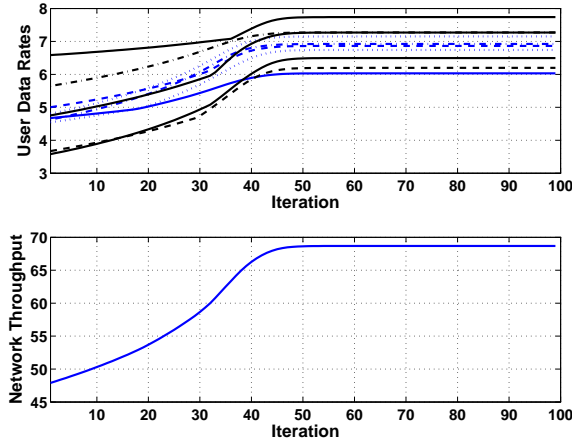


Fig. 1.7. Evolution of data rate for each user and user sum rate

timal solution obtained by solving the proposed optimization problem in the centralized manner.

Numerical Results for Joint Admission Control and Power Allocation

We also investigate the performance of the proposed joint admission control and power allocation algorithm with $P_{s_i} = 1$ and $P_{r_j}^{\max} = 10$. It is assumed that the channel gain is due to the path loss only and the locations of the source and destination nodes are fixed. Different values of $\gamma_i^{\min}/R_i^{\min}$ have been used. For reference, we also consider the *optimal* admission control and power allocation scheme using exhaustive search over all feasible user subsets. A *feasible* user subset contains the *maximum* possible number of users and is selected as the *optimum* user subset if it requires the *smallest* transmit power. The simulation parameters and the performance results for the *optimal* admission control and power allocation scheme, and the proposed heuristic scheme are recorded in the columns “optimum allocation” and “proposed algorithm” in Table I, respectively.

Note that the running time is measured in seconds. It can be seen that the proposed algorithm determines exactly the optimal number of admitted users in all cases. The transmit power required by our proposed algorithm is just marginally larger than that required by the optimal admission control and power allocation based on exhaustive search. However, the running time for the proposed algorithm is dramatically smaller than that required by the optimal one. This makes the proposed approach attractive for practical implementation. As expected, when γ_i^{\min} increases, a smaller number of users

is admitted with a fixed amount of power. For example, nine users and four users are admitted with SNR $\gamma_i^{\min} = 12$ dB and 14 dB, respectively.

Table 1.1. Results with $P_{s_i} = 1$, $P_{r_j}^{\max} = 10$ (Running time in seconds)

	Optimum Allocation	Proposed Algorithm
SNR/rate	12 dB/4.0746 b/s/Hz	12 dB/4.0746 b/s/Hz
# users served	9	9
Users served	1, 2, 3, 4, 6, 7, 8, 9, 10	1, 2, 3, 4, 6, 7, 8, 9, 10
Transmit power	20.3619	20.4446
Running time	18.72	5.39
SNR/rate	13 dB/4.3891 b/s/Hz	13 dB/4.3891 b/s/Hz
# users served	6	6
Users served	1, 2, 7, 8, 9, 10	1, 2, 7, 8, 9, 10
Transmit power	22.9531	23.0342
Users served	2, 3, 7, 8, 9, 10	-
Transmit power	23.7717	-
Running time	458.07	9.60
SNR/rate	14 dB/4.7070 b/s/Hz	14 dB/4.7070 b/s/Hz
# users served	4	4
Users served	7, 8, 9, 10	7, 8, 9, 10
Transmit power	25.6046	25.6195
Running time	850.28	11.78
SNR/rate	15 dB/5.0278 b/s/Hz	15 dB/5.0278 b/s/Hz
# users served	2	2
Users served	8, 10	8, 10
Transmit power	7.5310	7.5320
Running time	930.11	12.92
SNR/rate	16 dB/5.3509 b/s/Hz	16 dB/5.3509 b/s/Hz
# users served	1	1
Users served	8	8
Transmit power	9.8002	9.8025
Running time	931.11	13.15

1.6 Conclusions

In this chapter, we have presented a survey of cooperative diversity and discussed important resource allocation problems in wireless relay networks. Specifically, we have described fundamental cooperative protocols and pointed out some recent enhanced protocols available in the literature. Typical applications of cooperative diversity in multihop cellular networks, cluster-based

wireless ad hoc networks and broadcasting in ad hoc networks have been introduced. We have also presented the overview on resource allocation problems for single and multi-user wireless relay networks. For the multi-user case, we have investigated optimal relay power allocation and admission control problems with fairness consideration using centralized and distributed approaches. Simulation results are shown to confirm the theoretical developments.

References

1. K. Azarian, H. El Gamal, and P. Schniter, "On the optimality of the ARQ-DDF protocol," *IEEE Trans. Inform. Theory*, vol. 54, no. 4, pp. 1718-1724, Apr. 2008.
2. E. Beres and R. Adve, "Selection cooperation in multi-source cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 118-127, Jan. 2008.
3. D. P. Bertsekas, *Nonlinear Programming*, 2-nd Edition. Belmont, MA: Athena Scientific, 1999.
4. A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659-672, Mar. 2006.
5. S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. Cambridge, United Kingdom: Cambridge University Press, 2004.
6. J. Boyer, D. D. Falconer, H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1820-1830, Oct. 2004.
7. I. Cerutti, A. Fumagalli, and P. Gupta, "Delay models of single-source single-relay cooperative ARQ protocols in slotted radio networks with Poisson frame arrivals," *IEEE/ACM Trans. Networking*, vol. 16, no. 2, pp. 371-382, Apr. 2008.
8. M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proc. of IEEE*, vol. 95, no. 1, pp. 255-312, Jan. 2007.
9. T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572-584, Sept. 1979.
10. L. Dai and K. B. Letaief, "Throughput maximization of ad-hoc wireless networks using adaptive cooperative diversity and truncated ARQ," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1907-1918, Nov. 2008.
11. D. Gunduz and E. Erkip, "Opportunistic cooperation by dynamic resource allocation," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1446-1454, Apr. 2007.
12. X. Hong, K. Xu, and M. Gerla, "Scalable routing protocols for mobile ad hoc networks," *IEEE Network*, vol. 16, no. 4, pp. 11-21, Jul.-Aug. 2002.
13. W. -J. Huang, Y. -W. P. Hong, and C. -C . J. Kuo, "Lifetime maximization for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1800-1805, May 2008.
14. T. E. Hunter and A. Nosratinia, "Diversity through coded cooperation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 283-289, Feb. 2006.
15. Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524-3536, Dec. 2006.
16. Y. Jing and H. Jafarkhani, "Single and multiple relay selection schemes and their diversity orders," in *Proc. IEEE Inter. Conf. on Commun., ICC'08*, Beijing, China, May 2008.

17. F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *J. Operational Research Society*, vol. 49, no. 3, pp. 237-252, Mar. 1998.
18. A. E. Khandani, J. Abounadi, E. Modiano, and L. Zheng, "Cooperative routing in static wireless networks," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2185-2192, Nov. 2007.
19. Y. Ko, S. A. Vorobyov, and M. Ardakani, "How much multiuser diversity gain is required over large-scale fading?" in *Proc. IEEE Inter. Conf. on Commun., ICC'09*, Dresden, Germany, June 2009.
20. Y. Ko, S. A. Vorobyov, and M. Ardakani, "How much multiuser diversity is required for homogeneous multiuser systems?" *IEEE Trans. Wireless Commun.*, revised 2009.
21. R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782795, Oct. 2003.
22. J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 3062-3080, Dec. 2004.
23. J. N. Laneman and G. W. Wornell, "Distributed spacetime-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415-2425, Oct. 2003.
24. L. B. Le and E. Hossain, "Multihop cellular networks: Potential gains, research challenges, and a resource allocation framework," *IEEE Commun. Mag.*, vol. 45, no. 9, pp. 66-73, Sept. 2007.
25. L. B. Le, E. Hossain, and M. Zorzi, "Queueing analysis for GBN and SR ARQ protocols under dynamic radio link adaptation with non-zero feedback delay," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3418-3428, Sept. 2007.
26. L. B. Le and E. Hossain, "An analytical model for ARQ cooperative diversity in multihop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1786-1791, May 2008.
27. L. B. Le and E. Hossain, "Cross-layer optimization frameworks for multihop wireless networks using cooperative diversity," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2592-2602, July 2008.
28. S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 371381, Feb. 2003.
29. Y. Li, B. Vucetic, Z. Zhou, and M. Dohler, "Distributed adaptive power allocation for wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 948-958, Mar. 2007.
30. Y. Liang, V. V. Veeravalli, and H. V. Poor, "Resource allocation for wireless fading relay channels: Max-min solution," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3432-3453, Oct. 2007.
31. Z. Lin, E. Erkip, and A. Stefanov, "Cooperative regions and partner choice in coded cooperative systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1323-1334, July 2006.
32. J. Luo, R. S. Blum, L. J. Cimini, L. J. Greenstein, and A. M. Haimovich, "Decode-and-Forward cooperative diversity with power allocation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 793-799, Mar. 2007.
33. E. Matakani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Joint multiuser downlink beamforming and admission control: A semidefinite relaxation approach," in *Proc. IEEE Inter. Conf. Accous. Speech and Sig. Proc. (ICASSP)*, Honolulu, Hawaii, USA, Apr. 2007.

34. E. Matakani, N.D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Convex approximation techniques for joint multiuser downlink beamforming and admission control," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2682-2693, July 2008.
35. R. Madan, N. B. Mehta, A. F. Molisch, and J. Zhang, "Energy-efficient decentralized cooperative routing in wireless networks," *IEEE Trans. Aut. Control*, vol. 54, no. 3, pp. 512-527, Mar. 2009.
36. A. H. Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2020-2040, June 2005.
37. I. Maric and R. D. Yates, "Cooperative multihop broadcast for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1080-1088, Aug. 2004.
38. V.I. Morgenshtern, H. Bolcskei, "Crystallization in large wireless networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3319-3349, Oct. 2007.
39. R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and spacetime signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099-1109, Aug. 2004.
40. R. Narasimhan, "Delay-limited throughput of cooperative multiple access channels with hybrid-ARQ," in *Proc. ISIT'2008*, Toronto, Canada, July 2008.
41. T. Nechiporenko, K. T. Phan, C. Tellambura, and H. H. Nguyen, "Capacity of Rayleigh fading cooperative systems under adaptive transmission," *IEEE Trans. Wireless Commun.*, to appear in 2009.
42. A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68-73, Oct. 2004.
43. T. C.-Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in a cooperative cellular network," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 328-339, Feb. 2007.
44. T. K. Phan, T. Le-Ngoc, S. A. Vorobyov, and C. Tellambura, "Power allocation in wireless relay networks: A geometric programming based approach," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM'2008*, New Orleans, LA, USA, Nov.-Dec. 2008.
45. T. K. Phan, L. B. Le, S. A. Vorobyov, and T. Le-Ngoc, "Centralized and distributed power allocation in multi-user wireless relay networks," in *Proc. IEEE Inter. Conf. Commun., ICC'2009*, Dresden, Germany, June 2009.
46. T. K. Phan, T. Le-Ngoc, S. A. Vorobyov, and C. Tellambura, "Power allocation in wireless multiuser relay networks," *IEEE Trans. Wireless Commun.*, to appear in 2009.
47. K.T. Phan, L.B. Le, S.A. Vorobyov, and T. Le-Ngoc, "Power allocation and admission control in multi-user relay networks via convex programming: Centralized and distributed schemes," *EURASIP J. Wireless Commun. and Networking*, to appear in 2009.
48. A. Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1264-1273, May 2005.
49. A. Scaglione, D. Goeckel, and J. N. Laneman, "Cooperative communications in mobile ad-hoc networks: Rethinking the link abstraction," *IEEE Signal Processing Mag.*, vol. 23, no. 5, pp. 18-29, Sept. 2006.
50. A. Scaglione and Y. W. Hong, "Opportunistic large arrays: Cooperative transmission in wireless multihop ad hoc networks to reach far distances," *IEEE Trans. Sig. Proc.*, vol. 51, no. 8, pp. 2082-2092, Aug. 2003.

51. A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927-1938, Nov. 2003.
52. A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939-1948, Nov. 2003.
53. B. Sirkeci-Mergen, A. Scaglione, and G. Mergen, "Asymptotic analysis of multi-stage cooperative broadcast in wireless networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5166-5166, Nov. 2006.
54. B. Sirkeci-Mergen and A. Scaglione, "On the power efficiency of cooperative broadcast in dense wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 497-507, Feb. 2007.
55. B. Sirkeci-Mergen and A. Scaglione, "Randomized space-time coding for distributed cooperative communication," *IEEE Trans. Signal Proc.*, vol. 55, no. 10, pp. 5003-5017, Oct. 2007.
56. R. Tannious and A. Nosratinia, "Spectrally-efficient relay selection with limited feedback," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1419-1428, Oct. 2008.
57. V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.
58. V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456-1467, July 1999.
59. I. E. Telatar, "Capacity of multi-antenna Gaussian Channels," *European Trans. Telecommunications*, vol. 10, no. 6, pp. 585-595, Nov.-Dec. 1999.
60. S. Wei, D. L. Goeckel, and M. C. Valenti, "Asynchronous cooperative diversity," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1547-1557, June 2006.
61. L. Xiao, T. E. Fuja, J. Kliewer, and D. J. Costello, Jr., "A network coding approach to cooperative diversity," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3714-3722, Oct. 2007.
62. L. Xiao, M. Johansson, and S. P. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1136-1144, July 2004.
63. Y. Yao, X. Cai, and G. B. Giannakis, "On energy efficiency and optimum resource allocation of relay transmissions in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2917-2927, Nov. 2005.
64. B. Zhao and M. C. Valenti, "Practical relay networks: A generalization of hybrid-ARQ," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 7-18, Jan. 2005.
65. Y. Zhao, R. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks: Optimal power allocation versus selection," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3114-3123, Aug. 2007.
66. L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073-1096, May 2003.
67. P. L. Zhifeng, Z. Lin, E. Erkip, and S. Panwar, "Cooperative wireless communications: A cross-layer design," *IEEE Commun. Mag.*, vol. 13, no. 4, pp. 84-92, Aug. 2006.
68. H. Zhu and G. Cao, "rDCF: A relay-enabled medium access control protocol for wireless ad hoc networks," in *Proc. IEEE INFOCOM'05*, Miami, USA, Mar. 2005.