Cooperative Partial Detection Using MIMO Relays

Kiarash (Kia) Amiri, Michael Wu, Joseph R. Cavallaro, *Senior Member, IEEE*, and Jorma Lilleberg

***Abstract*—Using multiple-input multiple-output (MIMO) relays in cooperative communication improves the data rate and reliability of the communication. The MIMO transmission, however, requires considerable resources for the detection in the relay. In particular, if a full detect-and-forward (FDF) strategy is employed, the relay needs to spend considerable resources to perform the full MIMO detection. We propose a novel cooperative partial detection (CPD) strategy to partition the detection task between the relay and the destination. CPD modifies the tree traversal of the tree-based sphere detectors in a way where there is no need to visit all the levels of the tree and only a subset of the levels; thus, a subset of the transmitted streams are visited. The destination, then, combines the source signal and the partial relay signal to perform the final detection step and recover the transmitted vector. We study and compare the performance and complexity of FDF and CPD and show that by using the CPD approach, the relay can avoid the considerable overhead of MIMO detection while helping the source-destination link to improve its performance. More specifically, in the case of a 4×4 system, the relay complexity can be reduced by up to 80% of the conventional relaying scheme.**

***Index Terms*—Cooperative system, MIMO, radio communication baseband, receivers, wireless communication.**

I. Introduction

COOPERATIVE communications and in particular relay channels, were originally introduced and studied in [1] where lower and upper bounds on the capacity of relay channels were derived, which were later improved by [2]. User cooperation reemerged again in [3] as a form of diversity in uplink scenarios. Different relaying protocols were studied and compared in [4].

In order to facilitate user cooperation in practical scenarios, coded cooperation was proposed and studied in [5]. Furthermore, in order to reduce the overhead of decoding in the relay, various distributed decoding schemes have been proposed in [6], where the relay performs a partial decoding as opposed to the conventional full decoding of the message.

With the promising results of multiple-input multiple-output (MIMO) point-to-point communications [7], [8], MIMO systems have been playing a significant role in a wide variety of wireless standards and thus, various detection algorithms, mostly based upon sphere detection, have been proposed to reduce the complexity of detection in MIMO systems [9]–[13]. The sphere detection algorithm is based on performing a tree search algorithm to detect the MIMO streams. In this tree search, tree levels correspond to the number of antennas and the number of children of every tree node corresponds to the modulation order. More recently, there have been some attempts to study the theoretical benefits and bounds on deploying MIMO nodes in cooperative scenarios, both as relays and as source/destination pairs. In doing so, lower bounds and upper bounds for MIMO relay networks were given in [14] and [15] and capacity scaling factors were derived for multi-hop MIMO relays [16]. Optimal precoder designs for MIMO relays were discussed in [17]. In [15], full-duplex MIMO relay channels are studied and using message splitting and partial cooperation, rate bounds are derived.

MIMO relays can be mobile multi-antenna users that could choose to assist the active links in the environments during their idle time. Such idle MIMO users act as relays if such cooperation will not require significant processing battery power that they would need later for their own use. Full detect-and-forward in the relay can require a significant amount of resources in MIMO cooperative communications, particularly if the relay chooses to perform a close-to-optimum detection. This effect becomes more important when one considers the practical resource constraints of idle MIMO users operating as relays. Therefore, it is crucial to distribute the detection task between the relay and the destination in such a way that the relay does not need to spend too much of its processing and transmit power and yet can enhance the performance compared to a nonrelay scenario.

In order to address this challenge, we propose a novel cooperative partial detection (CPD) scheme in MIMO relay channels. In CPD, the relay, instead of applying the conventional full detection, performs a partial detection and forwards the detected parts of the message to the destination. Moreover, instead of making the impractical assumption of complete channel state information in all the nodes; the proposed cooperative partial detection strategy assumes that in each communication link, channel knowledge is available only at the receiver of that link. Detecting and transmitting a subset of the source streams from the relay reduces the total transmit power from the relay compared to the conventional relaying, where all the source signals

Manuscript received September 29, 2010; revised January 19, 2011 and April 21, 2011; accepted April 25, 2011. Date of publication May 23, 2011; date of current version September 14, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Tong Zhang. The authors would like to thank Nokia, Nokia Siemens Networks (NSN), Xilinx, and the U.S. National Science Foundation (CCF-0541363, CNS-0551692, EECS-0925942, CNS-0923479, and CNS-0619767) for their support of the research.

K. Amiri, M. Wu, and J. R. Cavallaro are with the Electrical and Computer Engineering Department, Rice University, Houston, TX 77005-1827 USA (e-mail: kiarash.amiri@gmail.com; mbw2@rice.edu; cavallar@rice.edu).

J. Lilleberg is with the Renesas Mobile Corporation, Oulu 990590, Finland (e-mail: jorma.lilleberg@renesasmobile.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2011.2157498

are retransmitted from the relay. Therefore, when there are many transmitters and receivers in the environment, CPD reduces interference due to the relay in the second wireless network.

We define *expansion factor* , ef, as the parameter that captures the number of streams of data detected in the relay and transmitted from the relay to the destination. In other words, ef represents a partial tree traversal, where only a subset of the tree levels are visited and the search is stopped before reaching the end of the tree. Using the ef parameter, we will show that this cooperative detection scheme improves the error performance compared to nonrelay scenarios with limited computational overhead in the relay. We will show that this technique can help in distributing the detection process between the relay and destination. Furthermore, the ef parameter provides the means for the relay so that it could choose, depending on its resource availability, how much of its processing power it should dedicate to helping the direct source–destination link.

Therefore, we propose a *partial* sphere detection scheme, which is designed and proposed based on the practical limitations of wireless devices. This detection scheme is used in the relay for partial detection. We also propose a detection scheme in the destination that is based on maximal ratio combining of the received data.

It is important to note that our proposed cooperative detection scheme can be applied to a wide variety of wireless communications systems. For instance, in the context of uplink scenarios, this scheme can be applied in the MIMO terminal transmitting its spatially multiplexed signals to the basestation. Also, this scheme may be used in assisting the basestation in uplink *multiuser* detection scenarios, where multiple users with a smaller number of antennas try to use the same channel for sending the data to the basestation using a combination of spatial multiplexing and spatial multiple-access techniques. As for the downlink, the MIMO relay can be used for communicating data from the basestation to terminals with multiple antennas. Note that multiple-antenna mobile nodes have been discussed and proposed for IEEE 802.16 [18] and IMT-advanced [19] standards and also for 3GPP LTE [20]. In all such scenarios, the relay node can be either a dedicated MIMO relay, or another idle MIMO user.

In this paper, we are building upon some of the preliminary results presented in [21] and [22] to demonstrate new complexity-power tradeoffs and the impact of channel coding on the complexity and performance of CPD. The rest of the paper is organized as follows. Section II covers the system model definition and the full detect and forward scheme is described in Section III. The proposed cooperative partial detection algorithm is presented in Section IV, and the computational complexity of this technique is studied in Section V. Monte Carlo simulation results of this scheme are presented in Section VI, and hardware architecture and over-the-air verifications are discussed in Section VII. Finally, the paper concludes with Section VIII.

II. System Model

Throughout the paper, we assume a three-node network: the source, relay, and destination, denoted by S, R, and D; respectively. We further assume that the source, relay, and destination are equipped with M_{s}, M_{r}, and M_{d} antennas, respectively. Given the practical limitations of deploying full duplex radios, we assume the relay operates in half-duplex mode. The communication between the source and the destination is performed over two *time slots* . In the first time slot, the source broadcasts its message to both the relay and the destination; and in the second time slot, the relay, using an $ef \leq M_{\mathrm{r}}$ subset of its antennas, transmits its message to the destination while the source is silent. The *expansion factor*, ef, corresponds to the number of *utilized* antennas in the relay during the second time slot. The choice of ef and its impact on the performance and complexity will be discussed in detail in the next sections. The transmitted vector from the source is of length M_{s} and the source uses a spatial multiplexing scheme to transmit different streams, i.e., modulation symbols, on different antennas.

We assume coded systems, where the bits are coded and spread across the transmit antennas in the source before modulation. The bits b_j, $j = 1, \ldots, K$ are passed through the channel coder of rate $\frac{K}{N}$ in the source node to generate u_l, $l = 1, \ldots, N$. The u_l bits are mapped to modulation points x_i and spread across the transmit antennas of the source to form the source transmit vector $\mathbf{x}_{\mathrm{s}}$. Therefore, the $M_{\mathrm{s}} \log w$-length bit vector $\mathbf{u} = [u_1, \ldots, u_{M_{\mathrm{s}} \log w}]^{\mathrm{T}}$ is constructed by concatenating the bit vectors of the transmitted symbols. For instance, $[u_1, \ldots, u_{\log w}]^{\mathrm{T}}$ is the mapping of the x_1 symbol to bits and $[u_{1+\log w}, \ldots, u_{2 \log w}]^{\mathrm{T}}$ is the mapping of the x_2 symbol to bits, where $\mathbf{x}_{\mathrm{s}} = [x_1, x_2, \ldots, x_{M_{\mathrm{s}}}]^{\mathrm{T}}$.

The received signals at the relay and destination at the end of the first time slot are given by

$$\mathbf{y}_{\mathrm{r}} = \mathbf{H}_{\mathrm{sr}} \mathbf{x}_{\mathrm{s}} + \mathbf{n}_{\mathrm{r}} \tag{1}$$

$$\mathbf{y}_{\mathrm{d}}^{(1)} = \mathbf{H}_{\mathrm{sd}} \mathbf{x}_{\mathrm{s}} + \mathbf{n}_{\mathrm{d}}^{(1)}. \tag{2}$$

The relay, then, detects all or part of the transmitted vector symbols and forwards them to the destination. Therefore, the received signal at the destination at the end of the second time slot is given by

$$\mathbf{y}_{\mathrm{d}}^{(2)} = \mathbf{H}_{\mathrm{rd}} \mathbf{x}_{\mathrm{r}} + \mathbf{n}_{\mathrm{d}}^{(2)} \tag{3}$$

where superscripts $^{(1)}$ and $^{(2)}$ are used to distinguish the first and second time slots. Since the relay receives only at the end of the first time slot, no superscript is used for the relay. The noise vectors, $\mathbf{n}_{\mathrm{r}}$, $\mathbf{n}_{\mathrm{d}}^{(1)}$, and $\mathbf{n}_{\mathrm{d}}^{(2)}$ are of size M_{r}, M_{d} and M_{d}, with each of their elements chosen from a complex symmetric Gaussian variable $\mathcal{C}N(0, 1)$. We also assume that each element of the $\mathbf{x}_{\mathrm{s}}$, $\mathbf{x}_{\mathrm{r}}$, and $\mathbf{x}_{\mathrm{d}}$ vectors are chosen from a QAM modulation, $\mathcal{O}$, with the modulation set size of $w = |\mathcal{O}|$ and average power constraint of $E[{x_i}^2] = 1$.

Note that the type of processing in the relay depends on the amount of available resources in the relay. The relay can choose the detection process and how much it is willing to detect the transmitted signals and whether or not, it should perform decoding and re-encoding of the transmitted signals. This is one of the contributions of this paper and will be discussed in more detail in the next sections.

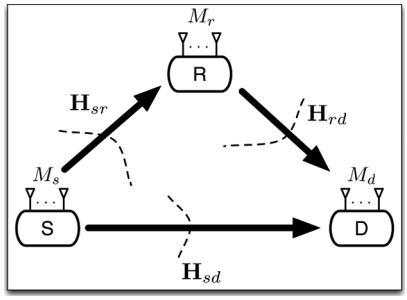


Fig. 1. A relay network with three nodes: source, relay, and destination. The respective channel matrices are denoted by $\mathbf{H}_{\rm sr}$, $\mathbf{H}_{\rm rd}$, and $\mathbf{H}_{\rm sd}$.

As illustrated in Fig. 1, the $\mathbf{H}_{\rm sr}$, $\mathbf{H}_{\rm rd}$, and $\mathbf{H}_{\rm sd}$ are matrices of sizes $M_{\rm r} \times M_{\rm s}$, $M_{\rm d} \times ef$, and $M_{\rm d} \times M_{\rm s}$, and correspond to the channel matrices between the source and the relay, relay and the destination and source and the destination, respectively. All these channel matrices, $\mathbf{H}_{\rm sr}$, $\mathbf{H}_{\rm rd}$ and $\mathbf{H}_{\rm sd}$, have independent elements, each drawn from a circularly symmetric Gaussian random distribution with zero mean and variances of $\sigma_{\rm sr}^2$, $\sigma_{\rm rd}^2$ and $\sigma_{\rm sd}^2$, respectively, where

$$\sigma_{\rm sr}^2 = \sqrt{\frac{\rm SNR_{sr}}{M_{\rm s}}},\ \sigma_{\rm rd}^2 = \sqrt{\frac{\rm SNR_{rd}}{ef}},\ \sigma_{\rm sd}^2 = \sqrt{\frac{\rm SNR_{sd}}{M_{\rm s}}}.$$

We make the practically feasible assumption that the $\mathbf{H}_{\rm sr}$ matrix is known in the relay; and $\mathbf{H}_{\rm sd}$ and $\mathbf{H}_{\rm rd}$ matrices are known in the destination node; thus, only the receivers of each communication link have complete channel knowledge.

The signal-to-noise ratios (SNRs) at each of the received antennas of the relay and destination are defined as

$$\mathrm{SNR_{sr}} = \frac{\mu P}{(d_{\rm sr})^{\alpha}},\ \mathrm{SNR_{rd}} = \frac{(1-\mu)P}{(d_{\rm rd})^{\alpha}} \mathrm{SNR_{sd}} = \frac{\mu P}{(d_{\rm sd})^{\alpha}} \quad (4)$$

where α is the path loss exponent, which usually varies between 2 and 6. The above SNR equations imply that the sum transmit power from the source and the relay is set to P and is split with a proportionality factor of $0 < \mu \leq 1$ such that the source uses μP and the relay uses $(1-\mu)P$. Therefore, if τ represents the symbol time, then the amount of energy per information bit is given by

$$E_b = \frac{\tau\mu P + \tau(1-\mu)P}{M_{\rm s}(\frac{K}{N})\log_2 w} = \frac{\tau P}{M_{\rm s}(\frac{K}{N})\log_2 w}[\text{Joules/bit}]. \quad (5)$$

III. Conventional Full Detect-and-Forward

In this section, we present the symbol-level detector in the relay and destination. In the full detect-and-forward (FDF), the source transmits $\mathbf{x}_{\rm s}$ in the first time slot and the relay and destination receive their copy of the transmitted vector, $\mathbf{y}_{\rm r}$ and $\mathbf{y}_{\rm d}^{(1)}$. Then, the relay performs full sphere detection, as described in the previous section, on its received vector, $\mathbf{y}_{\rm r}$, to find $\tilde{\mathbf{x}}_{\rm s}$, where $\tilde{\mathbf{x}}_{\rm s}$ is equal to $\mathbf{x}_{\rm s}$ in an error-free detection:

$$\tilde{\mathbf{x}}_{\rm s} = \underset{\mathbf{x}\in\mathcal{O}^{M_{\rm s}}}{\arg\min} \left|\left| \mathbf{y}_{\rm r} - \mathbf{H}_{\rm sr}\mathbf{x} \right|\right|^2. \quad (6)$$

The norm in (6) can be rewritten as [9]

$$\begin{aligned} D(\mathbf{x}) &= \| \mathbf{y}_{\rm r} - \mathbf{H}_{\rm sr}\mathbf{x} \|^2 \\ &= \| \mathbf{Q}^*\mathbf{y}_{\rm r} - \mathbf{R}\mathbf{x} \|^2 = \sum_{i=M_{\rm s}}^{1} \left| y_i' - \sum_{j=i}^{M_{\rm s}} R_{i,j}x_j \right|^2 \end{aligned} \quad (7)$$

where $\mathbf{H}_{\rm sr} = \mathbf{QR}$, $\mathbf{QQ}^* = \mathbf{I}$ and $\mathbf{y}' = \mathbf{Q}^*\mathbf{y}_{\rm r}$. Throughout this paper, we will use the superscript * to denote the matrix Hermitian transpose. This minimization process can be performed in a depth-first tree search sphere detection, where the tree levels correspond to the transmit antennas and the children of a tree node represent different modulation orders. The sphere detection has shown significant gains over the brute-force search scheme [9].

Finally, the relay transmits the $\mathbf{x}_{\rm r} = \tilde{\mathbf{x}}_{\rm s}$ in the second time slot to the destination, using the same modulation order. The received vector at the destination from the relay is denoted by $\mathbf{y}_{\rm d}^{(2)}$. The destination can now combine the received copies from the source and relay and perform a sphere detection on the newly formed combined vector.

We will now derive the combination procedure. Given the two received copies in the destination, the MAP detector, assuming that the stream from the relay is error-free, is

$$\underset{\mathbf{x}\in\mathcal{O}^{M_{\rm s}}}{\arg\max} P\{\mathbf{x}|\mathbf{y}_{\rm d}^{(2)}, \mathbf{y}_{\rm d}^{(1)}, \mathbf{H}_{\rm rd}, \mathbf{H}_{\rm sd}, \mathbf{x}_{\rm s} = \tilde{\mathbf{x}}_{\rm s}\} \quad (8)$$

which, given the equal noise power in different links, is equivalent to

$$\underset{\mathbf{x}\in\mathcal{O}^{M_{\rm s}}}{\arg\min} \left(\| \mathbf{y}_{\rm d}^{(2)} - \mathbf{H}_{\rm rd}\mathbf{x} \|_2^2 + \| \mathbf{y}_{\rm d}^{(1)} - \mathbf{H}_{\rm sd}\mathbf{x} \|_2^2 \right). \quad (9)$$

After expanding each of the norms in (9) and regrouping the different terms, (9) can be rewritten as

$$\underset{\mathbf{x}\in\mathcal{O}^{M_{\rm s}}}{\arg\min} (\| \mathbf{y}_{\rm FDF} - \mathbf{H}_{\rm FDF}\mathbf{x} \|_2^2) \quad (10)$$

where the equivalent channel matrix, $\mathbf{H}_{\rm eq}$, and the equivalent received vector, $\mathbf{y}_{\rm eq}$, are given by

$$\mathbf{H}_{\rm FDF} = (\mathbf{H}_{\rm sd}^*\mathbf{H}_{\rm sd} + \mathbf{H}_{\rm rd}^*\mathbf{H}_{\rm rd})^{\frac{1}{2}} \quad (11)$$

$$\mathbf{y}_{\rm FDF} = \mathbf{H}_{\rm FDF}^{-1}(\mathbf{H}_{\rm sd}^*\mathbf{y}_{\rm d}^{(1)} + \mathbf{H}_{\rm rd}^*\mathbf{y}_{\rm d}^{(2)}). \quad (12)$$

It is worth noting that (11) and (12) are essentially similar to performing a MIMO maximal-ratio combining (MRC), followed by whitening the colored noise [23]. The equivalent received vector and channel matrix can also be computed by concatenating the received signals and channel matrices

$$\mathbf{H}_{\rm FDF} = \begin{bmatrix} \mathbf{H}_{\rm sd} \\ \mathbf{H}_{\rm rd} \end{bmatrix}, \mathbf{y}_{\rm FDF} = \begin{bmatrix} \mathbf{y}_{\rm d}^{(1)} \\ \mathbf{y}_{\rm d}^{(2)} \end{bmatrix}. \quad (13)$$

While the concatenation process of (13) does not require the per-vector combining of [(11)–(12)], it increases the size of the effective channel matrix and thus, requires more resources for QR decomposition. However, since QR decomposition needs

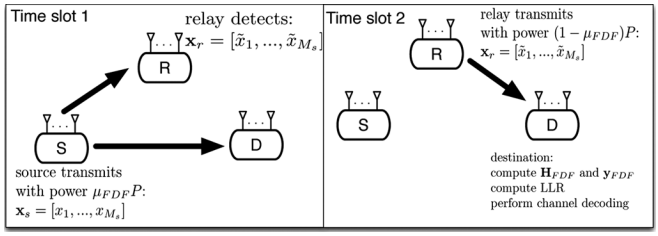


Fig. 2. Full detect-and-forward through MIMO relay node. In the first time slot, the relay receives a copy of the source multi-stream data and detects it and forwards the detected data. In the second time slot, the receiver combines the multiple copies as described earlier to compute the LLR values. We denote the power splitting ratio by μ_{FDF}.

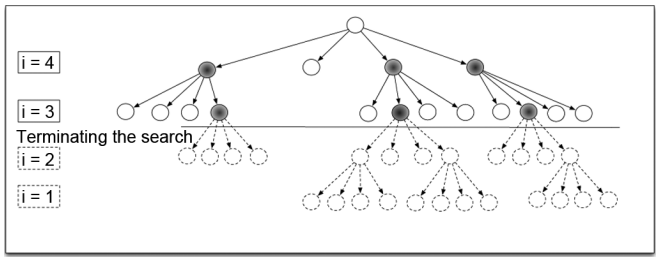


Fig. 3. The tree structure for a partial sphere detector with the expansion factor of two, $ef = 2$. Each node has 16 children for the example case of 16-QAM modulation.

to happen at the channel updating rate, as opposed to symbol vector rate, it generally leads to a less complex procedure.

The soft values can then be computed according to [24]

$$\begin{aligned} L_E(u_k|\mathbf{y}_{\mathrm{FDF}}) &\approx \frac{1}{2}\max_{\mathbf{x}\in\mathcal{L}\cap\mathbf{U}_{k,+1}}\left\{-\frac{1}{\sigma^2}||\mathbf{y}_{\mathrm{FDF}}-\mathbf{H}_{\mathrm{FDF}}\cdot\mathbf{x}||^2\right\} \\ &\quad -\frac{1}{2}\max_{\mathbf{x}\in\mathcal{L}\cap\mathbf{U}_{k,-1}}\left\{-\frac{1}{\sigma^2}||\mathbf{y}_{\mathrm{FDF}}-\mathbf{H}_{\mathrm{FDF}}\cdot\mathbf{x}||^2\right\} \quad (14) \end{aligned}$$

where $\mathcal{L}$ is the list of possible vectors and σ^2 is the noise variance. The $\mathbf{U}_{k,+1}$ is the set of $2^{M_{\mathrm{s}}\log w-1}$ bits of vector $\mathbf{u}$ with $u_k=+1$, while $\mathbf{U}_{k,-1}$ is similarly defined with $u_k=-1$.

Note that the performance of such a detector and decoder pair will be further improved if the detector and decoder, iteratively, pass the LLR information between each other [24]. However, since the focus of this paper is on the cooperative aspect of the detection process and in order not to complicate the parameters, we choose a no-iteration case. Fig. 2 summarizes the steps of the full detect-and-forward.

IV. Reducing Complexity Using Cooperative Partial Detection

In this section, we propose CPD as a low-complexity strategy for relays with limited resources. The c CPD is based on partial sphere detection in the relay to facilitate the cooperative detection strategy.

A. Partial Sphere Detection in the Relay

In order to reduce the relay overhead, we propose partial sphere detection (P-SD), where the relay visits only a subset of the tree levels as opposed to all the levels. Our proposed P-SD requires similar preprocessing operations as that of the conventional sphere detector: the QR decomposition triangularizes the channel matrix and the tree traversal starts from the top level, $i=M_{\mathrm{s}}$, where M_{s} is the number of transmit antennas and ends at $i=1$, i.e., the bottom layer of the tree. Unlike the conventional sphere detection method, the tree traversal of the partial sphere detection method terminates in one of the middle levels and the corresponding minimum distance at that level is considered as the *partial* detected symbol vector. We call the number of visited antennas the *expansion factor* , ef and, as pointed out in Section II, use ef antennas of the relay to transmit those messages. Fig. 3 shows this process for an example case with 16-QAM modulation and expansion factor of 2.

In other words, instead of transmitting $\mathbf{x}_{\mathrm{r}}=\tilde{\mathbf{x}}_{\mathrm{s}}$, as in FDF, the relay now transmits only ef symbols, $\mathbf{x}_{\mathrm{r}}=[\tilde{x}_1,\ldots,\tilde{x}_{ef}]^{\mathrm{T}}$, where the superscript $^{\mathrm{T}}$ denotes the vector transpose operation.

In order to understand the computational savings of the P-SD, we should note that the complexity of sphere detection, in terms of computation count, can be modeled as

$$C_{\mathrm{SD}}(M_{\mathrm{s}},w)=\sum_{i=M_{\mathrm{s}}}^{1}C_iE\{D_i\} \quad (15)$$

where C_i corresponds to the computation count for one node in level i, and $E\{D_i\}$ is the average number of visited nodes in level i. Based on (7), it is clear that C_i is larger for the nodes closer to the bottom of the tree, i.e., $C_{i+1}<C_i$. Therefore, P-SD reduces the total complexity in the relay by not only reducing the total number of visited nodes, but also by limiting the search to the nodes located at the top of the tree with less computation per node. We should note that in other detection mechanisms, e.g., [25], C_{i+1} could be equal or greater than C_i. In such cases, the total complexity is still reduced because of the smaller number of visited nodes.

B. CPD in the Destination

In the symbol combining method, the destination combines the two received vectors, $\mathbf{y}_{\mathrm{d}}^{(1)}$ and $\mathbf{y}_{\mathrm{d}}^{(2)}$, as shown below.

We first break the original transmitted vector into two parts:

$$\mathbf{x}_{\mathrm{s}}=\mathbf{x}=\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix} \quad (16)$$

where

$$\mathbf{x}_1=\begin{bmatrix}x_1\\ \cdot\\ \cdot\\ x_{ef}\end{bmatrix},\mathbf{x}_2=\begin{bmatrix}x_{ef+1}\\ \cdot\\ \cdot\\ x_{M_{\mathrm{s}}}\end{bmatrix} \quad (17)$$

and denote the relay's transmitted vector as

$$\mathbf{x}_{\mathrm{r}}=\tilde{\mathbf{x}}_1=\begin{bmatrix}\tilde{x}_1\\ \cdot\\ \cdot\\ \tilde{x_{ef}}\end{bmatrix}. \quad (18)$$

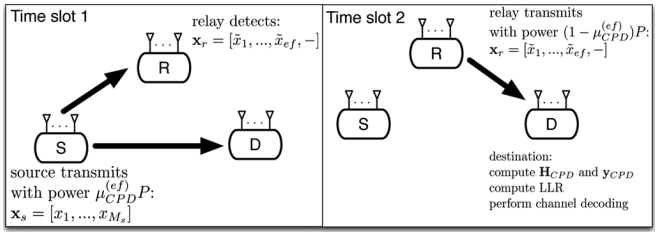


Fig. 4. Cooperative partial detection through MIMO relay node. In the first time slot, the relay receives a copy of the source multi-stream data and partially detects it and forwards the detected data. In the second time slot, the receiver combines the multiple copies as described earlier to compute the LLR values. We denote the power splitting ratio by $\mu_{\mathrm{CPD}}^{(ef)}$.

We also split the source-destination channel matrix into two parts according to (16):

$$\mathbf{H}_{\mathrm{sd}} = [\,\mathbf{H}_1 \mathbf{H}_2\,]. \tag{19}$$

Similar to (9), assuming perfect detection in the relay, i.e., $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$, the symbol level maximum-likelihood solution can be written as the following minimization problem:

$$\underset{\mathbf{x}\in\mathcal{O}^{M_{\mathrm{s}}}}{\arg\min} \left(\| \mathbf{y}_{\mathrm{d}}^{(2)} - \mathbf{H}_{\mathrm{rd}}\mathbf{x}_1 \|_2^2 + \| \mathbf{y}_{\mathrm{d}}^{(1)} - \mathbf{H}_1\mathbf{x}_1 - \mathbf{H}_2\mathbf{x}_2 \|_2^2 \right). \tag{20}$$

After rewriting and regrouping the terms in (20), as shown in the Appendix, we can summarize it as

$$\underset{\mathbf{x}\in\mathcal{O}^{M_{\mathrm{s}}}}{\arg\min} (\| \mathbf{y}_{\mathrm{CPD}} - \mathbf{H}_{\mathrm{CPD}}\mathbf{x} \|_2^2) \tag{21}$$

where the equivalent channel matrix, $\mathbf{H}_{\mathrm{CPD}}$ and the equivalent received vector, $\mathbf{y}_{\mathrm{CPD}}$, are given by

$$\mathbf{H}_{\mathrm{CPD}} = \begin{bmatrix} \mathbf{H}_1^*\mathbf{H}_1 + \mathbf{H}_{\mathrm{rd}}^*\mathbf{H}_{\mathrm{rd}} & \mathbf{H}_1^*\mathbf{H}_2 \\ \mathbf{H}_2^*\mathbf{H}_1 & \mathbf{H}_2^*\mathbf{H}_2 \end{bmatrix}^{\frac{1}{2}} \tag{22}$$

$$\mathbf{y}_{\mathrm{CPD}} = \mathbf{H}_{\mathrm{CPD}}^{-1} \begin{bmatrix} \mathbf{H}_1^*\mathbf{y}_{\mathrm{d}}^{(1)} + \mathbf{H}_{\mathrm{rd}}^*\mathbf{y}_{\mathrm{d}}^{(2)} \\ \mathbf{H}_2^*\mathbf{y}_{\mathrm{d}}^{(1)} \end{bmatrix}. \tag{23}$$

Similar to (13), the equivalent channel matrix and received vector can also be computed by concatenating the received signals and channel matrices:

$$\mathbf{H}_{\mathrm{CPD}} = \begin{bmatrix} \mathbf{H}_1\mathbf{H}_2 \\ \mathbf{H}_{\mathrm{rd}}0 \end{bmatrix}, \mathbf{y}_{\mathrm{CPD}} = \begin{bmatrix} \mathbf{y}_{\mathrm{d}}^{(1)} \\ \mathbf{y}_{\mathrm{d}}^{(2)} \end{bmatrix}. \tag{24}$$

After combining the effective $\mathbf{y}_{\mathrm{CPD}}$ and $\mathbf{H}_{\mathrm{CPD}}$, they are passed to a sphere detector to compute the LLR values and then passed to the channel decoder. Fig. 4 summarizes these steps.

V. Computational Complexity Comparison

In this section, we derive and compare the complexity of the proposed techniques. The channel usually changes at a smaller rate than the received signal and can be implemented with higher resource reuse in the hardware. Therefore, in computing the complexity, we mainly focus on the operations that happen in the symbol updating rate, as opposed to channel updating rate.

The complexity of a sphere detection operation, C_{SD} is given in (15). In order to compute C_i, we refer to the VLSI implementation of [11] and note that, for each node, one needs to compute the $R_{i,j}b_j$, multiplications, where, except for the diagonal element, $R_{i,i}$, the rest of the multiplications are complex valued. The expansion procedure, (7), requires computing $R_{i,j}b_j$ for $j = i+1, \ldots, M_{\mathrm{s}}$, which would require $(M_{\mathrm{s}} - i)$ complex multiplications and also computing $R_{i,i}b_i$ for all the possible choices of $b_j \in \mathcal{O}$. Even though there are w different b_j's, there are only $(\frac{\sqrt{w}}{2} - 1)$ different multiplications required for QAM modulations. For instance, for a 16-QAM system with $\{\pm 3 \pm 3j, \pm 1 \pm 1j, \pm 3 \pm 1j, \pm 1 \pm 3j\}$, computing only $(R_{i,i} \times 3)$ would be sufficient for all the choices of modulation points.[1] Note that computing $(R_{i,i} \times 3)$ can be done either using dedicated multipliers, or shift–add operations. We will assume an FPGA-based implementation with DSP48s and dedicated multipliers, such as in a Xilinx Virtex device, and we will assume multipliers are being used for implementing these multiplication operations. Finally, computing the $\| \cdot \|^2$ norm requires a squarer or a multiplier, depending on the architecture and hardware availabilities.

In order to compute the number of adders for each norm expansion in (7), we assume a depth-first based tree search. Therefore, there are $(M_{\mathrm{s}} - i)$ complex valued adders required for $f_i' - \sum_{j=i+1}^{M_{\mathrm{s}}} R_{i,j}b_j$ and w more complex adders to add the newly computed $R_{i,i}b_i$ values. Once the w different norms, $\left|f_i' - \sum_{j=i}^{M_{\mathrm{s}}} R_{i,j}b_j\right|^2$, are computed, they need to be added to the partial distance coming from the higher level, which requires w more addition procedures. Finally, unless the search is happening at the end of the tree, the norms need to be sorted, which assuming a simple sorter, requires $\frac{w(w+1)}{2}$ compare–select operations.

Therefore, keeping in mind that each complex multiplier corresponds to four real-valued multipliers and two real-valued adders and that every complex adder corresponds to two real-valued adders, C_i is calculated by

$$\begin{aligned} C_i(M_{\mathrm{s}}, w) = {} & \gamma\left(\left(\frac{\sqrt{w}}{2} - 1\right) + 4(M_{\mathrm{s}} - i) + 1\right) \\ & + \theta(2(M_{\mathrm{s}} - i) + 2w + w) \\ & + \beta\left(\frac{w(w+1)}{2}\right) \cdot f(i-1) \end{aligned} \tag{25}$$

where $f(i-1)$ is used to ensure sorting is counted only when the search has not reached the end of the tree and is equal to

$$f(t) = \begin{cases} 1 & t \geq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

Moreover, we use θ, β, and γ to represent the hardware-oriented costs for one adder, one compare-select and one multiplication operation, respectively. Based on FPGA and ASIC estimates, we choose $\theta = 1$, $\beta = 1$ and $\gamma = 10$ throughout this paper.

We note that this is only one method of implementing this architecture and depending on the architecture and timing requirements, other architectures could be used, which may lead to slightly different implementation and computation count. However, these differences will not produce significant impact on

[1]Note that if the wireless standard is using a modulation set other than $\{\pm 3 \pm 3j, \pm 1 \pm 1j, \pm 3 \pm 1j, \pm 1 \pm 3j\}$, an additional multiplication operations will be needed to renormalize the constellation point to the this set.

our comparisons since our goal is to *compare* different cooperative schemes, assuming that all of them use the *same* MIMO detector structures.

Therefore, the computational complexity in the relay for the CPD is given by

$$C_{\text{relay}}(M_{\text{s}}, ef, w) = C_{\text{SD}}(ef, w). \tag{27}$$

In order to compute the complexity in the destination, we extend the definition in (15) to the soft complexity of sphere detectors, C_{SSD}, that compute the LLR values for a list of size $|\mathcal{L}|$. The C_{SSD} is essentially similar to (15), except that $E\{D_i\}$ is now dependent on the target list size:

$$C_{\text{SSD}}(M_{\text{s}}, w, |\mathcal{L}|) = C_{\text{LLR}}(M_{\text{s}}, w, |\mathcal{L}|) + \sum_{i=M_{\text{s}}}^{1} C_i E\{D_i \| \mathcal{L}|\} \tag{28}$$

where $C_{\text{LLR}}(M_{\text{s}}, w, |\mathcal{L}|)$ is the number of operations required to compute the soft values (14).

Note that for similar M_{s}, w and list size $|\mathcal{L}|$, (14) remains the same. Therefore, for the sake of simplicity, we have not considered it in evaluating the C_{SSD}. Moreover, we use the concatenating methods of (13) and (24) in computing the destination complexity.

The total computational complexity in the destination for FDF and CPD are given, respectively, by

$$C_{\text{dest}}^{(\text{FDF})}(M_{\text{s}}, w) = C_{\text{dest}}^{(CPD)}(M_{\text{s}}, ef, w) = C_{\text{SSD}}(M_{\text{s}}, w, |\mathcal{L}|). \tag{29}$$

The simulation results for the complexity is shown in the next section.

VI. Simulation Results

We assume a three node relay network topology with the relay located between the source and destination, on the same line and thus $d_{\text{sd}} = d_{\text{sr}} + d_{\text{rd}} = 1$. We further assume that the path loss exponent is fixed to $\alpha = 3$. We fix the location of the relay and then optimize the performance of the full detect-and-forward network by varying the power splitting ratio μ, as defined in (4), from the discrete set of $\{0.05, 0.1, 0.15, 0.2, 0.25, \ldots, 0.9, 0.95\}$ and call it μ_{FDF}. In order to ensure that the savings in the relay are not limited to baseband processing saving, we also scale the transmit power of the relay by the ratio of the antennas being used. The power splitting ratio for the CPD case, $\mu_{\text{CPD}}^{(ef)}$, is, therefore defined as

$$\mu_{\text{CPD}}^{(ef)} = 1 - (1 - \mu_{\text{FDF}})\frac{ef}{M_{\text{s}}} \tag{30}$$

which implies that the relay transmit power in CPD scenario is scaled down by a factor of $(\frac{ef}{M_{\text{s}}})$ compared to the FDF case and source uses a higher transmit power in return. This choice of transmit power allocation to relay and source better models the real-world per-antenna power constraint and guarantees that by picking the partial detection strategy, the relay not only saves in the baseband computational processing, but also, in the transmit power. For the sake of completeness, we also present the BER performance for a full decode-and-forward scenario, where the relay fully detects and decodes the source signal and then, re-encodes the signal and transmits that to the destination. In the full decode-and-forward, which is presented here as a comparison baseline point, the destination performs a process similar to the full detect-and-forward (FDF) case. Obviously, for the full decode-and-forward scenario, the complexity and delay of the processing in the relay will be much higher than the full detect-and-forward (FDF) case due to the full soft sphere detection and decoding process in the relay.

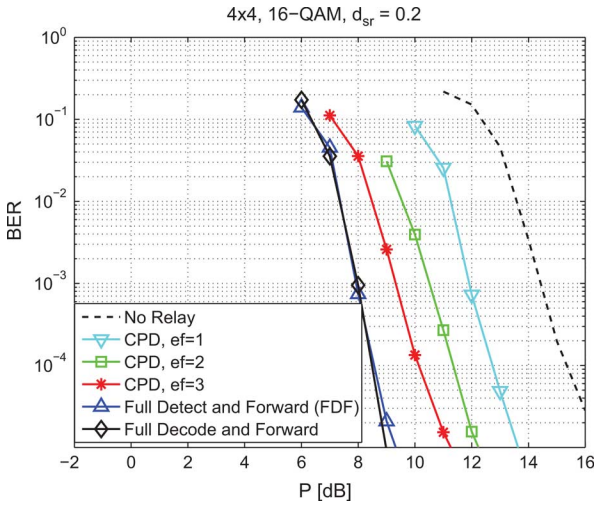


Fig. 5. BER Comparison for a system with $M_{\text{s}} = M_{\text{d}} = 4$ and 16-QAM. The relay is located at $d_{\text{sr}} = 0.2$. The power splitting ratios of the full detect-and-forward and full decode-and-forward is set to $\mu_{\text{FDF}} = 0.6$. The $\mu_{\text{CPD}}^{(ef)}$ for $ef = 3, 2$ and 1 is set to $0.7, 0.8$, and 0.9, respectively.

For this section's simulations, a rate $\frac{1}{2}$ Turbo code is used in the source with an interleaver of size 1355 and feedback polynomial $(1 + D + D^2)$ and feedforward polynomial $(1 + D^2)$. Rayleigh fading channel coefficients, as described in the previous sections of the paper, are used.

Fig. 5 shows the BER performance for a MIMO relay system with 4 antennas and a 16-QAM modulation. The relay is located at $d_{\text{sr}} = 0.2$ and the results are presented for different ef values. As ef increases, the performance gets closer to the full detection scenario. Therefore, the relay can adjust its level of complexity based on the available computational resources. The list size $|\mathcal{L}|$ is set to 100 in the destination for both the full detect-and-forward and full decode-and-forward cases. Note that since the relay is located relatively close to the source, it enjoys a very high SNR source-relay link and therefore, performing the decoding procedure in the relay does not improve the performance significantly compared to just detecting.

Assuming the same system configuration, Fig. 6 shows the required total transmit power required to achieve a BER of 10^{-4} for different power splitting ratios μ. The optimum μ_{FDF} is 0.6 and using 30 the power sharing between the source and the relay is determined.

Monte Carlo simulations are used to generate $E\{D_i\}$ and $E\{D_i \| \mathcal{L}|\}$ and in combination with ((25)–(29)) compute the overall complexity for different total transmit power P values. Figs. 7 and 8 show these results for a four-antenna system with

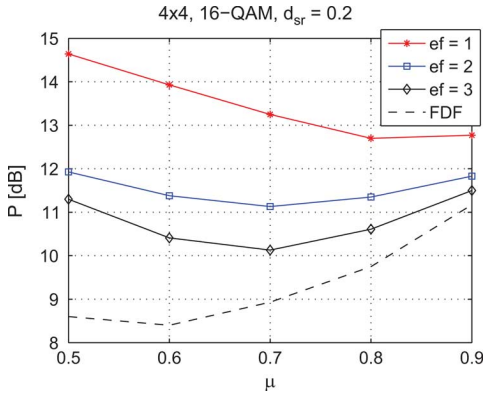


Fig. 6. The total transmit power required to achieve a BER of 10^{-4} for different power splitting ratios. A system with $M_{\rm s} = M_{\rm d} = 4$ and 16-QAM is assumed and the relay is located at $d_{\rm sr} = 0.2$.

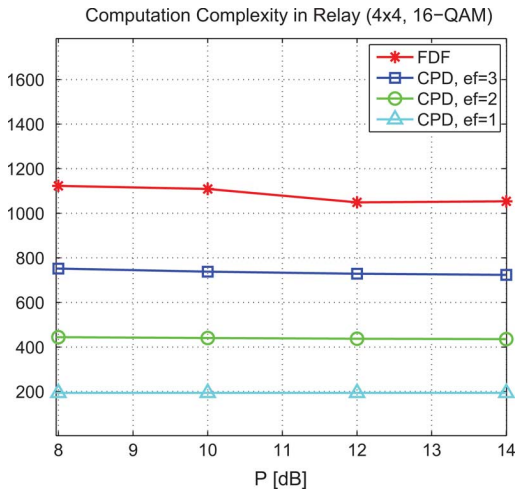


Fig. 7. Comparison between the complexity of detection in relay for FDF and CPD with expansion factors of 2 and 3. The relay is located at $d_{\rm sr} = 0.2$. The power splitting ratios of the full detect-and-forward and full decode-and-forward is set to $\mu_{\rm FDF} = 0.6$. The $\mu_{\rm CPD}^{(ef)}$ for $ef = 3, 2$, and 1 is set to 0.7, 0.8, and 0.9, respectively.

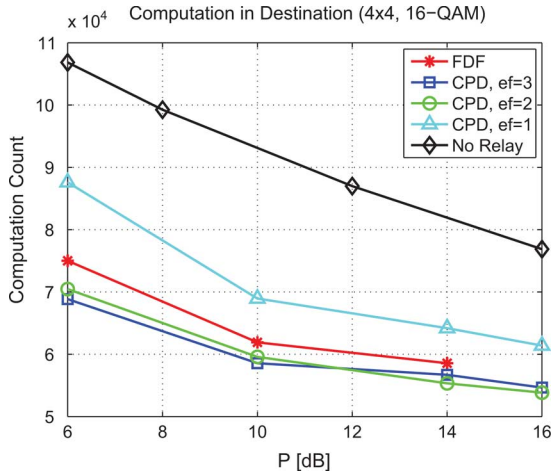


Fig. 8. Comparison between the complexity of detection in the destination for FDF and CPD, with expansion factors of 1, 2, and 3.

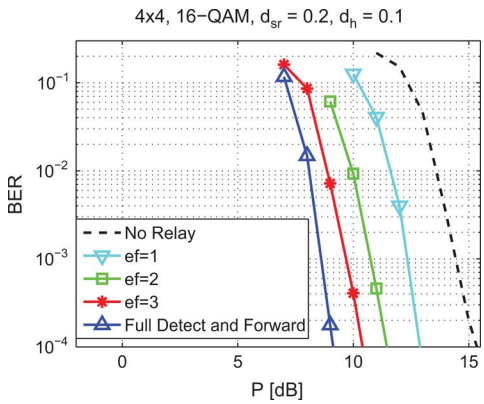


Fig. 9. BER Comparison for a system with $M_{\rm s} = M_{\rm d} = 4$ and 16-QAM. The relay is located at $d'_{\rm sr} = 0.2$ and $d_h = 0.1$. The power splitting ratios of the full detect-and-forward is set to $\mu_{\rm FDF} = 0.65$. The $\mu_{\rm CPD}^{(ef)}$ for $ef = 3, 2$ and 1 is set to 0.7375, 0.825, and 0.9125, respectively.

16-QAM modulation in both the relay and the destination. Note that the list size, in the destination, is set to $|\mathcal{L}| = 100$. The relay requires less computational overhead if it chooses to perform partial sphere detection with 1, 2, or 3 streams of data. Both the FDF and CPD methods require the destination to perform a full sphere detection besides combinations of (13) and (24). Depending on whether the relay complexity, or the total complexity, are the bottleneck, the expansion factor can be chosen. For instance, in this case, if the relay has limited transmit and processing power, the expansion factor can be selected such that the relay complexity, Fig. 7, is maintained within acceptable range while achieving the target BER at that transmit power.

Note that in Fig. 7, as more streams are detected, the effective SNR in the destination and hence the reliability of the detection, goes higher. Therefore, on average fewer nodes need to be visited and that results in reducing the overall complexity. However, for the full detect-and-forward case, the extra preprocessing costs result in a slightly higher computation cost.

We also consider a system where the relay is not located on the direct line connecting source to destination. In this case, we assume that the projection of the relay's location onto the direct line is at $d'_{\rm sr} = 0.2$ distance from the source and the distance between the relay and the direct line is $d_h = 0.1$. Therefore, the relay is 0.2236 away from the source and 0.8 away from the destination. Fig. 9 shows that with this system configuration, the cooperative partial detection shows similar behavior to the case where the relay is located on the same line connecting the source to destination.

Figs. 10 and 11 show similar BER results for 3×3 and 4×4 systems with the relay located at $d_{\rm sr} = 0.4$. The list size $|\mathcal{L}|$ is set to 60 in the destination for both the FDF and full decode-and-forward for the 3×3 case and $|\mathcal{L}| = 100$ for the 4×4 case. Note that the gap between the full decode-and-forward and FDF is wider in $d_{\rm sr} = 0.4$ cases compared $d_{\rm sr} = 0.2$ of Fig. 5. This effect is because of the stronger channel between source and relay in the first case, i.e., $d_{\rm sr} = 0.2$. In other words, since the source–relay channel is relatively stronger, the channel decoding in the relay does not improve the overall error performance, which is now dominated by other factors, such as the source–destination and relay–destination links.

In order to better understand the complexity-power tradeoff, we present the minimum total transmit power P required to

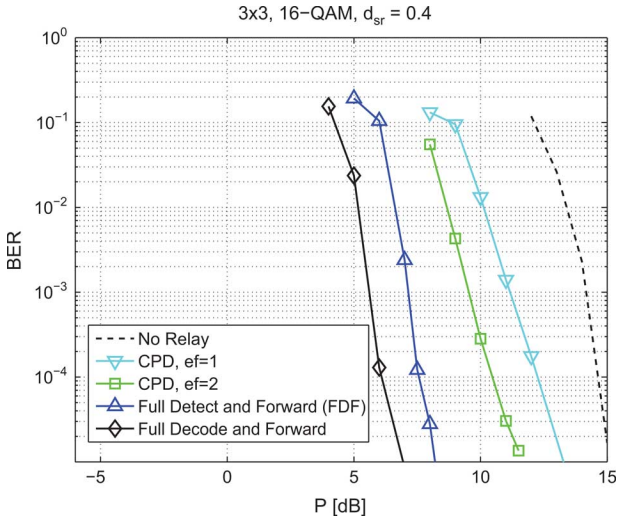


Fig. 10. BER Comparison for a system with $M_{\rm s} = M_{\rm d} = 3$ and 16-QAM. The relay is located at $d_{\rm sr} = 0.4$. The power splitting ratios of the FDF and full decode-and-forward is set to $\mu_{\rm FDF} = 0.6$. The $\mu_{\rm CPD}^{(ef)}$ for $ef = 2$ and 1 is set to 0.73 and 0.86, respectively.

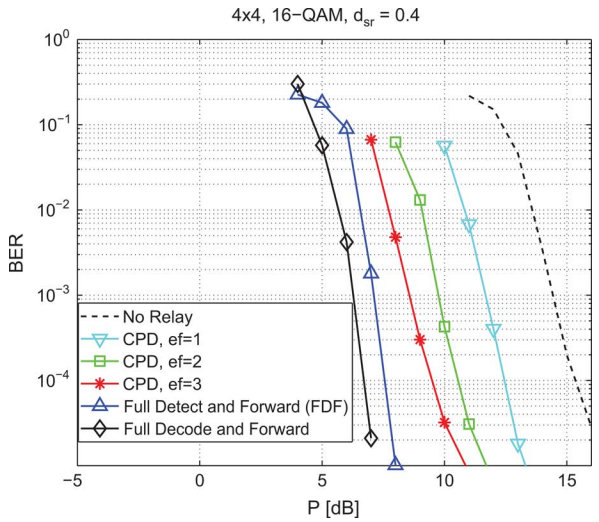


Fig. 11. BER Comparison for a system with $M_{\rm s} = M_{\rm d} = 4$ and 16-QAM. The relay is located at $d_{\rm sr} = 0.4$. The power splitting ratios of the FDF and full decode-and-forward is set to $\mu_{\rm FDF} = 0.65$. The $\mu_{\rm CPD}^{(ef)}$ for $ef = 3$, 2, and 1 is set to 0.7375, 0.825, and 0.9125, respectively.

achieve a target BER. This is shown in Fig. 12, where the required power is plotted versus the expansion factor, ef, which is a complexity measure. Similar to the earlier simulation results, for each relay location, the power splitting ratio, $\mu_{\rm FDF}$, that achieves a better performance for the full detect-and-forward is picked from the limited set of $\{0.1, 0.2, \ldots, 0.9\}$. Then, the corresponding $\mu_{\rm CPD}^{(ef)}$ for the partial detection schemes are chosen according to (30). We observe that detecting more streams in the relay, i.e., higher ef, improved the overall performance; therefore, higher ef translates into lower required power.

VII. Hardware Architecture and Over-the-Air Verification

In this section, we discuss a possible architecture for the implementation of the cooperative partial detection in the relay and

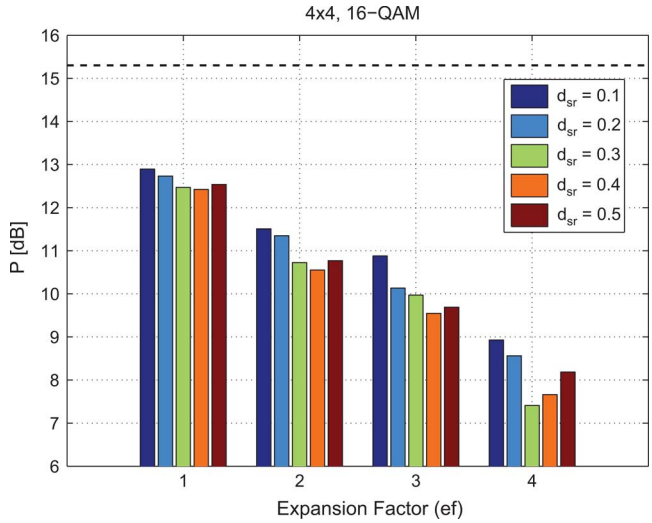


Fig. 12. Complexity-power tradeoff for a 4×4, 16-QAM system with relay located at different positions. The vertical axis corresponds to the required total transmit power to achieve a BER of 10^{-4} and the horizontal axis represents the expansion factor ef. The last set of bars, i.e., $ef = 4$, corresponds to the FDF, and the dashed line corresponds to the no relay scenario.

the destination. Flex-sphere architecture [26] can support different number of antennas and modulation orders. Therefore, it provides the flexibility needed in the relay to perform the detection for different values of ef. Fig. 13 shows the architecture of the detection block in the relay, where rows correspond to the different modulation orders and columns correspond to the number of antennas. Assuming a real-valued decomposition, two levels are needed per transmit antennas. Therefore, the inputs to the final MUXes come from the columns that correspond to the complete complex symbol.

Table I summarizes the synthesis results for a complete $M_{\rm s} = 4$, Virtex-5, implementation. As observed in this table, the flexible detector can support a wide range of scenarios, which makes it suitable for cooperative partial detection, while supporting data rates required by the many wireless standards.

A. Over-the-Air Verification With WARPLab

In this section, we describe the hardware platform to perform cooperative communication tests and demonstrate its applicability in practical scenarios. We will utilize WARPLab, which is a platform for rapid prototyping of physical layer algorithms over the air. It takes advantage of the WARP hardware [21], [27] and Matlab at the same time.

The three WARP boards are connected to a PC through Ethernet. In order to emulate channel behavior, an Azimuth ACE 400 WB wireless channel emulator [28] is used. The emulator can support up to two four-antenna boards. Therefore, for the 2×2 full MIMO relay setup, we use two inputs, four outputs, and 12 paths. For the first time slot, we designate one node as the source, one node as the relay and one node as the destination. In the second time slot, we designate one node as the relay and one node as the destination and connect the two nodes with four reverse links.

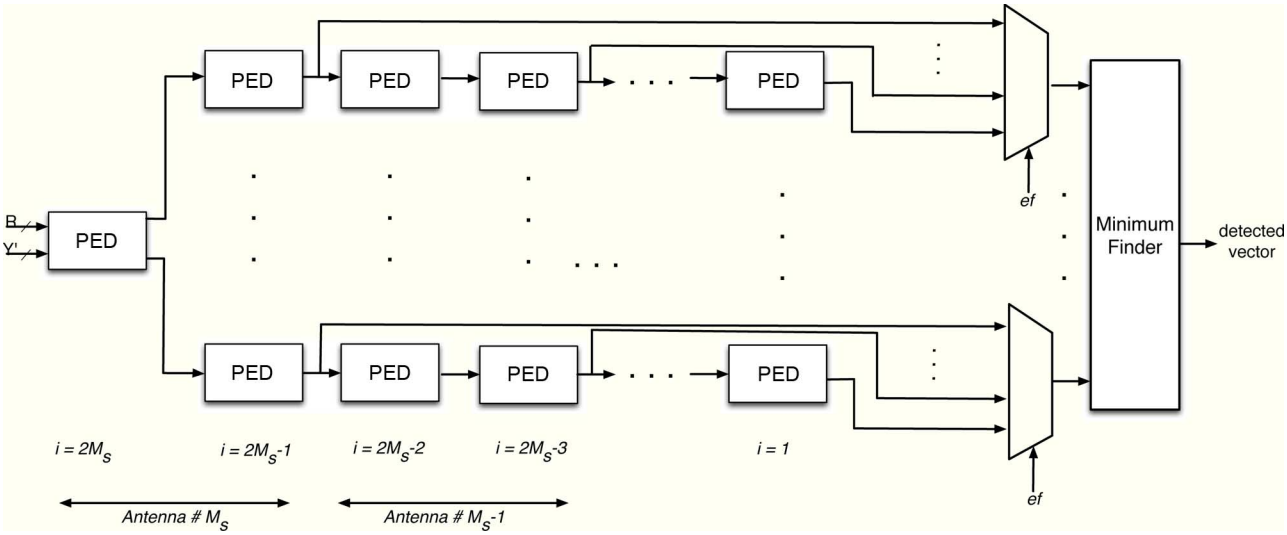


Fig. 13. The architecture for the cooperative partial detection in relay using Flex-Sphere.

TABLE I
SYNTHESIS RESULTS FOR A COMPLETE $M_s = 4$, VIRTEX-5, IMPLEMENTATION

Design	Flex-Sphere
Device	XC5VSX95
ef	1, 2, 3, 4
Max. Data Rate	857.1 Mbps
Number of Slices	11,604/14,720 (78 %)
Number of Registers/FFs	27,115/58,880 (46 %)
Number of Slice LUTs	33,427/58,880 (56 %)
Number of DSP48E/Multipliers	321/640 (50 %)
Number of block RAMs	0 (0 %)
Max. Freq.	285.71 MHz

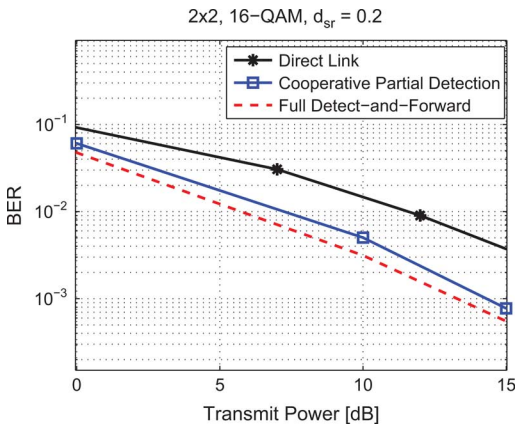


Fig. 14. BER comparison of the no-relay, CPD and FDF techniques using the WARP hardware platform at the 2.4 GHz band. The channel emulation is done using the Azimuth ACE 400 WB [28] channel emulator and the results include the RF effects.

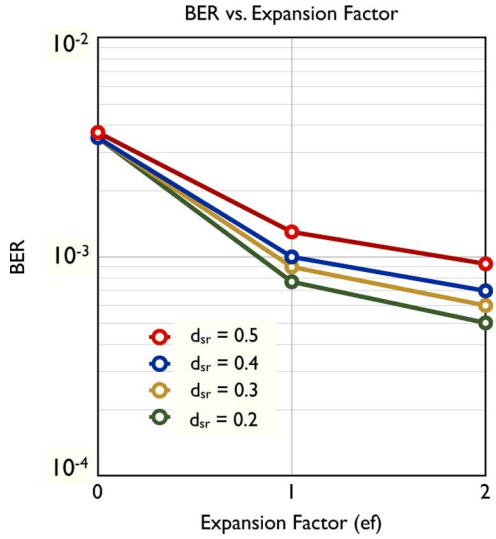


Fig. 15. WARPLab experimental results of BER versus different expansion factors for a 2×2 MIMO system and a total transmit power of 18 dBm. Note that $ef = 0$ correspond to no-relay and $ef = 2$ correspond to FDF.

The hardware emulation results using the platform are shown in Fig. 14 for a 2×2, 16-QAM system, where the relay is located at $d_{\rm sr} = 0.5$ and the power splitting ratio is $\mu = 0.5$ and the channel is a 3GPP Class B channel [28]. Fig. 15 also shows the BER performance comparison for different relay locations at fixed transmit power points. Since the tests are performed on a hardware platform, the performance curves take into account the effects of the baseband processing as well as the RF chain, e.g., the amplifiers, the AGC (automatic gain control), imperfect channel estimate, etc. In the presence of such effects, the CPD method provides a middle point that improves the performance compared to no-relay scenario while avoiding the larger complexity of the FDF method, which conforms with the simulation results for other systems dimensions.

VIII. CONCLUSION

This paper presented the novel CPD scheme for multiantenna relays. CPD utilizes the inherent structure of the tree-based sphere detectors and modifies the tree traversal so that instead of visiting all the levels of the tree, only a subset of the levels, thus a subset of the transmitted streams, are visited. The proposed scheme is based on architecture-friendly MIMO detection scenarios. We also developed a detection scheme based on combining of the received vectors. We analyzed the

complexity in the relay and destination and demonstrated, through simulations and WARPLab over-the-air verification, that this scheme can be used to distribute the computational processing between the source and the destination and more importantly, the relay can avoid the considerable overhead of MIMO detection while helping the source–destination link to improve its performance.

APPENDIX

In this section, we describe the derivation steps of (22) and (23). Starting from 20, we can expand the norms and keep the terms that depend on $\mathbf{x}$:

$$\arg\min_{\mathbf{x}\in\mathcal{O}^{M_s}} D(\mathbf{x}) = \tag{31}$$

$$\arg\min_{\mathbf{x}\in\mathcal{O}^{M_s}}(\| \mathbf{y}_d^{(2)} - \mathbf{H}_{rd}\mathbf{x}_1 \|_2^2 + \| \mathbf{y}_d^{(1)} - \mathbf{H}_1\mathbf{x}_1 - \mathbf{H}_2\mathbf{x}_2 \|_2^2) =$$

$$\arg\min_{\mathbf{x}\in\mathcal{O}^{M_s}}(A - B - B^* + g(\mathbf{y}_d^{(1)}, \mathbf{y}_d^{(2)})) \tag{32}$$

where $g(.)$ contains those terms that do not depend on $\mathbf{x}$ and, hence, will not affect the solution and A and B are given by

$$\begin{aligned} A =& \mathbf{x}_2^*\mathbf{H}_2^*\mathbf{H}_2\mathbf{x}_2 + \mathbf{x}_1^*\mathbf{H}_{rd}^*\mathbf{H}_{rd}\mathbf{x}_1 + \mathbf{x}_1^*\mathbf{H}_1^*\mathbf{H}_1\mathbf{x}_1 \\ &+ \mathbf{x}_1^*\mathbf{H}_1^*\mathbf{H}_2\mathbf{x}_2 + \mathbf{x}_2^*\mathbf{H}_2^*\mathbf{H}_1\mathbf{x}_1 \end{aligned} \tag{33}$$

$$\begin{aligned} B =& \mathbf{x}_2^*\mathbf{H}_2^*\mathbf{y}_d^{(1)} + \mathbf{x}_1^*\mathbf{H}_1^*\mathbf{y}_d^{(1)} + \mathbf{x}_1^*\mathbf{H}_{rd}^*\mathbf{y}_d^{(2)} \\ =& \mathbf{x}_1^*(\mathbf{H}_1^*\mathbf{y}_d^{(1)} + \mathbf{H}_{rd}^*\mathbf{y}_d^{(2)}) + \mathbf{x}_2^*\mathbf{H}_2^*\mathbf{y}_d^{(1)}. \end{aligned} \tag{34}$$

Comparing (32) with

$$\begin{aligned} &\| \mathbf{y}_{CPD} - \mathbf{H}_{CPD}\mathbf{x} \|_2^2 \\ &= \| \mathbf{y}_{CPD} \|_2^2 - \mathbf{x}^*\mathbf{H}_{CPD}^*\mathbf{y}_{CPD} - \mathbf{y}_{CPD}^*\mathbf{H}_{CPD}\mathbf{x} \\ &\quad + \mathbf{x}^*\mathbf{H}_{CPD}^*\mathbf{H}_{CPD}\mathbf{x} \end{aligned} \tag{35}$$

shows that the original problem in (31) is equivalent to

$$\arg\min_{\mathbf{x}\in\mathcal{O}^{M_s}} \left(\| \mathbf{y}_{CPD} - \mathbf{H}_{CPD}\mathbf{x} \|_2^2 \right) \tag{36}$$

if we set

$$\mathbf{H}_{CPD} = \begin{bmatrix} \mathbf{H}_1^*\mathbf{H}_1 + \mathbf{H}_{rd}^*\mathbf{H}_{rd} & \mathbf{H}_1^*\mathbf{H}_2 \\ \mathbf{H}_2^*\mathbf{H}_1 & \mathbf{H}_2^*\mathbf{H}_2 \end{bmatrix}^{\frac{1}{2}} \tag{37}$$

$$\mathbf{y}_{CPD} = \mathbf{H}_{CPD}^{-1} \begin{bmatrix} \mathbf{H}_1^*\mathbf{y}_d^{(1)} + \mathbf{H}_{rd}^*\mathbf{y}_d^{(2)} \\ \mathbf{H}_2^*\mathbf{y}_d^{(1)} \end{bmatrix}. \tag{38}$$

REFERENCES

[1] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, Electr. Eng. and Comput. Sci. Dept., Univ. of California, Berkeley, CA, 1968.

[2] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE. Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[5] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE. Trans. Signal Process.*, vol. 52, no. 2, pp. 362–371, 2004.

[6] M. Karkooti and J. R. Cavallaro, "Distributed decoding in cooperative communications," in *Proc. Asilomar Conf.*, Oct. 2007, pp. 824–828.

[7] I. E. Telatar, "Capacity of multiantenna Gaussian channels," in *Eur. Trans. Telecommun.*, Nov. 1999, vol. 10, no. 6, pp. 585–595.

[8] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," in *Wireless Pers. Commun.*, 1998, vol. 6, no. 3, pp. 311–335.

[9] E. Viterbo and J. Boutros, "A universal lattice decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

[10] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.

[11] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.

[12] M. Li, B. Bougard, W. Xu, D. Novo, L. van der Perre, and F. Catthoor, "Optimizing near-ML MIMO detector for SDR baseband on parallel programmable architectures," in *Proc. Conf. Design, Autom., Test in Eur.*, 2008, pp. 444–449.

[13] J. Antikainen, P. Salmela, O. Silvén, M. J. Juntti, J. H. Takala, and M. Myllyla, "Application-specific instruction set processor implementation of list sphere detector," *EURASIP J. Embedded Syst.*, vol. 2007, 2007 [Online]. Available: http://www.hindawi.com/journals/es/2007/054173/cta/, Article ID 54173

[14] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[15] C. K. Lo, S. Vishwanath, and R. W. Heath, "Rate bounds for MIMO relay channels using precoding," in *Proc. IEEE Global Telecommun. Conf.*, Nov. 2005, pp. 1172–1176.

[16] H. Bolcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006.

[17] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.

[18] C. Eklund, R. Marks, K. Stanwood, and S. Wang, "IEEE standard 802.16: A technical overview of the WirelessMANTM air interface for broadband wireless access," *IEEE Commun. Mag.*, vol. 40, no. 6, pp. 98–107, Jun. 2002.

[19] IMT-Advanced [Online]. Available: http://www.ieee802.org/21/doctree/IMT-Advanced/18-07-00xx-00-0000 IMT Advanced d3.doc

[20] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical solutions for the 3G long-term evolution," *IEEE Commun. Mag.*, vol. 44, no. 3, pp. 38–45, Mar. 2006.

[21] K. Amiri, M. Wu, M. Duarte, and J. R. Cavallaro, "Physical layer algorithm and hardware verification of MIMO relays using cooperative partial detection," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Jun. 2010, pp. 5614–5617.

[22] K. Amiri and J. R. Cavallaro, "Partial detection for multiple antenna cooperation," in *Proc. Conf. Inf. Sci. Syst.*, Mar. 2009, pp. 669–674.

[23] E. W. Jang, J. Lee, H. H. Lou, and J. M. Cioffi, "Optimal combining schemes for MIMO systems with hybrid ARQ," in *Proc. Int. Symp. Inf. Theory*, Jun. 2007, pp. 2286–2290.

[24] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," in *IEEE Trans. Commun.*, Mar. 2003, vol. 51, no. 3, pp. 389–399.

[25] Z. Guo and P. Nilsson, "Reduced complexity Schnorr–Euchner decoding algorithms for MIMO systems," in *IEEE Commun. Lett.*, May 2004, vol. 8, no. 5, pp. 286–288.

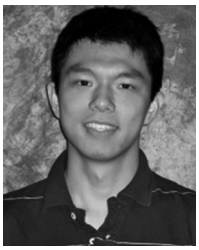
[26] K. Amiri, C. Dick, R. Rao, and J. R. Cavallaro, "A high throughput configurable SDR detector for multi-user MIMO wireless systems," *Springer J. Signal Process. Syst.*, vol. 62, no. 2, pp. 233–245, Feb. 2011.

[27] WARP [Online]. Available: https://www.warp.rice.edu

[28] Azimuth Systems [Online]. Available: http://www.azimuthsystems.com

Kiarash (Kia) Amiri received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2005 and the M.S. degree in electrical and computer engineering from Rice University, Houston, TX, in 2007. He is currently working towards the Ph.D. degree in electrical and computer engineering at Rice University, where he is a member of the Center for Multimedia Communication (CMC) Laboratory.

His research focus is in the area of physical layer design and hardware architecture for wireless communication. During summer and fall 2007, he worked on developing and implementing MIMO algorithms as part of the Advanced Systems Technology Group in Xilinx, San Jose, CA.



Michael Wu received the B.S. degree from Franklin W. Olin College, Needham, MA, in May 2007 and the M.S. degree from Rice University, Houston, TX, in May 2010, both in electrical and computer engineering. He is currently working towards the Ph.D. degree in the Electrical and Computer Engineering Department at Rice University.

His research interests are wireless algorithms, software defined radio on GPGPU and other parallel architectures and high performance wireless receiver designs.

Joseph R. Cavallaro (S'78–M'82–SM'05) received the B.S. degree from the University of Pennsylvania, Philadelphia, in 1981, the M.S. degree from Princeton University, Princeton, NJ, in 1982, and the Ph.D. degree from Cornell University, Ithaca, NY, in 1988, all in electrical engineering.

From 1981 to 1983, he was with AT&T Bell Laboratories, Holmdel, NJ. In 1988, he joined the faculty of Rice University, Houston, TX, where he is currently a Professor of electrical and computer engineering. His research interests include computer arithmetic, VLSI design and microlithography and DSP and VLSI architectures for applications in wireless communications. During the 1996–1997 academic year, he served at the National Science Foundation as Director of the Prototyping Tools and Methodology Program. He was a Nokia Foundation Fellow and a Visiting Professor at the University of Oulu, Finland, in 2005 and continues his affiliation there as an Adjunct Professor. He is currently the Director of the Center for Multimedia Communication at Rice University.

Dr. Cavallaro was Co-Chair of the 2004 Signal Processing for Communications Symposium at the IEEE Global Communications Conference and General/Program Co-Chair of the 2003, 2004, and 2011 IEEE International Conference on Application-Specific Systems, Architectures and Processors (ASAP).

Jorma Lilleberg was born in Rovaniemi, Finland, in 1953. He received the Diploma Engineer and Licentiate of Technology degrees in electrical engineering at the University of Oulu, Oulu, Finland, in 1979 and 1984, respectively, and the Doctor of Technology degree at the Tampere University of Technology, Tampere, Finland, in 1992.

During 1992–1993, he worked at the Technical Research Center of Finland, Oulu, Finland, as an acting Research Professor and Chief Scientist for signal processing. From 1993 to 2010, he worked at Nokia, Oulu, Finland, as a Principal Scientist, Technology Fellow, and Distinguished Research Leader. He is currently working at Renesas Mobile Corporation as Distinguished Research Leader. His research interests are in digital communications theory and application of statistical signal processing methods for digital radio networks. He has coauthored more than 90 research papers and holds more than 19 patents. He is a Docent at the University of Oulu, Oulu, Finland, and an Adjunct Professor at Rice University, Houston, TX. He was also a visiting Professor at the Chinese Academy of Science, Shanghai Research Center for Wireless Communications in Shanghai, China from 2006 to 2010.