Fixed and Random Access Cooperative Networks

Alejandro Ribeiro

Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA
Email: aribeiro@ece.umn.edu

Georgios B. Giannakis

Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA
Email: georgios@ece.umn.edu

Rich scattering of electromagnetic waves propagating through physical environments generates complex interference patterns. As such patterns go through maxima and minima, large variations in energy adversely affect wireless reception and thus deteriorate error probability performance of wireless communication systems. By providing multiple channels with independent (or at least uncorrelated) variations in time, frequency, and/or space, *diversity* techniques offer well-appreciated countermeasures mitigating such (so called fading) effects. With the deployment of multiple antennas effecting space diversity we create copies of the transmitted signal either at the receiver, at the transmitter, or both. In time or frequency (a.k.a. Doppler or multipath) diversity systems, we exploit the natural property of wireless channels to vary over time or frequency. The benefits of diversity are significant. In a typical (wireline) additive white Gaussian noise (AWGN) channel the error probability decays exponentially as the received signal-to-noise-ratio (SNR) increases; that is, error effects decrease as $e^{-\mathrm{SNR}}$. A wireless Rayleigh fading channel, however, exhibits errors decaying as SNR^{-1}. A κth-order diversity channel entails κ uncorrelated channels and exhibits error probability which decreases as $\mathrm{SNR}^{-\kappa}$. Needless to say the gap between the exponential decay in wireline channels and the inversely linear decay in wireless channel is enormous. Considering that for sufficiently large κ the $\mathrm{SNR}^{-\kappa}$ and $e^{-\mathrm{SNR}}$ functions are not very different, the value of diversity is clear: it can close the error performance gap between wireline and wireless channels.

Spatial and time-frequency diversity systems are at opposite ends of a deployment cost versus reliability curve. Spatial diversity is reliable but comes with hardware cost. Time-frequency diversity on the other hand exploits natural phenomena that may or may not be present in a particular link and is thus less reliable even if it comes for free when available. User *cooperation* is an alternative form of diversity which aims to strike a balance in this curve by providing diversity more reliable than natural time-frequency variations yet without requiring deployment of additional antennas. The basic idea is to have single-antenna terminals share information and cooperate in relaying it to intended destinations. If properly designed, cooperative protocols involving κ terminals can achieve κth-order diversity relying on relatively inexpensive software modifications of existing wireless protocols.

Since its introduction [12, 13, 21, 22], researchers in signal processing, wireless communications, and information theory have contributed major advancements to explore and realize the potential of cooperative networks. The main purpose of this tutorial is to cover recent developments in user cooperation for multiple access (MA) over fixed as well as random access (RA) channels. User cooperation was originally developed for point-to-point links and this is still the best setup to explain the basic concepts involved. These are covered in Section 1 where different cooperation strategies are outlined.

Even though a point-to-point link provides a simple setup to introduce the basic principles of cooperative communications, it is not until one considers multipoint links that the challenges of implementing user cooperation are exposed. To illustrate these challenges, we consider three classes of *fixed* cooperative multiple access (MA) protocols in Section 2. The opportunistic multipath (OM) class offers simple repetition protocols which build on the idea that repetitions of a source signal by cooperating relay nodes can be viewed as a form of multipath, since the destination cannot (and need not) differentiate between passive signal reflections and active cooperative repetitions [18]. OM capitalizes on the advantages of statistically orthogonal MA channels in dealing with frequency selectivity to realize an efficient cooperative protocol which alleviates the bandwidth loss associated with the original relay schemes (Section 2.1). Two-phase approaches constitute a second class of cooperative protocols which begin with a low-power transmission phase to disseminate information to nearby users, and follow up with a cooperative retransmission phase to reach the destination [20]. By exploiting the spatial separation among active users, a shared channel can be used in the first phase to minimize the bandwidth increase required to implement cooperation (Section 2.2). The third class of cooperative protocols entails schemes implementing multisource cooperation [23, 26], where users collaborate to create a distributed (e.g., convolutionally) coded transmission which enables diversity order equal to the number of active users (Section 2.3).

The last half of the paper is devoted to very recent results in user cooperation for wireless *random* access (Section 3). Considering that in RA networks users decide to transmit at random, only a few out of the total number of transmitters are active at any given time; thus, transmission hardware resources are inherently underutilized in wireless RA networks. We will see how user cooperation can exploit these resources to gain in diversity, without draining additional energy from the network and without bandwidth expansion [19]. This intuitively reasonable notion is reinforced if we take into account that the number of temporarily idle users increases with the size of the network. Building on this observation, we demonstrate that as the network size increases, there is an increasing diversity advantage to be exploited leading to a limiting scenario in which the throughput of cooperative RA over wireless fading channels approaches that of an equivalent system operating over a wireline AWGN channel.

1. Single source cooperation (SSC)

The core idea behind user cooperation is to create a virtual antenna array (VAA) for transmission by means of data sharing between users. With reference to Figure 1, consider source S_1 ($S2$) sending a data packet $\mathbf{d}_1$ ($\mathbf{d}_2$) to destination D_1 (D_2) through the wireless Rayleigh flat fading channel $h(S_1, D_1)$ [$h(S_2, D_2)$]. Due to the broadcast nature of the wireless channel, $\mathbf{d}_1$ transmitted by S_1 is not only received by D_1 but also by S_2 and D_2 through corresponding channels $h(S_1, S_2)$ and $h(S_1, D_2)$. Thus, if we let S_2 repeat the signal received from S_1 and vice versa, both destinations receive two independent copies of $\mathbf{d}_1$ and $\mathbf{d}_2$. Forgetting for a moment the channel $h(S_1, S_2)$ between sources, D_1 receives data from a 2×1 multiple input single output (MISO) channel which is capable of providing second-order diversity [2].

Even though similar, there are important differences between VAAs and MISO systems with multiple colocated antennas. One difference is that wireless terminals are halfduplex, and as such they cannot transmit and receive over the same frequency at the same time. This practical limitation is rooted in the need to isolate transmitter and receiver in order to avoid feedback from the transmitter to the receiver radio-frequency (RF) front end. If the terminal size is not enough to provide spatial isolation, this has to be achieved in time and/or frequency. The implication is that cooperation protocols have to follow a scheme like the one depicted in Figure 1 in which we have a slot assigned to S_1's transmission, a second slot assigned to S_2's, and a third slot for the cooperative transmission of the other terminal's data. Comparing this scheme with space-time codes [2] we recognize that different from MISO channels the diversity advantage of VAAs comes at the price of bandwidth increase. It is worth noting that this does not necessarily imply a penalty in communication rate, because the decrease in the amount of forward error correction (FEC) and/or number of retransmissions required can compensate for the bandwidth expansion [21].

A second difference is that in VAAs we cannot ignore the channel $h(S_1, S_2)$ between sources. To appreciate its effects, let $\hat{\mathbf{d}}_1$ denote S_2's estimate of $\mathbf{d}_1$ and consider the signals received by the destination D_1:

$$
\begin{aligned}
\mathbf{y}_{11} &= \sqrt{P}h(S_1, D_1)\mathbf{d}_1 + \mathbf{w}_{11}, \\
\mathbf{y}_{12} &= \sqrt{P}h(S_2, D_1)\hat{\mathbf{d}}_1 + \mathbf{w}_{12},
\end{aligned}
\tag{1}
$$

where $\mathbf{w}_{11}$ and $\mathbf{w}_{12}$ denote AWGN terms and P is the transmitted power. It is a surprising result that if D_1 uses a maximum ratio combiner (MRC) for estimating $\mathbf{d}_1$ as ($*$ stands for conjugation and $\|x\|$ for the magnitude of x)

$$
\hat{\mathbf{d}}_1^{\mathrm{MRC}} = \arg\min_{\mathbf{d}_1} \left\| h^*(S_1, D_1)\mathbf{y}_{11} + h^*(S_2, D_1)\mathbf{y}_{12} - \sqrt{P}\Big[|h(S_1, D_1)|^2 + |h(S_2, D_1)|^2 \Big]\mathbf{d}_1 \right\|,
\tag{2}
$$

then the diversity order of this two-branch VAA is only one. The reason for the lack of diversity in this so-called decode and forward (DF) strategy is that the VAA error probability is dominated by the error probability in the link $S_1 \to S_2$.

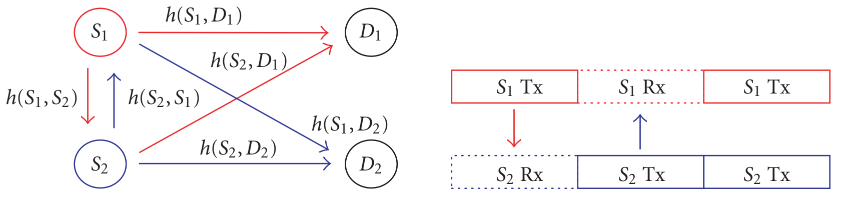

FIGURE 1: Source terminals S_1 and S_2 cooperate in transmitting to their respective destinations D_1 and D_2 by creating a distributed virtual antenna array (VAA).

While DF does not achieve diversity, three alternative strategies do achieve this goal:

(S1) *Selective forwarding (SF)*: instead of always repeating $\mathbf{d}_1$, S_2 will repeat the packet only if it is successfully decoded, that is, if $\hat{\mathbf{d}}_1 = \mathbf{d}_1$. This strategy is more complex than DF because it requires FEC decoding followed by a cyclic redundancy code (CRC) check to detect possible errors at S_2.

(S2) *Amplify and forward (AF)*: a seemingly simple alternative is to let S_2 amplify the analog-amplitude signal received from S_1. That is, the signal $\mathbf{y}_{21} = h(S_1, S_2)\mathbf{d}_1 + \mathbf{w}_{21}$ received by S_2 is transmitted after amplification as $A\mathbf{y}_{21}$. The amplification factor satisfies

$$A^2 = \frac{P}{P\,|\,h(S_1, S_2)\,|^2 + N_0}, \tag{3}$$

so that the power of the signal transmitted by S_2 is equal to P.

(S3) *Cooperative (C) MRC*: while the strategies (S1) and (S2) require operations at the cooperating terminal, a different approach is to adopt DF at the cooperating terminal but use a weighted version of the MRC demodulator in (2)

$$\hat{\mathbf{d}}_1^{\mathrm{CMRC}} = \arg\min_{\mathbf{d}_1} \|\alpha_{11}\mathbf{y}_{11} + \alpha_{12}\mathbf{y}_{12} - \sqrt{P}[\alpha_{11}h(S_1, D_1) + \alpha_{12}h(S_2, D_1)]\mathbf{d}_1\|. \tag{4}$$

By properly selecting α_{11}, and α_{12} as functions of $h(S_1, D_2)$, $h(S_2, D_2)$ and $h(S_1, S_2)$ the so-called C-MRC in (4) can be shown to achieve second-order diversity [27].

Each of the strategies (S1)–(S3) has its own merits. SF is the simplest one from the perspective of the destination but strains the digital processor at the cooperating terminal; also, even if the packet is not correctly decoded, there is still some information about $\mathbf{d}_1$ in the signal received at the cooperator that is not conveyed to the destination. When the link between sources ($S_1 \to S_2$) is expected to be much better than the links between sources and destination ($S_1, S_2 \to D_1$), S_2 will almost always correctly decode $\mathbf{d}_1$ making SF the method of choice for this case. AF requires minimal processing at the cooperating terminal, but necessitates storage of the analog-amplitude received signal thus straining memory resources. AF is appealing when the cooperating terminal is located close to the destination so that the link from the cooperating terminal to the destination ($S_2 \to D_1$) is strong and the link

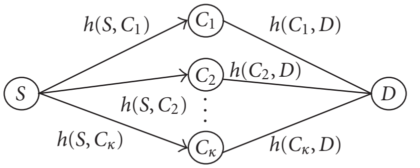

Figure 2: Multibranch cooperation.

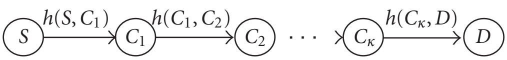

Figure 3: Multihop cooperation.

$S_1 \to S_2$ is comparable to the link $S_1 \to D_1$. Use of C-MRC for decoding DF relayed signals is the simplest strategy from the perspective of the cooperating terminal. Its drawback is that the channel realization $h(S_1, S_2)$ has to be transmitted to the destination since it is needed to compute α_{11} and α_{12}. If this can be accomplished by transmitting a few bits the overhead is not significant.

Pairwise cooperation can be generalized to groups of terminals. For a group of κ cooperating terminals we can build a protocol using any of the strategies (S1)–(S3) to achieve κth-order diversity. This may not be always the best approach considering that in cooperative networks—sometimes also referred to as relay networks—there is a tradeoff between multibranching (see Figure 2) and multihopping (see Figure 3). In multihopping, the source packet is relayed through a cascade of cooperating terminals; while not providing diversity, this approach saves energy by exploiting the smaller pathloss between cooperators as compared to the pathloss from source to destination. In multibranching, the packet is relayed to κ cooperators that retransmit the packet to the destination; this provides diversity but does not benefit from pathloss reduction. The configuration offering desirable tradeoffs in a general network is a combination of multihop and multibranch cooperation [17].

Remark. We have introduced only simple concepts of SSC necessary to study cooperation in multiple fixed and random access channels. Among topics we did not cover due to space limitations is the aforementioned bandwidth penalty VAAs incur relative to MISO systems with co-located antennas. A possible remedy is resorting to (e.g., turbo) coded cooperation whereby the source transmits the first subblock of the code, the cooperating terminal decodes the signal using only this first subblock and, if successful, transmits the second subblock of the turbo code. This does not incur bandwidth expansion to implement cooperation but requires coding at the relays which expands bandwidth, even though the latter is arguably needed anyways [9]. An additional issue is the use of coherent versus noncoherent reception. The use of noncoherent modulation in cooperative networks and its diversity benefits are reported in [5, 28]. Fundamental performance limits of cooperative links are closely related to the capacity of the relay channel, the evaluation of which remains an open problem in information theory [6]. It has been shown that the bandwidth penalty of cooperative protocols is not inherent to the relay channel but is due to the use of repetition coding [3]. In the low-power regime, achievable rates and optimum resource allocation issues for the relay channel have been studied in [4, 29]. For the Gaussian relay channel it is also known that DF and AF relay strategies can be outperformed by a quantize and forward (QF) scheme, whereby the cooperating relay forwards a quantized version of the source signal [11].

2. Cooperation in multiple access channels

In wireless MA channels, a group of users $\mathcal{J} = \{U_j\}_{j=1}^{J}$ communicates with an access point (AP) through independent Rayleigh fading channels $h(U_j, \mathrm{AP})$. For the purpose of our discussion, we differentiate between deterministically and statistically orthogonal MA techniques. To illustrate their differences consider the uplink of a code division multiple access (CDMA) system where the jth user spreads its L-bit data block $\mathbf{d}_{U_j} := \{d_{U_j}(l)\}_{l=0}^{L-1}$ with a T-chip code $\mathbf{c}_{U_j} := \{c_{U_j}(t)\}_{t=0}^{T-1}$ to construct the transmitted packet $\mathbf{x}_{U_j} := \{x_{U_j}(t)\}_{t=0}^{T-1}$. Defining the spreading gain as $S := T/L$, we have that

$$x_{U_j}(Sl+s) = \sqrt{P(U_j)}\, d_{U_j}(l) c_{U_j}(Sl+s). \tag{5}$$

We will use the notation $\mathbf{x}_{U_j} = \mathbf{d}_{U_j} \circ \mathbf{c}_{U_j}$ to represent the spreading operation in (5). Transmission of $\mathbf{x}_{U_j}$ requires S times more bandwidth than transmission of $\mathbf{d}_{U_j}$. Depending on the properties of the inner product $\langle \mathbf{c}_{U_{j_1}}, \mathbf{c}_{U_{j_2}} \rangle$ we have the following MA techniques.

(T1) *Orthogonal MA*: by selecting codes such that $\langle \mathbf{c}_{U_{j_1}}, \mathbf{c}_{U_{j_2}} \rangle = 0$ for $j_1 \neq j_2$—for example, short symbol-periodic Walsh-Hadamard codes—data of different users are transmitted through orthogonal channels. In the absence of multipath and asynchronism, the AP can perfectly separate users while error performance in demodulating each user's data is determined by the associated receive SNR

$$\gamma_j = \frac{SE\left[\,|h(U_j, \mathrm{AP})|^2\right] P_j}{N_0}, \quad j = 1, \ldots, J, \tag{6}$$

where $E(\cdot)$ denotes expectation, P_j the transmitted power and N_0 the noise power. We stress that when using orthogonal spreading sequences we must have $J \leq S$, since the number of spreading codes coincides with the spreading gain. Also, what we say here about CDMA also applies to time (T) and frequency (F) division multiple access (DMA).

(T2) *Statistically orthogonal MA*: here we require the spreading codes to be orthogonal in the statistical sense; that is, $E[\langle \mathbf{c}_{U_{j_1}}, \mathbf{c}_{U_{j_2}} \rangle] = 0$ for $j_1 \neq j_2$. In this case, the AP separates users only on the average and the associated error performance is determined by the signal-to-noise-plus-interference ratio (SINR)

$$\gamma_j = S \frac{E\left[\,|h(U_j, \mathrm{AP})|^2\right] P_j}{\sum_{i=1,\, i\neq j}^{J} E\left[\,|h(U_i, \mathrm{AP})|^2\right] P_i + N_0}, \quad j = 1, \ldots, J. \tag{7}$$

A distinct property of (T2) relative to (T1) is that the number of spreading codes available is much larger than S; for example, there are 2^S codes if we work with binary spreading sequences. This is not to say that we can afford $J \gg S$ since in this case the SINR would be too low. The relevant observation here is that the error probability performance depends on the SINR, but is not affected by the number of spreading codes used [8].

Another feature of practical MA networks is that users are not always active. In voice cellular systems, for example, ratios of active to idle users are typically larger than 10. This enables architectures in which temporarily idle users are available to serve as relays for temporarily active ones. In the following sections we describe two protocols that exploit idle users and one protocol that can capitalize on cooperation of active users alone.

2.1. *Opportunistic multipath*

With each active user $U_j \in \mathcal{J}$ we associate two idle terminals I_{j_1}, I_{j_2} capable of decoding U_j's data and relaying the information to the AP. For simplicity, we focus on a reference source $U_j \equiv U_1$, and set $I_1 \equiv I_{11}$ and $I_2 \equiv I_{12}$. As usual, time is divided into slots during which a frame is transmitted and the two terminals I_1, I_2 take turns in repeating the frames corresponding to odd and even time slots. Specifically, during time slot 0, U_1 transmits the data frame $\mathbf{d}_{U_1}(0)$ spread by the pseudo-noise (PN) spreading code $\mathbf{c}_{U_1}$. During the same time slot, I_1 listens to this transmission that is going to repeat in the next time slot 1, but with spreading code $\mathbf{c}_{I_1}$. Being in transmit mode during slot 1, I_1 misses the frame $\mathbf{d}_{U_1}(1)$, but this frame is received by I_2, which in turn retransmits it in time slot 2 using the code $\mathbf{c}_{I_2}$. This process continues while the transmission lasts. In general, for the $(2i)$th and the $(2i+1)$st time slots, the blocks transmitted by U_1, I_1, and I_2 are

$$\begin{aligned} \mathbf{x}_{U_1}(2i) &= \mathbf{d}_{U_1}(2i) \circ \mathbf{c}_{U_1}, & \mathbf{x}_{U_1}(2i+1) &= \mathbf{d}_{U_1}(2i+1) \circ \mathbf{c}_{U_1}, \\ \mathbf{x}_{I_1}(2i) &= \mathbf{0}, & \mathbf{x}_{I_1}(2i+1) &= \hat{\mathbf{d}}_{U_1}(2i) \circ \mathbf{c}_{I_1}, \\ \mathbf{x}_{I_2}(2i) &= \hat{\mathbf{d}}_{U_1}(2i-1) \circ \mathbf{c}_{I_2}, & \mathbf{x}_{I_2}(2i+1) &= \mathbf{0}, \end{aligned} \tag{8}$$

where $\mathbf{x}_{U_1}$ is the block transmitted from U_1, $\mathbf{x}_{I_i}$ the one from I_i, $i = 0, 1$, $\mathbf{d}_{U_1}(i)$ stands for the frame at time slot i, and $\hat{\mathbf{d}}_{U_1}(i)$ is estimate of $\mathbf{d}_{U_1}(i)$ using the SF rule (S1). Recall that when using SF, I_1 and I_2 forward received packets only when correctly decoded. The important observation here is that every three time slots, proper despreading allows one to recover three data blocks $\{\mathbf{d}_{U_1}(2i), \mathbf{d}_{U_1}(2i+1), \mathbf{d}_{U_1}(2i+2)\}$ directly from the source and three data blocks $\{\hat{\mathbf{d}}_{U_1}(2i-1), \hat{\mathbf{d}}_{U_1}(2i), \hat{\mathbf{d}}_{U_1}(2i+1)\}$ through the cooperating terminals; and by sliding this 3-slot window we obtain two independent copies of each data block. This implies that diversity of order two becomes available without consuming extra time or frequency slots compared with a noncooperative link between U_1 and the AP [18].

While the cooperative protocol in (8) applies to any MA technique, the *spectral efficiency* claim—that is, the fact that we do not need extra bandwidth with respect to a noncooperative MA channel—is valid only with statistically orthogonal spreading codes. Indeed, if we use deterministic spreading sequences, when implementing (8) we require three times as many codes and correspondingly three times as much bandwidth. It is only because there are up to 2^S codes satisfying $E[\langle \mathbf{c}_{U_{j_1}}, \mathbf{c}_{U_{j_2}} \rangle] = 0$ that we can implement (8) without increasing bandwidth. The repetition rule in (8) can be viewed as the introduction of intentional (opportunistic) multipath. Indeed, from the AP's perspective there is no difference between a (passive) reflection off a scatterer and an active repetition by a cooperating terminal. It is thus not surprising that CDMA can effect user cooperation without bandwidth penalty, since this is precisely the property that has made CDMA so popular in handling frequency-selective multipath wireless channels. The OM protocol can be generalized to involve multiple sources and multiple cooperating terminals per source-destination link in order to effect a diversity order equal to the number of cooperators plus 1. Also, a multicode alternative achieving the same diversity advantages while requiring a single cooperator per source-destination link can be devised as detailed in [18].

2.2. Two-phase cooperation

When using deterministically orthogonal (as opposed to statistically orthogonal) MA, the repetition protocol in (8) requires twice as much bandwidth as the one required for noncooperative transmission. This penalty stems from the need to use separable channels for the source and relay transmissions. An alternative approach is to exploit the spatial separation between source-cooperator pairs by assigning a *shared* channel for all the source-to-cooperator communications. This idea appeared first in the two-phase cooperative protocol of [20]. In the first phase, the source terminals $\{U_j\}_{j=1}^{J}$ transmit their information to preassigned cooperating terminals $\{C_j\}_{j=1}^{J}$, and in the second phase the pair (U_j, C_j) conveys the packet to the AP. A possible implementation is to consider a slot of duration T_1 for the J simultaneous communications $U_j \rightarrow C_j$ followed by J time slots of duration T_2 dedicated to the transmission from each (U_j, C_j) pair to the AP. In the J slots assigned to the $(U_j, C_j) \rightarrow$ AP transmission, we have the potential for diversity that we can enable with a space-time code to avoid further bandwidth expansion. Meanwhile, we have rate reduction equal to $JT_2/(JT_2 + T_1)$; but if we select $T_1 \ll JT_2$, then the spectral efficiency is almost equal to 1.

While a smaller T_1 leads to higher spectral efficiency, it poses challenges to the $U_j \rightarrow C_j$ communications. For this reason, the goal of two-phase cooperation protocols is to optimally balance these conflicting requirements. Letting each user transmit the packet $\mathbf{x}_j = \mathbf{F}_j\mathbf{d}_j$, for every set of matrices $\mathcal{F} := \{\mathbf{F}_j\}_{j=1}^{J}$ we have a corresponding set of rates $\mathcal{R}(\mathcal{F}) = \{R_j(\mathcal{F})\}_{j=1}^{J}$ between U_j and C_j. Given a power constraint $P_{\max}$ per user, a rate maximizing approach is to choose the set $\mathcal{F}^*$ so that the set $\mathcal{R}^*$ is optimal in the Pareto sense [20]

$$\mathcal{F}^* = \arg\max \mathcal{R}(\mathcal{F}), \qquad \text{subject to } \operatorname{tr}\left[\mathbf{F}_j\mathbf{F}_j^{\mathcal{H}}\right] \leq P_{\max}, \quad j = 1, \ldots, J. \tag{9}$$

Pareto optimality implies that any other set of matrices $\mathcal{F} \neq \mathcal{F}^*$ results in at least one user having a smaller rate $R_j < R_j^*$. In this sense, it is the maximum fair rate since an increase in R_j in any other rate allocation comes at the expense of a rate decrease for some other user $U_i \neq U_j$.

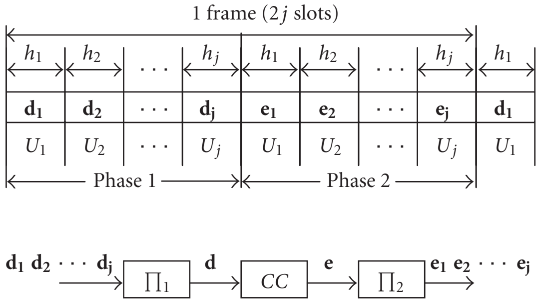

FIGURE 4: DCC-MSC with TDMA frame structure for J active users.

The optimization problem in (9) is difficult to solve in general; and, even if we find a solution, its implementation requires coordination among the sources in $\mathcal{J}$. This may not be feasible in certain situations, motivating a reformulation of (9) as a competitive noncooperative game. In this noncooperative game, the pairs $\{(U_j, C_j)\}_{j=1}^{J}$—"players"—have conflicting interests and compete for the resources through self-optimization. The optimal solutions are stable Nash equilibrium (NE) points. At these points, given the power allocation of other players, namely, the pairs $\{(U_i, C_i)\}_{i=1,\, i\neq j}^{J}$, each player pair (U_j, C_j) does not obtain any rate increase by changing its own power. While the NE is suboptimal in the Pareto sense, it has the advantage of being achievable by decentralized algorithms. Conditions for the existence of NE and a game achieving it based on successive waterfilling can be found in [20].

2.3. Multisource cooperation (MSC)

The OM protocol offers desirable tradeoffs in rate versus SINR (and thus error) performance by capitalizing on idle users. An alternative framework which is also flexible in trading off rate for diversity (and thus error) performance is the so-termed multisource cooperation (MSC) [23, 26]. While MSC protocols can also take advantage of idle users, their distinct feature relative to OM is that they allow for cooperation among active users only.

The general setup for MSC involving J users is depicted in Figure 4. Each transmission frame consists of two phases: direct transmission and relaying. The first phase includes J slots, during which the J sources transmit their information blocks $\{\mathbf{d}_j\}_{j=1}^{J}$, each with block length L. By the end of this phase, the destination as well as sources (during their listening slots) have received all messages from all sources. Although the SNRs of user-user pairs are typically higher than those of user-destination pairs (since users in the same cluster are close), it is possible that only a subset of users can correctly receive all messages $\{\mathbf{d}_j\}_{j=1}^{J}$. Let $\mathcal{D}$ denote such a subset. During the relay phase, each user in $\mathcal{D}$ reencodes these messages $\{\mathbf{d}_j\}_{j=1}^{J}$ jointly using a systematic code with rate R_c, and then transmits part of the parity check bits during its time slot. Systematic and parity check bits are, therefore, transmitted in the first and second phases, respectively. At the destination, joint decoding of multiple messages is performed. When idle users are willing to serve as relays, certain parity check bits can be transmitted by the relays.

This encoding process of each user in $\mathcal{D}$ is further detailed in the lower part of Figure 4. At the end of the first phase, each user first feeds the multiple data blocks $\{\mathbf{d}_j\}_{j=1}^{J}$ into an interleaver, Π_1, the role of which is to equally protect messages from multiple sources. The interleaved sequence $\mathbf{d}$ is encoded by a rate R_c systematic convolutional code (CC) [26]. With systematic bits ignored, the parity check bits $\mathbf{e}$ are then fed to a second interleaver Π_2. The role of the second interleaver is to distribute parity check bits to different channels in order to effect a high diversity order. The jth source or relay then transmits the jth segment of the parity check bits $\mathbf{e}_j$ during its time slot, during the second phase. If user U_j is not in $\mathcal{D}$, then $\mathbf{e}_j$ is not transmitted and the resultant code is a punctured CC with $\mathbf{e}_j$ missing.

With proper design of the interleavers involved in phase 2, it is possible to prove that the diversity order of this CC based MSC protocol does not depend on the cardinality of the set $\mathcal{D}$; but is given by the $\min(d_{\min}, 1 + \lfloor J(1 - R_c)\rfloor)$ where $d_{\min}$ denotes the free distance of the CC, J is the number of active users, and R_c is the code rate [26]. This expression for the diversity order shows that full diversity J cannot be achieved when distributed CC-based MSC relies on code rates $R_c > 1/J$. Existing results on linear block codes can be directly borrowed to search for codes with maximum $d_{\min}$, and thus enable the highest possible diversity order.

Relative to SSC with repetition coding, MSC based on distributed CC can also enhance coding gains because relay transmissions are coded across time and space. As each source in MSC is served by multiple relays, for the same spectral efficiency, MSC can achieve higher diversity gains than SSC. And since each relay serves multiple sources simultaneously, for the same diversity, MSC can offer higher spectral efficiency than SSC. To further improve data rates, distributed trellis coded modulation has been proposed recently to replace CC in MSC [26].

3. Cooperation in random access networks

Instead of agreeing on a fixed channel allocation, RA networks let users transmit at random contending to reach the common AP. Letting users transmit packets independently with probability p implies that successful packet delivery depends not only on the physical channel but on how many other users decided to transmit, leading to a packet delivery probability function $P_d(p)$. In turn, this implies that an average of $\mu(p) := pP_d(p)$ packets are delivered per time slot. A remarkable property of RA networks is that despite the lack of coordination among users, it is possible to achieve a reasonable average number of packets delivered by selecting p so as to achieve $\mu := \max[\mu(p)]$. In, for example, the slotted Aloha protocol, $\mu = 0.36$ which means that about 1 packet is delivered every 3 time slots.

The random nature of RA dictates that in any time slot only a fraction of potential users is active, the others having their transmissions deferred. But since only a few out of the total number of transmitters are active at any given time, transmission hardware resources are inherently underutilized in wireless RA networks. It is thus reasonable to expect that user cooperation can exploit these resources to gain a diversity advantage and intuition suggests that user cooperation appears to be a form of diversity well matched to RA.

Cooperation can be implemented with a two-phase RA protocol (see Figure 5). In the first phase, "phase-A," users send a packet with just enough power to be correctly decoded by nearby peers; while in the second phase, "phase-B," the set of peers that successfully decoded this packet transmit cooperatively with power sufficient to reach the AP. If we manage to balance conflicting power requirements, what happens in phase-A is that nearby users decode the original packet while the power received at the destination is negligible. This implies that (i) phase-A users do not interfere severely with concurrent phase-B nodes; and (ii) phase-A locally disseminates information so that subsequent phase-B transmissions are enriched with a certain degree of cooperative diversity.

From this high-level description, one may expect benefits from cooperation, but a proper assessment of these benefits requires studying the following conflicts.

(I1) Since cooperation clearly adds complexity to an RA network it is important to determine whether diversity provides a substantial advantage in terms of increasing throughput.

(I2) Cooperation is almost a synonym of coordination, but a cooperative random access protocol has to be faithful to the RA premise of minimal coordination between users.

(I3) Ideally, we would prefer the power during phase-A to be negligible and the number of phase-B cooperators to be very large. But as the phase-A power decreases, so does the number of phase-B cooperators.

(I4) The benefits of diversity can be compromised by the excess bandwidth and/or power required to implement cooperation.

In the rest of the section, we work with the model in Figure 5 where a set of J users, $\mathcal{J} = \{U_j\}_{j=1}^{J}$, communicates with an AP in a spread spectrum (SS) RA network. User j and its position in a coordinate system centered at the AP will be denoted by U_j, with these positions considered random and uniformly distributed within a circle of radius R. User positions are further assumed to be independent.

The link between any two users $h(U_{j_2}, U_{j_1})$ is modeled as a flat Rayleigh fading channel. The average power received at U_{j_1} from a source U_{j_2} transmitting with power $P(U_{j_2})$ is given by an exponential pathloss model

$$P(U_{j_2} \longrightarrow U_{j_1}) = \frac{\xi P(U_{j_2})}{||U_{j_1} - U_{j_2}||^{\alpha}}, \tag{10}$$

with $\|U_{j_1} - U_{j_2}\|$ denoting the 2-norm of the vector $U_{j_1} - U_{j_2}$ and ξ and $\alpha \leq 2$ constants. As a special case, the power received at the AP from U_{j_2} is $P(U_{j_2} \rightarrow \text{AP}) = \xi P(U_{j_2})/\|U_{j_2}\|^{\alpha}$.

Each of the J users has an infinite-length buffer for storing L-bit fixed length packets that arrive at a rate of λ packets per packet duration. The packet arrival processes are identically distributed (i.d.), not necessarily independent. The L bits of each packet are spread by a factor S (a.k.a. spreading gain) to construct a transmitted packet of $T := SL$ chips. Spreading is implemented using a long PN sequence $\mathbf{c} := \{c(t)\}_{t\in\mathbb{Z}}$ with period $\mathcal{P}$. If $\mathbf{d}_{U_j} := \{d_{U_j}(l)\}_{l=0}^{L-1}$

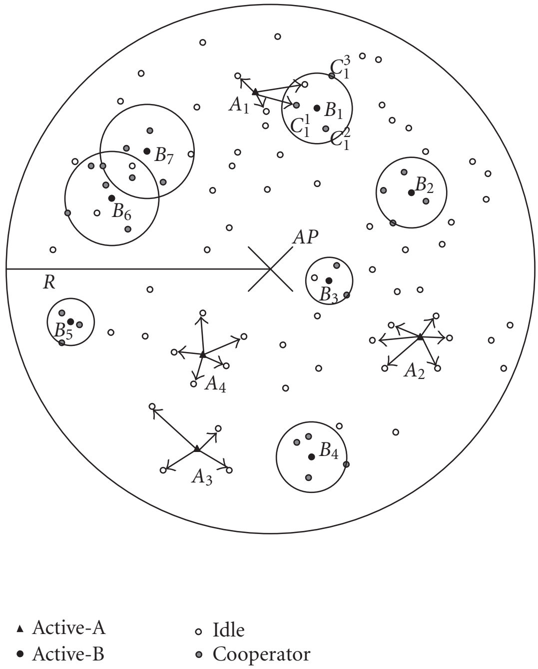

Figure 5: A cooperative RA network snapshot.

denotes a data packet of user U_j, and $\mathbf{x}_{U_j} := \{x_{U_j}(t)\}_{t=0}^{T-1}$ the corresponding transmitted packet, we have (c.f. (5))

$$x_{U_j}(Sl+s) = \sqrt{P(U_j)}d_{U_j}(l)c(Sl+s-\tau_{U_j}), \tag{11}$$

where $\mathbf{c}$ is a common long PN sequence *shared* by all users, τ_{U_j} is a user-specific shift applied to $\mathbf{c}$, and $P(U_j)$ is the power transmitted by node U_j.

Before detailing the RA protocols considered in this section a word is due on throughput. Strictly speaking, $\mu(p)$ is the departure rate of the RA system and throughput $\eta(p)$ is defined as the maximum arrival rate λ yielding stable queues. Under conditions that are valid in the subsequent discussion Loynes' theorem [15] asserts that $\eta(p) = \mu(p)$. Note that $\eta(p, J)$ is also a function of the number of users J. One is typically interested in the maximum stable throughput (MST) defined as $\eta_{\max}(J) = \max_p\{\eta(J, p)\}$ and achieved at $p = p_{\max}$. Here, we will be interested in the asymptotic MST that we define as $\eta_\infty = \lim_{J\to\infty}\eta_{\max}(J)$, and interpret as the average number of packets transmitted per unit time in a system with a very large number of users; see also [7, 25].

3.1. Noncooperative SSRA and the role of diversity

Let us begin by describing a noncooperative SSRA system. In such a system, each user transmits a packet constructed according to (11) with probability p. If the packet is successfully decoded by the AP, this is acknowledged through a common feedback channel. The resultant SSRA protocol is defined by the following rules.

(R0) The period of the PN sequence is $\mathcal{P} = T$.

(R1) Time is divided into slots, each comprising T chip periods. If users decide to transmit, they do so at the beginning of a slot.

(R2) Packets are spread for transmission according to (11). The shift τ_{U_j} is selected at random by each user; and $P(U_j) = P_0\|U_j\|^{\alpha}/\xi$ effects average power control so that all users are received at the AP with the same average power P_0 (c.f. (10)).

(R3) If a given user's queue is not empty, the user transmits the first queued packet in the next slot with probability p.

(R4) The AP acknowledges correctly decoded packets through a feedback channel. If an acknowledgment is not received, the packet is placed back in U_j's queue. As usual (see, e.g., [1]) feedback is assumed to be instantaneous and free of errors.

Rule (R1) defines a slotted system and its purpose is to simplify throughput analysis, (R2) effects statistical user separation and power control, (R3) controls the transmission rate, with p adjusted so as to maximize throughput, and (R4) determines the procedure for a packet to leave the system when it is successfully decoded.

Packets in an SSRA network are incorrectly decoded either when two users choose the same PN shift, $\tau_{U_{j_1}} = \tau_{U_{j_2}}$, or when the interference is too high. This motivates a distinction between hard and soft collisions. We say that U_{j_1} experiences a "hard collision" (HC), if $\tau_{U_{j_1}} = \tau_{U_{j_2}}$ for some $j_2 \neq j_1$. Given that U_{j_1} does not experience a hard collision, we say that it experiences a "soft collision" (SC) when the packet is lost due to interference.

Interestingly, throughput is mainly limited by soft collisions—a manifestation of the interference limited nature of SSRA networks. The probability of experiencing a soft collision is determined by the SINR, which in turn depends on the number of active users during the slot under consideration. Indeed, at any given slot the set of users $\mathcal{J}$ is divided into a set of temporarily active users $\mathcal{A} = \{A_j\}_{j=1}^{J_A}$ and a set of temporarily idle ones $\mathcal{I} = \{I_j\}_{j=1}^{J_I}$ with $J_A + J_I = J$. Given the number of active users J_A, we have that the probability of a packet being successfully received by the AP is [19]

$$P_s(J_A) = \left(1 - \frac{1}{T}\right)^{J_A - 1}\left[1 - P_e(\gamma)\right], \quad \gamma = \frac{1}{N_0/P_0 + (J_A - 1)/S}. \tag{12}$$

The first factor in (12) accounts for the HC probability and the second one for the SC probability. The function P_e maps the SINR to packet error probability and is determined by the channel model and the transmission/reception schemes which include the type of modulation, type of receiver, and FEC code.

Diversity manifests itself in changing the function $P_e(\gamma)$ in (12). If we consider different models for the channels $h(U_j, \text{AP})$, then we will have different functions $P_e(\gamma)$ resulting in different packet success rates P_s and respective throughputs.

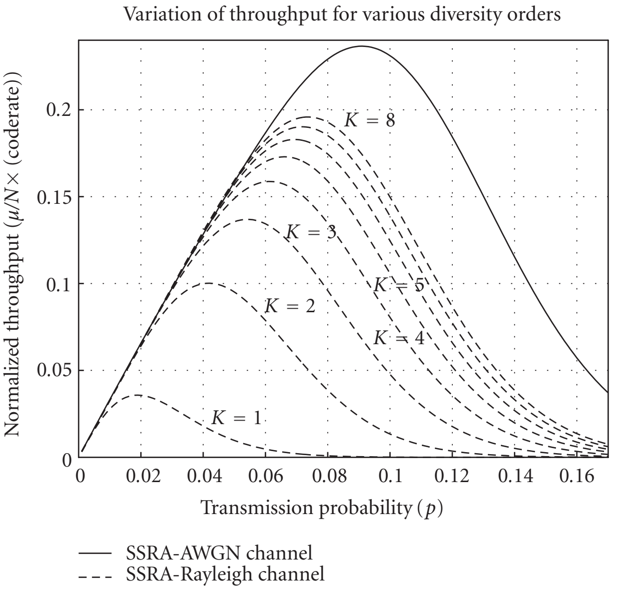

FIGURE 6: High-order diversity closes the enormous gap between the performance of RA over wireless Rayleigh fading channels with respect to wireline AWGN channels ($J = 128$, $S = 32$, $L = 1024$, 215/255 BCH code capable of correcting $t = 5$ errors).

For a fixed FEC code, we consider three different models for the channel $h(U_j, \mathrm{AP})$, corresponding to an AWGN channel, Rayleigh fading channels and diversity channels. The best possible scenario is when $h(U_j, \mathrm{AP})$ is a deterministic constant (AWGN channel). A better model for the wireless environment, however, is a Rayleigh fading channel where $|h(U_j, \mathrm{AP})|^2$ is random Rayleigh distributed variable. The throughput over wireless channels can be increased with diversity techniques. A channel with κth-order diversity is one in which the AP decodes κ copies received through uncorrelated Rayleigh channels, $\{h_k(U_j, \mathrm{AP})\}_{k=1}^{\kappa}$, with each $|h_k(U_j, \mathrm{AP})|^2$ Rayleigh distributed, yielding the aggregate channel model $|h(U_j, \mathrm{AP})|^2 := \sum_{k=1}^{\kappa} |h_k(U_j, \mathrm{AP})|^2$ when MRC is used.

For each of these channels, we depict in Figure 6 the normalized throughput as a function of the transmission probability p. It comes as no surprise that the MST over a wireless (Rayleigh) channel is miserable, being almost an order of magnitude smaller than the MST of the wireline AWGN channel. This sizeable gap can be closed by diversity techniques, as hinted by the twofold increase observed with 2nd-order diversity and the close-to-AWGN MST enabled with 8th-order diversity. Eventually, as κ keeps increasing the κth-order diversity channel approaches an AWGN channel. Thus, if we denote the throughput over an AWGN channel as η^G, and the throughput over an ∞-order diversity channel as η^∞, we can

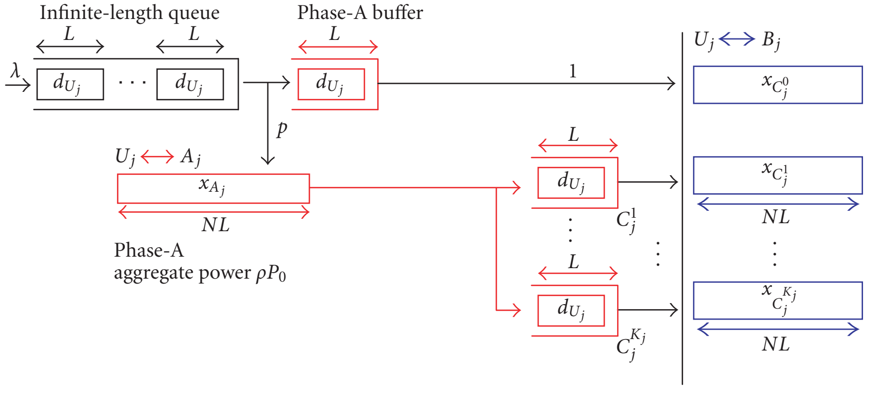

FIGURE 7: OCRA is a two-phase cooperative RA protocol.

write

$$\eta^{\infty} = \eta^{G}. \tag{13}$$

Meaning that diversity has the potential to yield wireline-like throughputs in wireless RA channels.

3.2. *Opportunistic cooperative random access*

The previous section established that diversity offers the potential for a large throughput increase in RA networks; the point is, of course, whether and how this diversity can be enabled. Since users transmit at random in RA networks, a number of users remain idle over any given slot. The opportunistic cooperative random access (OCRA) protocol introduced in this section exploits the good reception opportunities of this large set of idle users. OCRA is a two-phase protocol defined by the following operating conditions; see also Figure 7.

(O0) Let κ be an upper bound on the achievable diversity. The period of the PN code $\mathbf{c}(t)$ is chosen to be $\mathcal{P} = \kappa T + 1$.

(O1) At the beginning of each slot, if U_j's queue is not empty, U_j enters phase-A with probability p and moves the first packet in the queue, $\mathbf{d}_{U_j} := \{d_{U_j}(l)\}_{l=0}^{L-1}$, to a single packet buffer that we term phase-A buffer.

(O2) *Phase-A*: when in phase-A, we say that $U_j \leftrightarrow A_j$ is an active-A user and transmits a packet $\mathbf{x}_{A_j} := \{x_{A_j}(t)\}_{t=0}^{T-1}$ spread according to (11) with PN-shift and power given by

$$\tau_{A_j} = 0, \qquad P(A_j) = \frac{\rho P_0 \|A_j\|^{\alpha}}{\xi}, \tag{14}$$

with $\rho \in (0, 1)$. The PN shift is deterministically chosen and the transmission power is so that the packet is received at the AP with fractional power ρP_0. A random integer,

$\tau_{B_j} \sim \mathcal{U}[1, T]$, uniformly chosen over $[1, T]$ is included in the packet header to coordinate PN-shifts during phase-B.

(O3) *Phase-A handshake*: any idle user I_k that decodes $\mathbf{x}_{A_j}$ becomes a cooperator $I_k \leftrightarrow C_j^k$ and places $\mathbf{d}_{U_j}$ in a single-packet buffer designated for cooperation purposes. This successful decoding is acknowledged to A_j who collects a total of K_j acknowledgments and feeds forward this number K_j to the cooperators. Similar to, for example, [10, 16], this handshake is assumed to be instantaneous and error free.

(O4) User U_j enters phase-B in the slot immediately after entering phase-A.

(O5) *Phase-B*: let $\mathcal{C}_j = \{C_j^k\}_{k=0}^{K_j}$ be the set of cooperators comprising $C_j^0 = B_j \leftrightarrow U_j$ and the K_j cooperators recruited in phase-A. Each of the C_j^k transmits the packet $\mathbf{d}_{U_j}$ spread according to (11) using

$$\tau_{C_j^k} = \tau_{B_j} + \tau_k T, \qquad P(C_j^k) = \frac{P_0}{K_j + 1} ||C_j^k||^{\alpha} / \xi, \tag{15}$$

with τ_{B_j} the number received in phase-A's packet header, and the integer $\tau_k \sim \mathcal{U}[0, \kappa - 1]$. Power scaling is implemented so that the total received power at the destination is P_0. Let $\mathbf{x}_{C_j^k} := \{x_{C_j^k}(t)\}_{t=0}^{T-1}$ denote these transmitted packets.

The number of cooperators K_j is termed the "*cooperation order*" of B_j and the number κ_j of PN shifts chosen by at least one cooperator is called the "*diversity order*" of B_j.

(O6) *AP acknowledgment*: the AP acknowledges successful reception of the superimposed phase-B packets corresponding to B_j through a feedback channel. If an acknowledgment is not received, the packet $\mathbf{d}_{B_j}$ is placed back in B_j's queue; cooperators discard this packet in any event.

(O7) *Idle operation*: when not transmitting, $U_j \leftrightarrow I_j$ correlates the received signal with $\{c(t)\}_{t=0}^{T-1}$ to detect phase-A packets transmitted by other (nearby) users.

By rule (O2), U_j becomes the active-A user A_j and transmits $\mathbf{x}_{A_j}$ with low power so as to reach nearby users while not interfering with the AP (if $\rho \ll 1$). Phase-B is defined by rule (O5) in which the packet is transmitted with κ_j-order diversity by $U_j \leftrightarrow B_j$ plus K_j cooperators corresponding to the K_j idle users that successfully decoded U_j's transmission during phase-A. The opportunistic nature of the protocol manifests in the random diversity order κ_j which depends on the number K_j of cooperators recruited and the random selection of shifts τ_k.

Rules (O1), (O4), and (O6) govern the transition between idle and active-A/B states. Users move from idle to active-A with probability p as per (O1); after entering phase-A, the user moves deterministically to phase-B in the first upcoming slot (O4), and back to idle in the second one (O6). A lost packet does not alter this transition but only determines whether the packet is put back in queue or not. Also, (O6) dictates that cooperators discard B_j's packet regardless of the transmission success.

Rules (O0), (O3), and (O7) guarantee logical consistency of the protocol. Rule (O0), provides sufficient number of PN shifts to enable the selection rule in phase-B (c.f. (15)), (O3) disseminates the number of cooperators recruited to allow proper power scaling during phase-B as required by (15), and (O7) ensures that idle users are listening for phase-A packets.

Users in OCRA are temporarily divided into four classes; see also Figure 5. A set of N_A "active-A" users, $\mathcal{A} = \{A_j\}_{j=1}^{N_A}$, operating in phase-A of their transmission trying to reach nearby users; a set of N_B active-B users, $\mathcal{B} = \{B_j\}_{j=1}^{N_B}$, communicating their packets to the AP; and N_I idle users $\mathcal{I} = \{I_j\}_{j=1}^{N_I}$ that either have empty queues or decided not to transmit. The fourth class of users, comprises the sets of cooperators $\mathcal{C}_j = \{C_j^k\}_{k=0}^{K_j}$ associated with each active-B user B_j. The set $\mathcal{C}_j$ contains $C_j^0 = B_j$, and the K_j users that correctly decoded B_j's phase-A packet in the previous slot.

3.3. OCRA's throughput

Similar to the noncooperative SSRA in Section 3.1, OCRA's throughput can be derived from the packet success probability. As with noncooperative SSRA, packets in OCRA are not correctly decoded either when two users choose the same PN shift, $\tau_{B_{j_1}} = \tau_{B_{j_2}}$ (HC), or when the interference is too high (SC). It is not difficult to see that OCRA's HC probability is equal to that of noncooperative SSRA. The SC probability, though, depends on the number of active-A and active-B users,

$$P_s(J_A, J_B) = \left(1 - \frac{1}{T}\right)^{J_B - 1} \left[1 - P_e(J_A, J_B)\right]. \tag{16}$$

Different from (12), the function $P_e(J_A, J_B)$ in (16) is difficult to express in closed form since the packet error probability depends on the random cooperation and diversity orders of each individual user.

Instead of trying to find expressions for $P_e(J_A, J_B)$ we can take an asymptotic approach and relate the throughput of OCRA with the throughput of noncooperative SSRA as the number of users grows large. It can be proved that for an arbitrary diversity order κ the asymptotic throughput of OCRA, $\eta_\infty^{\text{OCRA}}(\kappa)$, and the asymptotic throughput of noncooperative SSRA operating over a κ-order diversity channel, η_∞^κ, are equal [19]; that is,

$$\eta_\infty^{\text{OCRA}}(\kappa) = \eta_\infty^\kappa. \tag{17}$$

Thus, a network of single-antenna terminals cooperating according to rules (O0)–(O7) is equivalent to a network of κ-antenna terminals communicating without cooperation according to rules (R0)–(R4). As we can see from Figure 6, this yields a significant increase in throughput when we compare OCRA with noncooperative SSRA operating over Rayleigh fading channels.

The relation in (17) can be obtained by formalizing the following argument. As the number of users grows large ($J \to \infty$) we let the phase-A power vanish ($\rho \to 0$) and consider an increasing function $K(J) \to \infty$. It can be proved that for appropriate convergence ratios ($\rho^{2/\alpha} J/K \to \infty$) the cooperation orders K_j of all active-B users are greater than K with high probability ($\Pr\{K_j \geq K, \ \forall j\} \to 1$) [19]. This establishes that every active-B user is receiving cooperation by a large number of users; moreover, as long as the convergence rates of ρ and $K(J)$ are adequate, the cooperation order K_j becomes arbitrarily large while the active-A transmitted power becomes arbitrarily small. Consequently, the seemingly conflicting requirements of recruiting an infinite number of cooperators with a vanishingly small power *are* compatible as $J \to \infty$ implying that very large diversity orders are achievable by OCRA.

After recalling that high-order diversity is tantamount to an AWGN channel (see also [19]), it is fair to state that with κ sufficiently large

$$\eta_{\infty}^{\mathrm{OCRA}}(\kappa) \approx \eta_{\infty}^{G}, \tag{18}$$

which is an accurate statement of our intuitive assertion that user cooperation can improve throughput to the point of achieving wireline-like throughput in a wireless RA environment. This is a subtle but significant difference relative to point-to-point user cooperation in fixed access networks, where the diversity advantage typically comes at the price of bandwidth expansion.

Other interesting properties of OCRA are the following.

(P1) *Average power constraint*. It can be proved that cooperation is limited to nearby idle users and accordingly the total transmitted power by any active communication is

$$\sum_{k=0}^{K_j} P(C_j^k) \approx (K_j+1)\frac{P_0}{K_j+1}||B_j||^{\alpha}/\xi = P_0||B_j||^{\alpha}/\xi. \tag{19}$$

Comparing (19) with noncooperative SSRA, we observe that the average transmitted power in noncooperative SSRA is equal to OCRA's phase-B power (c.f. rule (R2)). The sole power increase is due to the phase-A power used to recruit cooperators yielding the relation $P^{\mathrm{OCRA}}(U_j) \approx (1+\rho)P^{\mathrm{SSRA}}(U_j)$ between the power required by OCRA and noncooperative SSRA. Since $\rho \to 0$, we deduce that OCRA enables high order diversity with a small increase in average transmitted power.

(P2) *Maximum power constraint*. A maximum power constraint $P(U_j) \leq P_{\max}$ determines the AP's coverage area, since power control dictates that $\|U_j\|^{\alpha} \leq (\xi P_{\max}/P_0) := R_c^{\alpha}$. But the power in OCRA is contributed by K_j cooperators and accordingly

$$R_c^{\mathrm{OCRA}} = (K_j)^{1/\alpha} R_c^{\mathrm{SSRA}}. \tag{20}$$

This increase in coverage stems from the fact that in OCRA users are transmitting less power during more time.

(P3) *Fairness*. At a given slot, active-B users increase their throughput by "borrowing" power from cooperating terminals, raising fairness concerns. If the random processes involved are ergodic so that time averages equal ensemble averages; then, the time average of the power that any U_j user spends cooperating with other users coincides with the time average of power that other users spend cooperating with U_j. OCRA is thus a fair protocol in ergodic settings. Strictly speaking practical networks are nonergodic, but behave like ergodic ones when observed over long periods.

(P4) *Unslotted OCRA*. Slotted operation requires packet level synchronization that can be avoided with unslotted operation. An unslotted version of OCRA is developed in [19] whose relation with unslotted SSRA is the same as the relation between the corresponding slotted versions in (18).

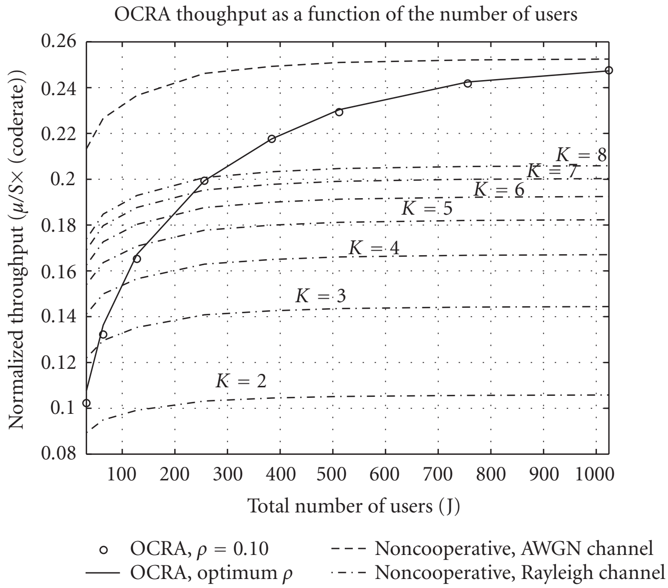

FIGURE 8: The MST for $J = 128$ is 2/3 the MST of SSRA over an AWGN channel ($\kappa = 10$, $S = 32$, $L = 1024$, 215/255 BCH code capable of correcting $t = 5$ errors).

3.4. Simulations

The question we address in this section is how large the number of users should be to achieve a significant throughput increase. For that matter, we refer to Figure 8 where we depict OCRA's MST, $\eta_{\max}^{\text{OCRA}}$, as a function of the number of users J in a network with spreading gain $S = 32$, packet length $L = 1024$, and a 215/255 BCH code capable of correcting $t = 5$ errors used for FEC. A quick inspection of Figure 8 reveals that convergence to AWGN throughput is rather slow since for J as large as 512 there is still a noticeable gap. Notwithstanding, the throughput increase is rather fast; for $J = 64$ there is a threefold throughput increase ($\eta_{\max} = 0.04$ if the channel is Rayleigh), and for $J = 128$ OCRA's MST is 2/3 of the MST achieved by noncooperative SSRA over an AWGN channel. Thus, while collecting the full diversity advantage requires an inordinately large number of users, OCRA can collect a significant percentage of it in moderate size networks, with a ratio $J/S \approx 4$.

Similar conclusions can be drawn from the simulation with $J = 128$ users depicted in Figure 9. For this case study, we show throughput and average diversity as a function of the transmission probability p. For the range of probabilities close to the MST, OCRA's throughput remains between the curves for 4th- and 5th-order diversities, consistent with the fact that the average degree of cooperation that users receive is between 4 and 5.

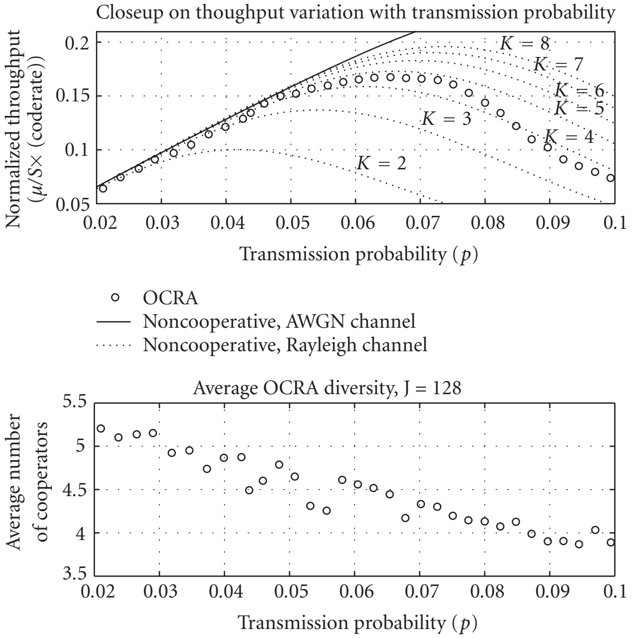

Figure 9: OCRA's throughput is between the throughput of 4th- and 5th-order diversities, consistent with the fact that the cooperation order is between 4 and 5 ($\rho = 0.01, J = 128$, the same as in Figure 8).

4. Conclusions

In this tutorial we presented recent developments in user cooperation protocols for fixed and random multiple access channels. Even though the concept of user cooperation in point-to-point links has reached reasonable maturity, important challenges in cooperative networking remain largely unresolved.

We outlined three cooperative multiple access protocols adhering to as many different paradigms. The opportunistic multipath protocol regards cooperation as a form of (intentionally induced) multipath and resorts to well-known tools for dealing with frequency selective channels. In two-phase cooperation protocols, the challenge is to design an effective method for sharing the common channel used for delivering information between sources and their respective cooperators. In multisource cooperation, the aim is to achieve high-order diversity by constructing a distributed convolutionally coded packet. All three approaches have advantages and shortcomings. The advantage of opportunistic multipath is spectral efficiency, and it is thus well suited for heavily loaded networks. Two-phase cooperation is a versatile approach that can be of interest in asymmetric networks requiring different rate/quality links to different users. Multisource cooperation is an alternative if we do not want to exploit temporarily idle users, but implementation may be more complex than the previous alternatives.

Cooperation in random access networks looks particularly promising given the inherent underutilization of radio resources and the natural match between cooperation and random access. We outlined the OCRA protocol which we showed capable of effecting a considerable throughput increase with respect to equivalent noncooperative random access protocols. Testament to this significant advantage is the fact that as the number of users in the network increases, OCRA's throughput over Rayleigh fading links approaches that of the corresponding SSRA protocol over AWGN links, without an energy penalty. Accordingly, OCRA offers the potential for rendering a wireless RA channel equivalent to a wireline one from the throughput perspective. The price paid is a modest increase in complexity (and therefore cost) of the baseband circuitry. Simulations demonstrated that certain asymptotic claims bear practical relevance to networks with realistic size.

Acknowledgments

Work in this paper was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

References

[1] N. Abramson, "Multiple access in wireless digital networks," *Proceedings of the IEEE*, vol. 82, no. 9, pp. 1360–1370, 1994.

[2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[3] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4152–4172, 2005.

[4] X. Cai, Y. Yao, and G. B. Giannakis, "Achievable rates in low-power relay links over fading channels," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 184–194, 2005.

[5] D. Chen and J. N. Laneman, "Modulation and demodulation for cooperative diversity in wireless systems," to appear in *IEEE Transactions on Wireless Communications*.

[6] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, 1979.

[7] G. Dimic, N. D. Sidiropoulos, and R. Zhang, "Signal processing and queuing tools for MAC-PHY cross-layer design," *IEEE Signal Processing Magazine*, vol. 21, pp. 40–50, 2004.

[8] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatley, "On the capacity of a cellular CDMA system," *IEEE Transactions on Vehicular Technology*, vol. 40, pp. 303–312, 1991.

[9] M. Janani, A. Hedayat, T. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 362–371, 2004.

[10] K. Joseph and D. Raychaudhuri, "Throughput of unslotted direct-sequence spread-spectrum multiple-access channels with block FEC coding," *IEEE Transactions on Communications*, vol. 41, no. 9, pp. 1373–1378, 1993.

[11] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, 2005.

[12] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.
[13] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415–2425, 2003.
[14] R. Lin and A. P. Petropulu, "Cooperative transmission for random access wireless networks," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 19–23, Philadelphia, Pa, USA, March 2005.
[15] R. Loynes, "The stability of a queue with non-independent interarrival and service times," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 58, pp. 497–520, 1962.
[16] R. K. Morrow Jr. and J. S. Lehnert, "Packet throughput in slotted ALOHA DS/SSMA radio systems with random signature sequences," *IEEE Transactions on Communications*, vol. 40, no. 7, pp. 1223–1230, 1992.
[17] A. Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1264–1273, 2005.
[18] A. Ribeiro, X. Cai, and G. B. Giannakis, "Opportunistic multipath for bandwidth-efficient cooperative networking," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, pp. 549–552, Montreal, Canada, May 2004.
[19] A. Ribeiro, N. Sidiropoulos, G. B. Giannakis, and Y. Yu, "Achieving wireline random access throughput in wireless networking via user cooperation," submitted to *IEEE Transactions on Information Theory*, http://www.ece.umn.edu/users/aribeiro/research/pubs.html.
[20] G. Scutari, S. Barbarossa, and D. Ludovici, "Cooperation diversity in multihop wireless networks using opportunistic driven multiple access," in *Proceedings of the IEEE International Workshop on Signal Processing Advances for Wireless Communications*, pp. 170–174, Rome, Italy, June 2003.
[21] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - part I: system description," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927–1938, 2003.
[22] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - part II: implementation aspects and performance analysis," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1939–1948, 2003.
[23] O. Shalvi, "Multiple source cooperation diversity," *IEEE Communications Letters*, vol. 8, no. 12, pp. 712–714, 2004.
[24] M. K. Tsatsanis, R. Zhang, and S. Banerjee, "Network-assisted diversity for random access wireless networks," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 702–711, 2000.
[25] B. Tsybakov and V. Mikhailov, "Ergodicity of a slotted Aloha system," *Problems of Information Transmission*, pp. 301–312, 1980, (translated from Russian original in *Problemy Peredachi Informatsii*, vol. 15, no. 4, pp. 72–87, 1979).
[26] R. Wang, W. Zhao, and G. B. Giannakis, "Multi-source cooperative networks with distributed convolutional coding," submitted to *IEEE Transactions on Wireless Communications*.
[27] T. Wang, A. Cano, and G. B. Giannakis, "Efficient demodulation in cooperative systems using decode-and-forward relays," in *Proceedings of the 39th Asilomar Conference on Signals, Systems and Computers*, Monterey, Calif, USA, October-November 2005.
[28] T. Wang, Y. Yao, and G. B. Giannakis, "Non-coherent distributed space-time processing for multiuser cooperative transmissions," to appear in *IEEE Transactions on Wireless Communications*, see also *Proceedings of Globecom Conference*, St. Louis, Mo, USA, November-December 2005.
[29] Y. Yao, X. Cai, and G. B. Giannakis, "On energy efficiency and optimum resource allocation of relay transmissions in the low-power regime," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2917–2927, 2005.