Interference Suppression with Physical-Layer Network Coding and MIMO for Multi-Way Channels

Flávio S. Brás
Instituto Superior Técnico - Universidade de Lisboa
Portugal

Filipe E. Ferreira, and Francisco A. Monteiro
ISCTE - Instituto Universitário de Lisboa
and
Instituto de Telecomunicações
Portugal
francisco.monteiro@lx.it.pt

António Rodrigues
Instituto Superior Técnico - Universidade de Lisboa
and
Instituto de Telecomunicações
Portugal

Abstract—Orthogonal techniques (such as TDMA – time division multiple access) are traditionally used when a set of terminals wants to exchange data and a relay is required when wireless terminals cannot communicate directly. This paper proposes two schemes to address multi-way channels when a set of terminals wants to exchange data. Both proposed schemes consist of a first phase based on virtual MIMO with lattice reduction-aided (LRA) detection at the relay, establishing a multiple access channel (MAC). Two new strategies are then proposed for the second (multicast) phase; one for single antenna terminals and another for multiple-input multiple-output (MIMO) terminals. The second strategy achieves interference suppression by successively combining physical layer network coding and MIMO detection (with LRA detectors also at the terminals). The paper assesses a 3 terminals network with a relay for which a TDMA-based scheme would require 6 time-slots, while the proposed strategy for MIMO terminals is able to exchange the information between all terminals using 2 time-slots only. This latter scheme is easily scalable to a higher number of terminals, and always able to accomplish the exchange of messages between all of them in just 2 time-slots, although limited by the number of antennas that the terminals can fit and also the MIMO processing complexity they can afford.

Keywords—Multi-way channel; MIMO; interference cancelation; physical-layer network coding.

I. INTRODUCTION

A concept to further enhance the throughput of wireless networks has recently emerged, the so-called physical layer network coding (PLNC). It appears to have been independently proposed by several research groups in 2006 by Zhang, Liew and Lam [1], Popovski and Yomo [2] for the two-way relay channel (TWRC), later developed in [3], and also by Nazer and Gastpar [4] [5], the latter under a distinct framework (named compute and forward). Both concepts exploit what naturally occurs when electromagnetic waves interfere and differ from previous applications of network coding to the TWRC [6].

This paper chiefly deals with interference cancelation when a receiver knows the channel that each of the interfering messages have gone through. Notice that this is also the case in “pure” MIMO spatial multiplexing, but in this paper the concept of successive interference cancelation (SIC) is concatenated with physical layer network coding (PLNC), interpreting PLNC as a first stage of SIC. It should be highlighted that after exploring all the dimensions where orthogonality can exist (time, frequency, spreading codes, and space) the use of non-orthogonal multiple access (NOMA) in conjunction with interference-cancelation techniques is becoming a promising new dimension to separate users [7].

The TWRC is the simplest example to illustrate the PLNC concept. Consider two nodes which cannot communicate directly, typically due to some propagation limitations. Eventually, they establish communication via a relay that can receive and send to both (Fig. 1a). Applying traditional TDMA in the TWRC each terminal takes one time-slot to transmit its message to the relay on its turn and then the relay takes more two time-slots to distribute the messages to the users. Using “pure” network coding at the relay in the TWRC, the number of time-slots can be reduced to three (e.g., [6]). Initially node 1 sends its message to the relay on time-slot one and then node 2 also sends its message to the relay on the second time-slot. Then the relay computes the sum of the messages \( w_1 \oplus w_2 \) and, in the third time-slot, broadcasts this sum (which is received by both nodes). However, because the relay only needs the sum of the signals, the additional concept in physical-layer network coding is that the nodes can simultaneously transmit to the relay (e.g., [1]). This reduces the number of time-slots to only two, as shown in Fig. 2b, which doubles the throughput of the TWRC in comparison to TDMA.

Figure 1. The simplest PLNC setup: the two-way relay channel.
New strategies combining MIMO with PLNC in scenarios that move beyond the traditional TWRC are proposed in this paper. The first proposed strategy involves three time-slots and matched filtering detection after interference cancelation. The second strategy requires two time-slots only while using MIMO terminals in a network with a MIMO relay. One considers a network of terminals wanting to exchange information between all of them. The paper analyses the performance for a three-terminal network but it will be seen that one of the two proposed schemes can be generalized to a larger number of terminals, as long as they can accommodate more antennas and afford the complexity involved.

It is assumed that the nodes are synchronized and share the same frequency band. It is imposed a half-duplex constraint which means that each terminal can either transmit or receive during each of the time-slots. However, it is worth mentioning that this restriction may be waived in the future by applying interference suppression techniques that will open doors to full-duplex relays [8].

II. SYSTEM MODEL

In the two strategies proposed in this paper the uplink or phase 1 (from the terminals to the relay) consists of decentralized transmitters and a MIMO relay, which in the downlink, or phase 2, becomes a MIMO transmitter sending to the decentralized terminals (which can have a single antenna or multiple antennas). Let us consider terminals transmitting with a single transmit antenna (even though they may have more antennas incorporated to later receive – used in phase 2) and a relay receiving the signals using \( N \) antennas, which forms the distributed MIMO system depicted in Fig. 2.

![Figure 2. Distributed MIMO in the uplink phase.](image)

In this distributed MIMO system one considers \( T \) terminals, each of which equipped with \( n \) antennas, and a relay with (with \( N \geq T \)). The baseband received signal is

\[
y = Hx + n
\]

where \( y = [y_1, \ldots, y_N]^T \in \mathbb{C}^{N \times 1} \) is the received signals vector and \( x = [x_1, \ldots, x_N]^T \in \mathbb{C}^{N \times 1} \) is the transmitted symbols vector. The links between each pair of transmit and receive antennas are the elements of the channel matrix \( H \in \mathbb{C}^{N \times T} \) in which its entries \( h_{ij} \) represent the complex coefficient associated with the link between the pair of a \( i^{th} \) receive antenna and the \( j^{th} \) transmit antenna. Each \( h_{ij} \) is taken from a zero-mean circularly symmetric complex Gaussian distribution with unit variance, i.e., a variance equal to \( 1/2 \) in both real and imaginary components. The noise vector is \( n = [n_1, \ldots, n_N]^T \in \mathbb{C}^{N \times 1} \), with elements taken from an independent circularly symmetric complex Gaussian with zero average and variance \( \sigma_n^2 \), so that both its real and imaginary components have variance \( \sigma_n^2 / 2 \).

Channel state information at the receiver (CSIR) is assumed and all the channels between antennas are considered reciprocal, i.e., they are the same in the uplink and downlink phases, which in practice simplifies the channel estimation phase.

\( M \)-ary squared quadrature amplitude modulation (\( M \)-QAM) constellations are assumed to be transmitted both from the terminals and from each antenna in the relay. The symbols are taken from a finite complex constellation \( \mathbb{C} \) constructed from the Cartesian product \( C = C_R \times C_R \), where \( C_R \) is the real set

\[
C_R = \left\{ - \left[ \sqrt{M} - 1 \right], \ldots, -3, -1, +1, +3, \ldots, + \left[ \sqrt{M} - 1 \right] \right\}
\]

Without loss of generality, that the filters at the receiver have impulse response \( h(t) \) normalized to \( \int |h(t)|^2 \, dt = 1 \).

The symbol error rate (SER) of the downlink phase is obtained comparing the messages decoded by each terminal with the original messages sent by each of them in the uplink phase. The proposed strategies will be assessed in terms of symbol error rate (SER) as function of the signal-to-noise ratio (SNR) defined as

\[
E\{y^2\} = E\{Hx^2\} = N \sigma_n^2.
\]

Note that the downlink performance accumulates the errors occurred during the two phases, i.e.,

\[
\text{SER}_{\text{total}} = 1 - (1 - \text{SER})^2.
\]

III. PROPOSED STRATEGIES

In the following two strategies are proposed for the exchange of information among all the terminals via a relay, combining a MAC stage called phase 1 using distributed MIMO, common to both proposals, which is then followed by phase 2, involving PLNC to partly cancel interference and then detect the remaining streams as in conventional MIMO spatial multiplexing. The paper proposes and assesses the performance of the multi-way relay channel with three terminals, and at the end generalizes the architecture for networks with more terminals.

In phase 2 (downlink, multi-cast) one has to differentiate two cases: i) when the terminals are equipped with one antenna, which will require three time-slots in total, or ii) when there are two antennas in the terminals (even though they only used one during the MAC stage), which requires
two time-slots in total to exchange all the messages between users.

A. Phase 1: multiple-access channel

In the MAC stage in the uplink (see Fig. 2) the relay should apply a robust detection technique, and for that reason lattice reduction-aided (LRA) – see Algorithm 1 – followed by ordered successive interference cancelation with minimum mean square error (OSIC-MMSE) was used given that its performance is near optimal and captures the diversity order available in MIMO spatial multiplexing [9]. The complex LLL algorithm was used to perform the lattice reduction given its lower complexity when compared to LLL performed on the real-equivalent MIMO model [9], [10].

Algorithm 1: Lattice reduction-aided receiver.

1: Shift and scale the QAM symbols constellation to have a lattice problem (i.e., including the zero vector)

\[
y_{\text{rel}} = \frac{1}{2}(y + H \cdot 1) , \quad \text{where } 1 = [1,1,\cdots,1]^T ;
\]

2: Reduce the lattice basis \( H \) using the Complex–LLL (CLLL) algorithm [10]

\[
(H_{\text{rel}}, M) = \text{CLLL}(H), \quad \text{with } M \text{ unimodular};
\]

3: Apply some detector to the closest vector problem [9] defined by \( (y_{\text{rel}}, H_{\text{rel}}) \), where OSIC-MMSE is used in the paper

\[
z = \text{detect}(y_{\text{rel}}, H_{\text{rel}});
\]

Using \( z \) and \( M \) estimate the symbol in the original coordinate system

\[
\hat{x} = 2M Q_z(z) - 1, \quad \text{where } Q_z \text{ is quantization to integers.}
\]

After phase 1, the MIMO relay has detected the messages \( x_1, x_2, \) and \( x_3 \), and these estimates at the relay are denoted \( \hat{x}_1, \hat{x}_2, \) and \( \hat{x}_3 \). This MAC stage consumes one time-slot in the overall messages exchanging process.

B. Phase 2 with single-antenna terminals

The configuration with single-antenna terminals required a phase 2 with two time-slots in the multicast phase, as illustrated in Fig. 3, in addition to the time-slot spent in phase 1.

In the second time-slot (i.e., the first time-slot in the multicast phase), the relay broadcasts from antenna 1 and 2, \( \hat{x}_1 \) and \( \hat{x}_2 \) respectively, which had been previously detected by the relay. The received signals at each terminal after this second time-slot are

\[
\begin{align*}
y_1^{(1)} & = h_{11}\hat{x}_1 + h_{12}\hat{x}_2 + n \\
y_2^{(1)} & = h_{21}\hat{x}_1 + h_{22}\hat{x}_2 + n \\
y_3^{(1)} & = h_{31}\hat{x}_1 + h_{32}\hat{x}_2 + n
\end{align*}
\]

where the coefficients \( h_{ij} \) are the ones in Fig. 3, which are respectively known by each of the terminals (considering CSIR), and the superscripts denote the processing time-steps (notice that these do not correspond to time-slots).

Considering that no errors occurred at the relay in phase 1 (hence, letting \( \hat{x} = x \), simplifying the notation), and because there is CSIR in all terminals, each terminal can cancel their respective message contributions in the received signal, and therefore

\[
\begin{align*}
y_1^{(2)} & = y_1^{(1)} - h_{11}x_1 \\
y_2^{(2)} & = y_2^{(1)} - h_{22}x_2
\end{align*}
\]

(6)

From (6) and by using the respective matched filters to the remaining signals at terminals 1 and 2, followed by the quantization \( Q_c \) to the constellation symbols, terminal 1 is able to detect \( \hat{x}_1 \) and terminal 2 detects \( \hat{x}_2 \):

\[
\begin{align*}
\hat{x}_1(1) & = Q_c \left[ \frac{h_{11}^{*}}{\|h_{11}\|} y_1^{(2)} \right] \\
\hat{x}_2(2) & = Q_c \left[ \frac{h_{22}^{*}}{\|h_{22}\|} y_2^{(2)} \right]
\end{align*}
\]

(7)

In (7), and in the following, the detected message \( x_i \) at terminal \( j \) is denoted as \( \hat{x}_i(j) \). At this stage, from (7),
terminal 1 has \( \hat{x}_1 \) and terminal 2 has \( \hat{x}_2 \) while terminal 3 still keeps in memory the information \( y_3^{(1)} \) extracted from (5).

In the third time-slot, the relay transmits the messages \( x_1 \), and \( x_3 \) from antennas 1 and 3, respectively (see Fig. 2), and therefore the signals received by the terminals are

\[
\begin{align*}
    y_1^{(3)} &= h_{11} x_1 + h_{13} x_3 + n, \\
    y_2^{(3)} &= h_{21} x_1 + h_{23} x_3 + n, \\
    y_3^{(3)} &= h_{31} x_1 + h_{33} x_3 + n.
\end{align*}
\]

(8)

At this point both terminals 1 and 3 can cancel out their respective uplink original messages using the PLNC principle, whereas terminal 2 can use the previously detected \( \hat{x}_1 \). Hence, the next step amounts to canceling these contributions respectively at each terminal:

\[
\begin{align*}
    y_1^{(4)} &= y_1^{(3)} - h_{11} x_1, \\
    y_2^{(4)} &= y_2^{(3)} - h_{21} \hat{x}_1(2), \\
    y_3^{(4)} &= y_3^{(3)} - h_{31} x_3
\end{align*}
\]

(9)

and again, using the respective matched filters to the remaining signals at each of the terminals, they can all detect the remaining missing message at each one:

\[
\begin{align*}
    \hat{x}_1(1) &= Q_c \begin{bmatrix} h_{13}^* \cr h_{13} \end{bmatrix} y_1^{(4)}, \\
    \hat{x}_2(2) &= Q_c \begin{bmatrix} h_{23}^* \cr h_{23} \end{bmatrix} y_2^{(4)}, \\
    \hat{x}_3(3) &= Q_c \begin{bmatrix} h_{33}^* \cr h_{33} \end{bmatrix} y_3^{(4)}.
\end{align*}
\]

(10)

Terminal 1 and 2 have now all the three messages while terminal 3 has \( x_1(3) \) and \( x_3(3) \). To end the process, terminal 3 can now apply the PLNC concept re-using (5) and cancel \( x_1(3) \) in it, thus computing

\[
y_3^{(5)} = y_3^{(1)} - h_{31} \hat{x}_3(3)
\]

(11)

and finally apply a simple matched filter to estimate \( \hat{x}_2 \):

\[
\hat{x}_2(3) = Q_c \begin{bmatrix} h_{22}^* \cr h_{22} \end{bmatrix} y_3^{(5)}.
\]

(12)

C. Phase 2 with multiple-antenna terminals

One possible way to reduce the three slots to just two is to apply joint decoding to the two unknown messages right after phase 1, which amounts to perform a detection in a lower dimensional space (using the projections of the signals in just one dimension). That obviously introduces a large power penalty in such message exchange mechanism (which for space reasons is not presented in this paper).

This paper proposes to accomplish phase 2 with one time-slot only (in the downlink) by considering terminals with more than one antenna. In the example of a network with three terminals one requires two antennas in each terminal in order to capture the two dimensions.

Considering the same distributed MIMO uplink, the downlink phase now makes use of the two antennas. Since the detection process will be the same for the three terminals, in the following only the case concerning terminal one will be shown (as seen in Fig. 4).

![Figure 4. Downlink phase with two time-slot strategy with MIMO terminals](example of the reception in terminal k=1).

The MIMO relay decodes the uplink messages \( \hat{x}_1, \hat{x}_2 \), and \( \hat{x}_3 \), and afterwards all these messages are simultaneously transmitted by the relay in the second time-slot (with the \( i \)th antenna transmitting \( x_i \)). The sum of the received signal at the two antennas of each terminal is given by

\[
y_{k,i}^{(l)} = h_{k1} \hat{x}_1 + h_{k2} x_i + h_{k2} \hat{x}_2 + h_{k3} \hat{x}_3 + n_{k,l}
\]

(13)

where \( y_{k,l}^{(l)} \) and \( n_{k,l} \) denote respectively the signal and the noise at the \( l \)th antenna of terminal \( k \), and the superscript indicates once again a processing step. Considering, as in section B, no errors in phase 1, and applying the PLNC principles, as each terminal has CSIR and knows its own message sent in the uplink, PLNC can be applied in each terminal, each of which cancels out its own contribution in the received signal, obtaining

\[
y_{k,2}^{(2)} = y_{k,2}^{(1)} - h_{k2} \hat{x}_1
\]

(14)

The remaining problem can now be seen as a 2×2 MIMO spatial multiplexing problem with

\[
\begin{bmatrix} y_{k,1}^{(2)} \\ y_{k,2}^{(2)} \end{bmatrix} = \begin{bmatrix} h_{k1} & h_{k1} \\ h_{k2} & h_{k2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} n_{k,1} \\ n_{k,2} \end{bmatrix}
\]

(15)

where \( k \) and \( \bar{k} \) are the indexes of the two remaining terminals, i.e., \( \bar{k} \neq \bar{k} \neq k \). The problem (15) can be solved by any MIMO detection algorithm [9]. Given its excellent performance versus complexity tradeoff, this paper will present the results using an LRA-OSIC-MMSE detector, similar to the one used at the relay.
IV. PERFORMANCE RESULTS

The strategies with two and three time-slot were both simulated with LRA with OSIC-MMSE detection at the relay (also used in the terminals in the case of the second strategy).

Fig. 5 and Fig. 6 show respectively the performance results in terms of the overall symbol error rate with 4-QAM and 16-QAM, using the same constellations in the uplink and downlink phases. For comparison purposes, the performance with conventional TDMA message exchange is given in Fig. 7 (in this case presenting redundant simulations for each terminal/user). Note that in the TDMA case, the gap between the performance at the relay and at the terminals is simply the result of the accumulation of errors in the two-hop connection.

It is possible to observe that the relay with LRA OSIC-MMSE (in the uplink) is always able attain the channel diversity (i.e., the slope $\log(\text{SER}) / \log(\text{SNR}) = -3$). However, the single-antenna terminals can only attain a slope corresponding to a diversity 1 in the downlink, while in the
strategy with MIMO terminals they are able to achieve diversity 2, attained by the two antennas when using LRA detectors.

The strategy with MIMO terminals triples the throughput of the network in comparison to the traditional TDMA scheme, reducing the communication from 6 time-slots to 2 time-slots only. The strategy with single-antenna terminals only doubles the throughput of the network and offers a poorer performance; however, this is a consequence of the much less complexity burden in the single-antenna terminals, which only need to apply matched filters after PLNC cancelation rather than a lattice detector.

The performance of assessed network with three terminals showed that having MIMO reception in the terminals attains a performance that is 12.5 dB (for 4-QAM) or 7.5 dB (for 16-QAM) better than using a strategy with one antenna and three time-slots. There are also performance gains when comparing to the conventional TDMA scheme, which come from the use of MIMO at the relay and also at the terminals in the second proposal (for both proposals the comparisons should be made to the performance curves in Fig. 7).

This later scheme can be generalized to a network with a larger number of MIMO terminals, T. Let us consider T terminals, each with n = T − 1 antennas, and a relay with N ≥ T antennas. Note that in the case of the example described so far with three terminals, one has T = N = 3 and n = 2. The schemes are feasible as long as the relay has a lattice detector able to cope with the detection complexity associated of N complex dimensions. For any number of terminals T the scheme with two time-slots remains valid: the uplink becomes a distributed MIMO channel with T terminals and in phase 1 the relay detects all the T incoming messages using an adequate MIMO detector such as LRA OSIC-MMSE.

After the relay broadcasting $[x_1, x_2, \ldots, x_N]^T$, each terminal applies PLNC to the respective message and leaves the remaining interference suppression to be dealt with in the spatial domain by some MIMO detection algorithm. For such networks with N terminals, the number of time-slots required to exchange all messages among the terminals with conventional TDMA is $2N$, while in the proposed scheme that number is always fixed to 2 time-slots only.

V. CONCLUSIONS

The paper presented two strategies for the multi-way channel based on a concatenation of physical-layer network coding and MIMO spatial multiplexing, which lead to considerable reductions in the number of time-slots necessary to exchange information between all the terminals through a relay. One proposal considers single-antenna terminals while the second considers multi-antenna terminals: the first doubles the throughput in respect to conventional (single-antenna) TDMA, while the latter triples the throughput, while maintaining the overall end-to-end transmission error (from terminal to terminal).

Both proposals apply the physical-layer network coding concept as the initial stage, and later the detection is complemented by performing interference-cancelation in the spatial domain. The energy performance gains in respect to (single-antenna) TDMA are obviously due to MIMO and lattice reduction-aided receivers were considered both at the terminals and at the relay, allowing the MIMO detection to operate close to the one achieved by maximum likelihood detectors. The strategy proposed with MIMO terminals is preferable if the terminals can afford the extra antennas and the extra processing complexity, and it can be generalized to configurations with T terminals, each of which having $n = N − 1$ antennas and with the MIMO relay equipped with $N ≥ T$ antennas. In that case the number of time-slots required to exchange all messages using TDMA is 2T, while with the proposed strategy the message exchange can be always accomplished in just two time-slots.

ACKNOWLEDGMENTS

Filipe E. Ferreira is a scholar at Instituto de Telecomunicações funded by FCT (Foundation for Science and Technology) under project PEst-OE/EEI/LA0008/2013. Francisco A. Monteiro is grateful to the European COST Action IC1104 – “Random Network Coding and Designs over GF(q)” for providing funding and learning opportunities in the field of network coding.

REFERENCES