Cellular Relaying Networks: State of the Art and Open Issues

Jian Zhao and Armin Wittneben
Swiss Federal Institute of Technology (ETH) Zurich
Communication Technology Laboratory, CH-8092 Zurich, Switzerland
Email: {zhao, wittneben}@nari.ee.ethz.ch

Abstract

Future wireless communication systems are expected to provide more stable and higher data rate transmissions in the whole network. Recent research shows that it is advantageous to use relay stations in wireless networks in order to satisfy that need. But due to the limited power supply of mobile users, it may be difficult for the users to act as relays. Thus placing additional fixed relay stations to assist the communications is an attractive way for the cellular network solution. We present an overview of the state of the art for relaying networks, focusing on cellular relaying networks. Furthermore, we discuss in this paper the open issues to be solved in the future.

I. INTRODUCTION

The concept of relaying has been a well-known technique for transmission of signals over very long distance. Research results on relay channels can be found in the 1970's [1], [2]. Recently, due to the explosion of interests in ad hoc networks and the research activity on the fourth generation (4G) wireless communication systems, the concept of relay networks was again brought to the mainstream of research in mobile communication academy and industry [3], [4]. That is because of the following reasons: First, the envisioned transmission rate of 4G systems is several orders of magnitude higher than that of the current mobile communication systems. With the same power level, the transmitted power per bit will be just a small fraction of that in current systems. It is unlikely that 4G mobile systems can cover the same service area without additional infrastructure. Second, the spectrum to be released for 4G systems will be far above the spectrum of current mobile systems. The problem of pathloss will prohibit the base station from communicating with users far away [5]. One way to overcome those problems is to use relays to extend the coverage area in cellular networks.

In addition to the previously mentioned reasons, recent research has shown that the deployment of relays can enhance the capacity of wireless networks and improve the overall data rate. By introducing relay stations, additional transmission paths are created. Signals carrying the same information are transmitted over independently fading channels and can be combined at the receiver. Thus multihop communication can realize spatial diversity in distributed antenna systems such as cellular networks to improve the stability of radio links between the base station and users. The integration of relay networks into conventional cellular networks is perhaps the most promising wireless architecture in the years to come. Due to the limited power supply of mobile users, it is difficult for the users to act as relays in future wireless networks. One feasible solution is to place fixed relay nodes on the top of lampposts or high buildings where good wireless connection with the base station is possible. They can have sufficient power supply and can be equipped with powerful signal processing hardware as well as multiple antennas.

This paper aims to present the state of the art for cellular relaying networks. The remainder of the paper is organized as follows: in section II, we describe the basic and the most commonly used relaying protocols that combat the fading channel induced by multipath propagation in wireless networks. We discuss why it is advantageous to introduce relays to improve radio link quality. Related work can be found in [6] and [7]. In section III, a cooperative transmission scheme for cellular downlink is discussed. We assume the base station is equipped with multiple antennas. Using space-time coding techniques, the benefits of multiple-input-multiple-output (MIMO) communication can be achieved [8]. In section IV, a strategy to combine the benefits of cellular and *ad hoc* networks using relaying stations is discussed [9]. Such a network can successfully divert traffic from one "hot spot" cell to another cell which is not congested. In section V, we discuss an example of the cellular relaying network called wireless media system (WMS), which was proposed in [5]. Such a network divides the cells into pico cells with the help of fixed relays, which can be easily integrated into next generation cellular networks. Some open problems will be discussed in section VI.

II. Cooperative Diversity

Diversity techniques have been widely used to increase the reliability of radio links in wireless communication. The basic idea of diversity is to transmit signals carrying the same information over different paths and combine the independently fading replicas at the receiver. Different techniques have been proposed to exploit diversity in time, frequency and space with colocated antennas. Recently, cooperative diversity has been proposed to realize spatial diversity in distributed antenna systems using relays stations. Such schemes exploit the broadcast property of wireless communication.

A possible two-hop relay channel model is depicted in Fig. 1. The source and destination can be a mobile user or the base station. Because a terminal's transmit signal is typically 100-150dB above its received signal and there is no sufficient electrical isolation between the transmit and receive circuitry up to now, the relay node cannot transmit and receive signals using the same channel. They have to be separated in time, frequency or code. We just consider half-duplex operations as depicted in Fig. 2. In the first time slot n, the source sends its signal $x_s[n]$. For a flat fading channel, the received signal at the relay and the destination can be expressed as

$$\begin{aligned} y_r[n] &= a_{s,r}x_s[n] + z_r[n] \\ y_d[n] &= a_{s,d}x_s[n] + z_d[n], \end{aligned} \tag{1}$$

where $y_r[n]$ and $y_d[n]$ are the relay and destination received signals, respectively. And in the second time slot $n+1$, the source keeps silent, while the relay transmits signal $x_r[n+1]$ to the destination.

$$y_d[n+1] = a_{r,d}x_r[n+1] + z_d[n+1], \tag{2}$$

where $y_d[n+1]$ denotes the received signal at the destination in the time slot $n+1$. In the above equations, $a_{i,j}$ is the channel coefficient between the transmitter and receiver and z_j is the additive noise at the receiver, where $i \in \{s,r\}$ and $j \in \{r,d\}$. We can model $a_{i,j}$ as independent $\mathcal{CN}(0, \sigma_{i,j}^2)$ random variables and z_j as independent $\mathcal{CN}(0, N_0)$ random variables.

A. *Amplify-and-Forward*

Amplify-and-Forward (AF) relays resemble the traditional analog relays, which transmit an amplified version of the previously received signal, i.e. [6]

$$x_r[n+1] = \beta y_r[n], \tag{3}$$

with the power constraint

$$\beta \leq \sqrt{\frac{P}{|a_{s,r}|^2 P + N_0}}, \tag{4}$$

where P is transmit power constraint for the source and the relay. In order to achieve the highest capacity available, the amplification factor β in (4) should be met with equality. Since the destination receives the same information from the source and the relay at different time slots, the two versions of the received signal can be decoded using the

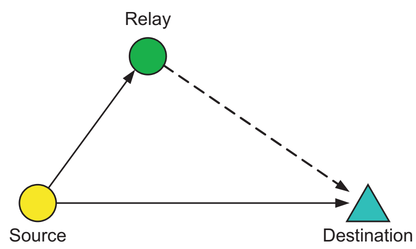

Fig. 1. Two Hop Relay Channel Model

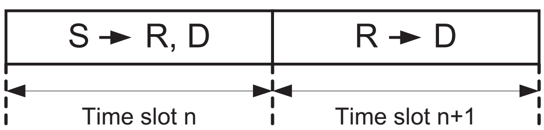

Fig. 2. Half-duplex Transmission

maximum-ratio combining technique at the destination. Such relaying scheme is easy to implement and transparent to the coding and modulation scheme.

The maximum average mutual information and the outage behavior has been analyzed in [6]. The result shows that the outage probability declines proportional to SNR^{-2} in the high SNR region. Thus amplify-and-forward relaying achieves full diversity.

B. Decode-and-Forward

Another relaying scheme is called Decode-and-Forward (DF) relaying. Here, the relay first decodes the received signal from the source and retransmits the decoded symbol to the destination. The retransmitted symbols from the relay can be written as

$$x_r[n+1] = \hat{x}_s[n], \tag{5}$$

where $\hat{x}_s[n]$ is the decoded symbol at the relay.

It has been shown in [6] that such DF relaying scheme does not offer diversity gain for large SNR if we require the relay to fully decode the source message. The outage probability can only decay proportionally to SNR^{-1} in the high SNR region. This is because the channel between the source and the relay limits the performance for the whole system.

C. Selection Relaying

Selection relaying tries to improve the performance of AF or DF relaying by adaptively choosing between the direct transmission path and the relay path. Since the fading coefficient $a_{s,r}$ is known at the relay, we can decide whether it is of advantage to use the relay path. If $|a_{s,r}|$ is below a certain threshold, using the relay will not improve the performance. So we can use the source instead of the relay to transmit in the second time slot using repetition coding. If $|a_{s,r}|$ is large, the relay forwards in the second time slot what it received in the first time slot, using AF or DF scheme. Since we assume that full channel knowledge is known at the receiver but not at the transmitter, we may require the relay to send back one bit to the source indicating whether or not it will transmit in the second time slot. And because neither the source nor the relay knows its channel quality to the destination, it is not possible to decide whether it is more advantageous to choose the source or the relay to transmit in the second time slot if the source-relay channel $|a_{s,r}|$ is large.

The selection relaying scheme offers full diversity because outage is caused either because both $|a_{s,r}|$ and $|a_{s,d}|$ are small or because both $|a_{s,d}|$ and $|a_{r,d}|$ are small. Thus it offers diversity order of two for the three nodes' system. It has been shown in [6] that the performance of selection decode-and-forward relaying is identical to that of previously mentioned amplify-and-forward scheme in the high SNR region.

[10] proposed *Coded Cooperation* scheme which is a variation of the selection decode-and-forward scheme. In such a scheme, the data from the source is divided into blocks augmented with cyclic redundancy check (CRC) bits. Using rate-compatible punctured convolutional (RCPC) coding, the block codes are re-encoded and partitioned into two parts. The first part of N_1 bits is the punctured convolutional code which is a valid code in itself. The N_2 bits of the second part are the bits punctured off in the coding process and they can form a stronger code together with the first part. In the communication process, the source transmits the first part of the code. Upon reception, the relay decode that part and check the CRC code. If the CRC code is correct, the relay forms the second part (punctured bits) and transmits it to the destination. Otherwise the second part will not be transmitted by the relay. CRC code is powerful in checking errors and has been implemented in most communication systems. The coded cooperation scheme achieves excellent performance when the channel between the source and the relay is good.

D. Incremental Relaying

We observe that all the previous schemes are variations of repetition coding, because they transmit the same information in two time slots. That is inefficient for high data rate transmission. Incremental relaying scheme tries to be more efficient in transmission and resembles the automatic-repeat-request (ARQ) protocol. Upon reception in the first time slot, the destination decodes the transmitted signal and check the codes using for example the CRC method. If the decoded code is without error, the destination transmits one bit back to the source and relay indicating the source to transmit the next message. Otherwise the relay transmits what it received in the first time slot using for example the AF scheme, and the destination combines the received signal and decodes again. Such scheme achieves high spectral efficiency and full diversity gain.

III. Cooperative Multiple-Input-Multiple-Output (MIMO) Transmission

For both AF and DF schemes, one user connection requires two orthogonal channels. In order to improve the transmission efficiency, it is mandatory to reuse the relay channels in a cellular system. In [8], the authors proposed one channel reuse scheme for the downlink in a TDMA cellular system. As is depicted in Fig. 3, the base station serves K users, with K' relay stations assisting the transmission. The base station first allocates K time slots for transmission. The users and relays receive dada in their corresponding time slots. In the last time slot, the K' relays retransmit the received information simultaneously. Therefore, each user connection equivalently occupies $1 + 1/K$ time slots instead of 2 time slots. During the relaying slot, the relays generate multi-user interference to other users because the relaying channels are not mutually orthogonal. But if each relay is close enough to its end user, high gains in capacity can also be achieved [11]. In order to manage the interference level, the authors considered a decentralized strategy proposed in [12], which is based on game theory. That strategy tries to maximize the theoretical capacity of the downlink channel under the total energy constraint, including the transmission power from the base station and the relay stations. Thus fair comparison between cooperative and noncooperative transmission schemes can be made in terms of equal total transmitted power.

Typically, the base station can be equipped with multiple antennas. But due to space constraints, Each mobile user normally has only one antenna on the equipment. Let us denote the number of antennas at the base station, at relay station and at user equipment as M, R and N respectively. The authors of [8] also considered to build Distributed Space Time Codes (DSTC) to improve the performance in cellular downlink. DSTC is based on Linear Dispersion Codes (LDC) proposed in [13], which is a special kind of space-time block code and it disperses the transmitted symbols both across space and time. The LDC, $\mathbf{S}$, can be expressed as

$$\mathbf{S} = \sum_{q=1}^{Q} (\mathbf{A}_q \alpha_q + j\mathbf{B}_q \beta_q), \tag{6}$$

where $(\alpha_q + j\beta_q)$ denotes the q-th uncoded data symbol. $\mathbf{A}_q$ and $\mathbf{B}_q$ are $T \times N_{TX}$ matrices, where T is the block length and N_{TX} is the number of transmitting antennas. They are called dispersion matrices (DMs). Each row of $\mathbf{S}$ denotes the symbols associated with each specific antenna, and each column of $\mathbf{S}$ denotes the symbols to be transmitted in each time slot. Different space-time coding techniques, such as V-BLAST and Quasi Orthogonal Designed (QOD) codes [14], can be expressed by (6).

For the AF strategy, each end user receives two versions of the transmitted symbol, one from the base station and the other from the relay stations. DSTC for the system can be constructed in the same way as designing LDC for an $M \times 2N$ MIMO system. Similarly, the DF strategy can be considered equivalent to an $(M + R) \times 2N$ MIMO system. If the relay stations use the same DMs as the base station, such DF is called repetition code approach (DF-RC). If the two kinds of DMs are different, such DF is called unconstrained code approach (DF-UC).

Simulations are performed on both AF and DF strategies with downlink relaying channel reuse. When $(M, R, N) = (2, 1, 1)$, the AF approach for cooperative transmission outperforms the DF approach because the available transmission rate for DF is limited by the source to relay channel. In AF scheme, the V-BLAST type DSTC can perform as a 2×2 MIMO system and double the achievable rate than a non-cooperative system. Those results have been achieved with only one antenna at the user equipment. When $(M, R, N) = (2, 2, 1)$, the DF outperforms the AF

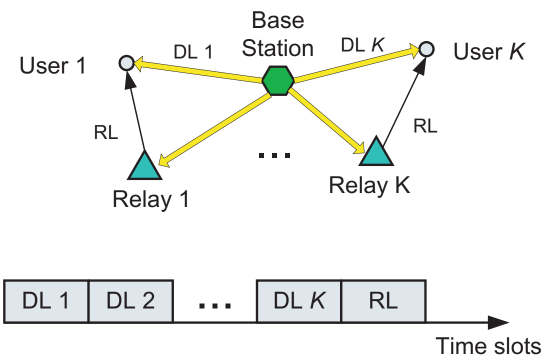

Fig. 3. Cooperative transmission for TDMA cellular downlink

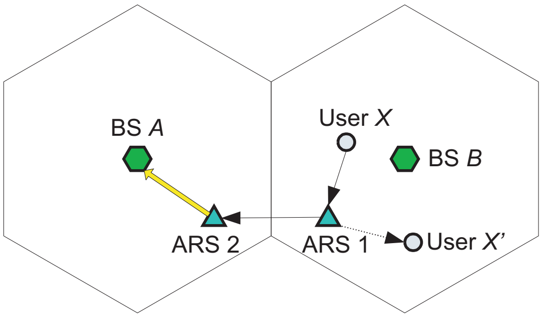

Fig. 4. Primary Relaying

approach. DF-UC approach achieves the best ergodic capacity and it can be implemented by Quasi Orthogonal Designed (QOD) codes.

IV. Integrated Cellular and *Ad Hoc* Relaying Systems

Integrated cellular and *ad hoc* relaying systems (iCAR) was proposed in [9], which tries to integrate the benefits of modern *ad hoc* networks into conventional cellular networks. It tries to alleviate the congestion problems caused by unbalanced traffic and provide interoperability for heterogeneous networks such as *ad hoc* and cellular system using relay stations. The basic idea is to dynamically forward the traffic from one cell to another using *ad hoc relaying stations* (ARS). In [9], the authors assumed that each ARS has two air interfaces, the *cellular channel* for communication with the base station and the *relaying channel* for communication with a mobile host or another ARS. Each mobile host should also have two air interfaces: the cellular channel for communication with the base station and the relaying channel for communication with an ARS. The cellular channel can operate at around 1900 MHz (PCS), and the relaying channel can use the unlicensed band at 2.4 GHz (the ISM band).

Primary Relaying: The simplest relaying strategy is called primary relaying and is depicted in Fig. 4. User *X* is in a congested cell (hot spot). When it tries to establish a connection with its base station (BS) *B* using the cellular channel, the request will be blocked. But user *X* can submit a request to its nearest relaying station, for example ARS 1, using the relaying channel. Upon reception, ARS 1 uses the relaying channel to forward the request to an relaying station (ARS 2) in a noncongested cell nearby. Finally, ARS 2 connects with its BS *A* using the cellular channel. Thus the connection path between user *X* and BS *A* can be established. In the *ad hoc* modes, user 1 can also bypass the base station and establish connection directly with other users (e.g. user *X'*) in the same cell using the relaying channel as depicted in Fig. 4.

Secondary Relaying: If there is no free relaying stations nearby, the primary relaying strategy will not work. Secondary relaying tries to free up an cellular channel for new users by forcing another user who is currently involved in communicating using the cellular channel to switch over to the relaying channel. As is depicted in Fig. 5, user *X* tries to communicate with BS *B*, but there is no free cellular channel available nor is there any free relaying station nearby. User *Y* is currently involved in communicating with the BS *B* using one cellular channel. In order to help the poor user *X*, user *Y* switches over to the relaying channel and establish connection with BS *A* through ARS. User *X* can now connect with BS *B* using the cellular channel previously occupied by user *Y*. Since

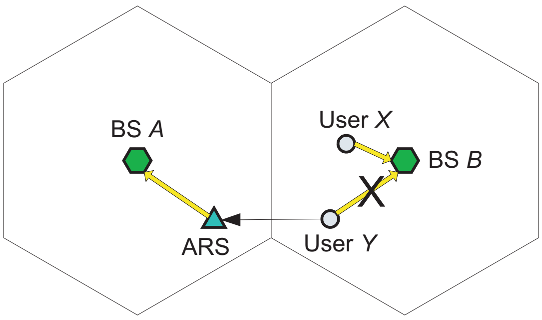

Fig. 5. Secondary Relaying

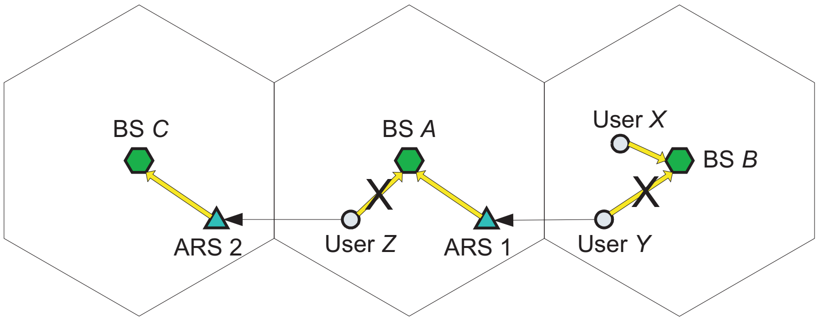

Fig. 6. Cascaded Relaying

it is unlikely that every user can have free relaying stations nearby, the probability of using secondary relaying strategy is larger than that of using primary relaying in a cellular system.

Cascaded Relaying: The previously mentioned two relaying strategies require some noncongested cells nearby. If all the surrounding cells are also ”hot spots”, it is still possible for the new users to establish connections to the base station using cascaded relaying strategy. That scheme is depicted in Fig. 6, which tries to establish one ”relaying path” to a noncongested cell maybe far away by repeatedly using secondary relaying strategy. Once the path is established, one cellular channel will be free in the current cell and the new user can be served.

Simulations show that the blocking/dropping probability can be dramatically reduced even just using the secondary relaying strategy. The call blocking probability falls to 2% if the coverage area of each ARS is 20% of the whole cell.

V. A Prototypical Cellular Relaying Network

A prototypical cellular relaying network, the wireless media system (WMS), has been proposed in [5]. Such a network aims to use fixed relay stations to extend radio coverage area and achieve high spectrum efficiency with the lowest required power. Such an architecture is made up of pico cell base stations, access points and relay stations of around 1 Watt using broadband air interface. An example of WMS is illustrated in Fig. 7.

WMS will provide very high multiplexing data rate of 100-1000 Mb/s for terminals traveling at medium velocity. The candidate spectrum is expected to be beyond 3 GHz, thus allowing the integration of small size equipments (including multiple antennas). The whole system shares an IPv6 based core network and all the functions of the cellular network. In such a system, both access points and relay stations act as the base stations for their own pico cell. Multiple pico cells can be grouped up and form micro cells. In this way, such a network provides very flexible infrastructure and can also be integrated into future generation cellular networks.

VI. Open Issues

As have been discussed above, many new schemes have been proposed in the past few years to improve the performance of cellular networks using fixed relays. But several open issues still need to be solved in the future.

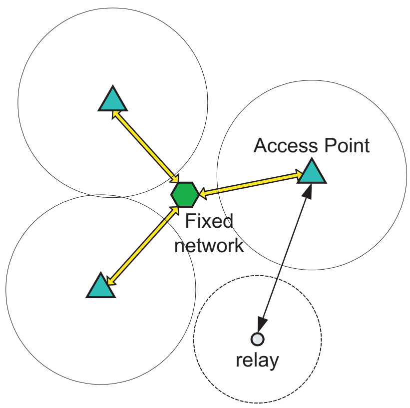

Fig. 7. Wireless Media System

Due to practical constraints, each relay station must transmit and receive signals using orthogonal channels. Otherwise, strong self-interference will be generated. As the number of relays increases in a cell, frequency reuse among different relays may be necessary in order not to cause interference during the relay transmission. Several schemes have been proposed to reuse frequency channels in adjacent cells. But as the number of users in the cell increases, the cell will be split into smaller cells. The interference between the relays and adjacent cells using the same frequency will become large. If power control at the base station is applied, the received power of its users will also decrease. Thus, how to deploy relays in a cell and reuse channels in an optimum way is still an open problem.

The relays can increase the coverage area of the base station. But different users in the same cell may still experience different channel quality. For example, users near the base station can use the direct transmission from the base station as a diversity path, while users on the border of a cell can only receive the signals from the relays. How to guarantee quality of service (QoS) among different users is also an open problem.

VII. Conclusions

This paper provides an overview of the state of the art for cellular relaying networks. Because the envisioned future cellular network cannot cover the same radio range as today's cellular network, it is necessary to deploy relays for radio range extension. And it is maybe the only way to cover those otherwise shadowed area. In addition to that, introducing relays to cellular networks can reduce transmission power for mobile users and offer additional diversity for improving the end-to-end radio links.

We reviewed different relaying protocols in detail. But searching for the optimum relaying scheme will be a long term research area. On the network level, relaying stations can divert congested traffic and reduce the blocking probability. Future cellular relaying protocols may also require joint optimization of cross layer design.

References

[1] E. C. Van der Meulen, "A survey of multi-way channels in information theory," *IEEE Trans. Inform. Theory*, vol. 23, pp. 1–37, 1977.

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

[3] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, pp. 388–404, Mar. 2000.

[4] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-Part I: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[5] R. Pabst, B. Walke, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, pp. 80–89, Sept. 2004.

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

[7] T. E. Hunter and A. Nosratinia, "Performance analysis of coded cooperation diversity," in *Proc. ICC*, vol. 4, (Anchorage, AK), pp. 2688–2692, May 11–15, 2003.

[8] O. Muñoz, A. Agustín, and J. Vidal, "Cellular capacity gains of cooperative MIMO transmission in the downlink," in *Int. Zurich Seminar on Communication (IZS)*, (Zurich, Switzerland), Feb. 2004.

[9] H. Wu, C. Qiao, S. De, and O. Tonguz, "Integrated cellular and *Ad Hoc* relaying systems: iCAR," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 2105–2115, Oct. 2001.

[10] T. E. Hunter and A. Nosratinia, "Cooperative diversity through coding," in *Proc. IEEE Int. Symposium on Inf. Theory*, (Lausanne, Switzerland), p. 220, June 30 – July 5, 2002.

[11] G. Scutari, S. Barbarossa, and D. Ludovici, "Cooperation diversity in multihop wireless networks using opportunistic driven multiple access," in *Proc. 4th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC) 2003*, (Rome, Italy), June 2003.

[12] A. Agustín, O. Muñoz, and J. Vidal, "A game theoretic approach for cooperative MIMO schemes with cellular reuse of the relay slot," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04)*, (Montreal, Canada), May 2004.

[13] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804–1824, July 2002.

[14] A. Agustín, J. Vidal, and A. Pagès-Zamora, "LDC construction with a defined structure," in *IEEE 58th Vehicular Technology Conference, 2003. VTC 2003-Fall*, (Orlando, USA), Oct. 2003.