Low Complexity Resource Allocation for QF VMIMO Receivers with a Shared Backhaul

Tim Rüegg, Yahia Hassan, Armin Wittneben
Swiss Federal Institute of Technology (ETH) Zurich,
Communication Technology Laboratory, CH-8092 Zurich, Switzerland
Email: {rueegg,hassan,wittneben}@nari.ee.ethz.ch

***Abstract*—In this paper, we consider the resource allocation problem of a quantize-and-forward virtual MIMO receive cooperation scheme, where N sources simultaneously communicate over N relays to a destination. Different to previous work, the second hop links are considered to be non-orthogonal (i.e. a shared backhaul) but with individual rate constraints. This requires a careful resource allocation, leading to an involved non-convex optimization problem. The goal of this paper is to provide and evaluate computationally simple and robust resource allocation algorithms. In this context, we devise two low complexity schemes which simply select a subset of relays and assign equal resources to them. In thorough numerical simulations we then demonstrate that for moderate backhaul rates (as to be expected in wireless virtual MIMO), it is beneficial to serve only a subset of the relays. Furthermore, while large gains are achievable by selecting the optimal relays and assigning equal resources to them, only minor additional gains can be achieved by further optimizing the resource allocation among them, e.g., with a gradient search. Hence, the proposed relay selection schemes lead to very good performance and are a reasonable alternative to costly optimization techniques if low complexity is required.**

I. Introduction

In this paper, we consider a wireless virtual multiple-input multiple-output (MIMO) receive cooperation scheme with a setup as shown in Fig. 1. Multiple non-cooperating sources individually transmit over multiple quantize-and-forward (QF) relays to a final destination which jointly decodes the source data streams based on the quantized observations at the relays. Such a setup is of high practical relevance in military as well as civil mobile ad hoc networks (MANETs) and sensor networks in order to increase the transmission range and/or the spectral efficiency (e.g. [1]–[3]). It could reflect a direct communication between two spatially separated clusters of nodes (e.g. a mobile military unit with multiple sources transmitting to another mobile unit or the headquaters over large distances), or equivalently the communication between many independent nodes to one destination supported by multiple relays (e.g. multiple soldiers distributed in space simultaneously transmitting to their commander without a direct link).

In the following, we will focus on the military application and thus consider all nodes to be half-duplex with

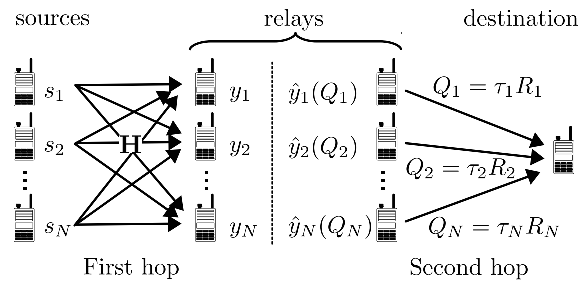

Fig. 1. System setup.

a single antenna, motivated by the typical downward compatibility and evolutionary growth requirements of military networks [4]. That is, the relays can not simultaneously receive and transmit, and the first and second hop have to be orthogonalized *in time*. The wireless links between the relays and the final destination (further called backhaul links) are assumed to support given individual rates R_i. However, as the backhaul is shared among the relays, it is divided into orthogonal resource blocks, e.g. using a time division multiple access (TDMA) scheme. The sizes τ_i of the resource blocks have a strong impact on the performance of the system as they determine the total duration of the second hop, as well as the number of bits $Q_i = \tau_i \cdot R_i$ which each QF relay can deliver per time unit to the destination (i.e. the τ_i determine the quantization rate per relay). Hence, the backhaul resources need to be assigned carefully in order to optimize the system performance.

This problem has not been considered so far. Related work has primarily been performed in the context of distributed MIMO receivers and cloud radio access networks (CRANs) (e.g. [5]–[9]). In [5] and [6], a two user setup with orthogonal backhaul links is considered. While [5] analyses the performance of QF and compress-and-forward (CF) for given backhaul link rates, [6] considers the same problem with random fluctuations on the backhaul. Different to QF, CF includes distributed source coding, leading to much higher complexity at the encoding and decoding and requiring full channel state information (CSI) at the relays [5]. In [7], a CRAN uplink scenario with orthogonal backhaul links is considered and the optimal quantization noise levels for different CF schemes are derived. Similarly but for a shared backhaul with a sum-capacity constraint, in [8] the optimal quantization noise level for each relay is derived at high signal-to-quantization-noise-ratio (SQNR) for QF and CF, and low complexity algorithms are presented for the quantization noise level design. In [9], the setups of [7] and [8] are extended to multi-antenna terminals and investigated in terms of their achievable rate regions.

Different to the setup at hand, in [5]–[9] all relays are considered to be full duplex (i.e. first and second hop are performed simultaneously and the second hop duration is thus fixed), and the second hop is either already considered orthogonalized (i.e. no resource allocation is necessary [5]–[7], [9]), or with a sum capacity constraint [8], [9]. However, especially in wireless virtual MIMO setups non-orthogonal backhaul links as well as individual rate constraints are of high practical relevance, as each link interferes with the others and experiences independent shadowing and fading. This motivates the work presented in this paper.

Assigning the backhaul resources is an involved non-convex optimization problem. However, the goal of this paper is to provide simple and robust resource allocation algorithms, which are well suited for military applications. Hence, as military equipment tends to boost long life time cycles [4], the proposed algorithms should be implementable as an incremental update on existing hardware (half-duplex nodes with single-antennas), requiring only low computational complexity.

In this context, we introduce the non-convex resource allocation problem, discuss the mathematical challenge of it and devise two suboptimal but computationally simple resource allocation approaches. These approaches simply select a subset of active relays (either based on their signal-to-noise-ratio (SNR) or on their SQNR) and assign equal resources to them. Furthermore, we derive the analytical gradient for a gradient search with multiple initializations. In numerical simulations, these approaches are thoroughly evaluated in different operating regimes and the resulting insights are discussed. Large gains can be achieved by applying the virtual MIMO scheme compared to the conventional approach of using a single relay only. However, in order to utilize its full potential, a proper resource allocation is crucial. This is illustrated in Fig. 2. It shows the empirical cumulative distribution function (CDF) of the achievable throughput of the conventional approach, a virtual MIMO scheme with equal resources for all relays, and virtual MIMO with optimized τ_i (details are provided in Sec. IV). We furthermore demonstrate, that while it is crucial to select the optimal subset of active relays, the optimization of the τ_i among them only leads to minor additional gains. The proposed relay selection schemes thus achieve a performance close to the gradient search while requiring substantially less computational complexity and CSI.

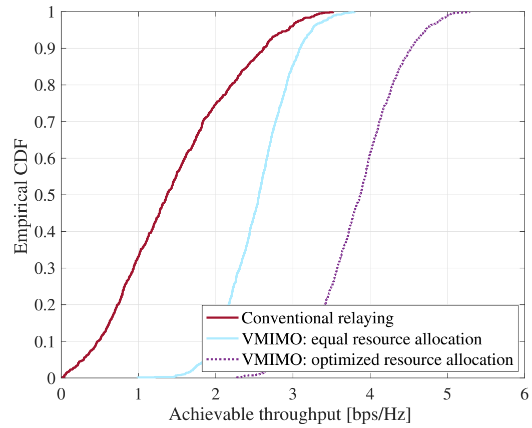

Fig. 2. Achievable throughput of a conventional approach (one relay only), of virtual MIMO with equal resources for all relays and of virtual MIMO with optimized resource allocation.

II. System Model and Problem Formulation

Throughout the paper, we consider the system model as shown in Fig. 1 with N sources and N half-duplex relays. All nodes have a single antenna. The half duplex assumption imposes quite stringent constraints on the feasible traffic patterns. Specifically, we assume that:

- Relays cannot receive on the first hop while transmitting on the second hop.
- If any one relay transmits, non of the other relays can receive simultaneously.

In order to maximize the compatibility with existing equipment and support evolutionary development, we furthermore assume that the relays transmit one at a time (thus the destination is not required to resolve collisions). The allocated time to relay i is denoted by τ_i. For the clarity of exposition we normalize without loss of generality all time intervals to the duration of 1 channel use of the first hop.

The wireless channel between the sources and the relays is denoted by $\mathbf{H} \in \mathbb{C}^{N \times N}$, with h_{ij} the channel from source j to relay i. The transmit symbol of source i is assumed to be $s_i \sim \mathcal{CN}(0, 1) \ \forall \ i \in \{1, \dots, N\}$. The received signal of the relays is then a superposition of all transmit signals

$$y_i = \sum_{j=1}^{N} h_{ij} \cdot s_j + n_i, \qquad (1)$$

where $n_i \sim \mathcal{CN}(0, \sigma_n^2)$ denotes the additive white Gaussian noise (AWGN), and the variance of the received signal is given as

$$\sigma_{y_i}^2 = \sum_{j=1}^{N} |h_{ij}|^2 + \sigma_n^2. \qquad (2)$$

The backhaul link from relay i to the destination is assumed to support rate R_i with negligible probability of error (i.e. relay i can transmit R_i information bit to the destination in the time which one channel use takes on the first hop). If we let τ_{tot} be the normalized duration of the second hop, the *effective quantization rate* of the relays is given as $Q_i = \tau_i \cdot R_i$, with $\sum_{i=1}^{N} \tau_i = \tau_{\text{tot}}$ and $\tau_i \geq 0 \ \forall \ i \in \{1, \ldots, N\}$. Considering a vector quantizer at the relays, the quantization noise is additive and Gaussian distributed [10]. The *achievable decoding rate* R^{QF} at the destination in bits per first hop channel uses is thus given as

$$R^{\text{QF}} = \log_2 \det \left(\mathbf{I} + (\mathbf{D}_n + \mathbf{D}_q)^{-1} \mathbf{\Lambda}_s \right), \tag{3}$$

with $\mathbf{I}$ the identity matrix, the signal covariance matrix $\mathbf{\Lambda}_s = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ and the two diagonal noise covariance matrices $\mathbf{D}_n = \mathbf{I} \cdot \sigma_n^2$ and $\mathbf{D}_q = \operatorname{diag}\left(\sigma_{q_1}^2, \ldots, \sigma_{q_N}^2\right)$ of the AWGN and the quantization noise respectively.

$$\sigma_{q_i}^2 = \frac{\sigma_{y_i}^2}{2^{Q_i} - 1} = \frac{\sigma_{y_i}^2}{2^{\tau_i R_i} - 1} \tag{4}$$

denotes the quantization noise variance of relay i, which depends on the effective quantization rate Q_i and thus on τ_i.

In total, a 2-hop transmission cycle comprises $1 + \tau_{\text{tot}}$ time units. Hence, the resulting *overall throughput* R_{MS} of the system is given as

$$R_{\text{MS}} = \frac{1}{1 + \tau_{\text{tot}}} \cdot R^{\text{QF}}. \tag{5}$$

Thereby, τ_{tot} has two opposing effects on the throughput R_{MS}. For increasing τ_{tot}, the achievable decoding rate at the final destination R^{QF} is increasing (if the τ_i are reasonably assigned), but the factor in front of R^{QF} is decreased, due to the longer second hop. That is, the trade-off leading to the maximal throughput has to be found. This optimization problem can be stated as

$$\begin{aligned} [\hat{\tau}_1, \ldots, \hat{\tau}_N] \quad &= \quad \underset{\tau_1, \ldots, \tau_N}{\arg\max}\, R_{\text{MS}} \\ \text{s.t.} \quad & \quad \tau_i \geq 0 \ \forall i \in \{1, \ldots, N\}, \end{aligned} \tag{6}$$

which inherently finds the optimal $\tau_{\text{tot}} = \sum_{i=1}^{N} \tau_i$.

Note: Different to [8] where the quantization noise level is optimized for full-duplex relays (i.e. $\tau_{\text{tot}} = 1$) under a sum capacity constraint C_{sum} (i.e. all Q_i could take on any value in $[0, C_{\text{sum}}]$ as long as $\sum_i Q_i \leq C_{\text{sum}}$), τ_{tot} is variable in our setup and the quantization rates are limited to $Q_i \in [0, \tau_{\text{tot}} \cdot R_i]$, which is different for every relay.

III. Backhaul Resource Allocation

The optimization problem (6) can be approached with the Karush-Kuhn-Tucker (KKT) conditions. Reformulating the inequality constraint $\tau_i \geq 0, \ \forall i \in \{1, \ldots, N\}$ into $-\tau_i \leq 0, \ \forall i \in \{1, \ldots, N\}$, we get the Lagrange function as

$$\mathcal{L}(\boldsymbol{\tau}, \boldsymbol{\mu}) = R_{\text{MS}} + \sum_{i=1}^{N} \mu_i \tau_i, \tag{7}$$

and the KKT conditions as

$$\begin{aligned} \nabla \mathcal{L}(\boldsymbol{\tau}, \boldsymbol{\mu}) \quad &= \quad 0 && (8) \\ -\tau_i \quad &\leq \quad 0 \ \forall i \in \{1, \ldots, N\} && (9) \\ \mu_i \quad &\geq \quad 0 \ \forall i \in \{1, \ldots, N\} && (10) \\ \mu_i \tau_i \quad &= \quad 0 \ \forall i \in \{1, \ldots, N\}. && (11) \end{aligned}$$

The partial derivatives of the gradient are thereby given as

$$\frac{\partial \mathcal{L}(\boldsymbol{\tau}, \boldsymbol{\mu})}{\partial \tau_i} = \frac{\partial R_{\text{MS}}}{\partial \tau_i} + \mu_i. \tag{12}$$

Hence, a relay is either active (i.e. $\tau_i > 0$), leading to $\mu_i = 0$ (from (11)) and thus $\partial R_{\text{MS}} / \partial \tau_i = 0$ (from (8)), or it is inactive ((i.e. $\tau_i = 0$), leading to $\mu_i \geq 0$ and thus $\partial R_{\text{MS}} / \partial \tau_i \leq 0$. That is, there is no gain in the achievable rate if incremental resources are reassigned from one relay to another and a local maximum is achieved.

The partial derivatives of R_{MS} can be found as

$$\frac{\partial R_{\text{MS}}}{\partial \tau_i} = R^{\text{QF}} \frac{\partial}{\partial \tau_i} \frac{1}{1 + \sum_{j=1}^{N} \tau_j} + \frac{1}{1 + \sum_{j=1}^{N} \tau_j} \frac{\partial}{\partial \tau_i} R^{\text{QF}}, \tag{13}$$

with

$$\frac{\partial}{\partial \tau_i} \frac{1}{1 + \sum_{j=1}^{N} \tau_j} = -\frac{1}{(1 + \sum_{j=1}^{N} \tau_j)^2} \tag{14}$$

and

$$\begin{aligned} \frac{\partial R^{\text{QF}}}{\partial \tau_i} &= \frac{\partial}{\partial \tau_i} (\log_2 \det (\mathbf{D}_q + \mathbf{D}_n + \mathbf{\Lambda}_s) \\ &\quad - \log_2 \det (\mathbf{D}_q + \mathbf{D}_n)) \\ &= \operatorname{tr}\Big((\mathbf{D}_q + \mathbf{D}_n + \mathbf{\Lambda}_s)^{-1} \frac{\partial}{\partial \tau_i} \mathbf{D}_q \Big) \\ &\quad - \operatorname{tr}\Big((\mathbf{D}_q + \mathbf{D}_n)^{-1} \frac{\partial}{\partial \tau_i} \mathbf{D}_q \Big) \\ &= \Big(((\mathbf{D}_q + \mathbf{D}_n + \mathbf{\Lambda}_s)^{-1})_{ii} - ((\mathbf{D}_q + \mathbf{D}_n)^{-1})_{ii} \Big) \\ &\quad \cdot \frac{R_i \sigma_{y_i}^2 2^{\tau_i R_i}}{(2^{\tau_i R_i} - 1)^2}, \end{aligned} \tag{15}$$

where the last step follows from the fact that in $\partial / \partial \tau_i \, \mathbf{D}_q$ only the element in row i and column i (denoted by $(\cdot)_{ii}$) is unequal to zero. It can be seen, that the partial derivatives of R_{MS} are coupled in all the τ_i and the problem can not be solved elementary. Furthermore, the problem is non-convex with multiple local optima.

Nevertheless, a numerical gradient search with multiple initializations can be applied. The higher the number of random initializations, the higher is the probability that the optimal solution is found.

A suboptimal solution can be found by approximating $\mathbf{\Lambda}_s$ by its diagonal and then first maximizing R^{QF} for a fixed τ_{tot} (with the KKT conditions), and afterwards optimizing over τ_{tot} by a one dimensional line search, consistently adapting the τ_i. However, due to space limitations, we omit the discussion of this approach in this paper and rather focus on two low complexity relay selection schemes in the following. They are then compared to the reference approach of equal time shares for all relays and to a gradient search with multiple initializations. For all of them, the τ_i are allocated for a fixed τ_{tot} and then various τ_{tot} are considered in order to illustrate the system performance in dependence of the second hop duration.

A. Reference Approach: Equal Time Allocation

As a baseline we consider equal resource allocation for all relays, i.e. $\tau_i = \tau_{\text{tot}}/N \ \forall \ i \in \{1, \ldots, N\}$. Not taking the information content of the observations of the relays and the backhaul rates into account can lead to highly suboptimal results for setups with strong variations in backhaul rates and signal strength at the relays.

B. First hop relay selection (1RS)

Depending on the backhaul rates and the available resources, it is crucial to only consider a subset of relays for the resource allocation. To keep the approach as simple as possible we just select N_a active relays based on their receive $\text{SNR}_i = \frac{\lambda_{ii}}{\sigma_n^2}$, with λ_{ii} the element of $\mathbf{\Lambda}_s$ in row i and column i, and assign equal $\tau_i = \tau_{\text{tot}}/N_\text{a}$ to them. The N_a as well as the τ_{tot} are thereby considered to be pre-determined. That is, the only necessary information for the relay selection is the SNR of the first hop. The crucial part however, is to choose τ_{tot} and N_a in advance (e.g. based on experience).

Note: For $N_\text{a} = N$, it corresponds to the reference approach of equal time-shares for all relays. For $N_\text{a} = 1$, no virtual MIMO is applied and simply the best relay is selected based on its SNR.

C. Second hop relay selection (2RS)

Similar to the previous approach, the second hop relay selection considers only a subset of N_a active relays. However, as the backhaul rates might have a significant impact on the performance, the relays are chosen based on their $\text{SQNR}_i = \frac{\lambda_{ii}}{\sigma_n^2+\sigma_{q_i}^2}$, with $\tau_i = \tau_{\text{tot}}/N_\text{a}$. That is, both hops are incorporated into the choice and relays with high backhaul rate but similar received signal power as the others are favored. Again, the N_a as well as the τ_{tot} are considered to be pre-determined. Additional to the SNRs also the backhaul rates have to be known for the choice.

For $N_\text{a} = N$ it also corresponds to the reference approach.

D. Gradient Search

The optimal solution of (6) can be found e.g. by a gradient search with sufficient initializations at the price of strongly increased computational effort. Nevertheless, as a reference for the proposed resource allocation schemes, a gradient search which optimizes R^{QF} for a given τ_{tot} is considered in the evaluations as well.

TABLE I
OPERATION REGIMES.

	Weak shadowing	Strong shadowing
clustered relays	I	IV
scattered relays	II	III

E. Complexity Comparison

While the gradient search requires in each step the evaluation of N partial derivatives which either include inverses of full matrices (analytical gradient) or determinants of full matrices (numerical gradient), both with high computational complexity, the computational complexity of the relay selection approaches is extremely low (determine N SNRs, respectively SQRNs and choose the strongest relays). Furthermore, the relay selection schemes only require the λ_{ii} and R_i (for the second hop relay selection), while the gradient approach requires the full knowledge of $\mathbf{\Lambda}_s$ and the R_i. Hence, more data needs to be exchanged.

IV. Numerical Evaluations

In the following simulations we consider a setup as described in Section II with $N = 10$ sources and QF relays. On the first hop, the channel coefficients from all sources to a given relay are considered i.i.d. complex Gaussian. We define $\overline{\text{SNR}}_i$ as the average SNR at relay i if all sources are transmitting and distinguish two relay deployment scenarios:

- Relay cluster: The relays are spatially close compared to the source-relay distances and thus have similar average receive SNR, $\{\overline{\text{SNR}}_i\} = [13, 13.22, \ldots, 15]^\mathsf{T}$ dB.
- Scattered relays: The relays are spatially dispersed and thus have substantially different average receive SNR of $\{\overline{\text{SNR}}_i\} = [2, 4, \ldots, 20]^\mathsf{T}$ dB.

The resulting channel matrix is then determined as

$$\mathbf{H} = \text{diag}(\sqrt{\mathbf{snr}/N}) \cdot \mathbf{F}, \tag{16}$$

with $\mathbf{F}$ the fading matrix with i.i.d. $\sim \mathcal{CN}(0,1)$ elements and the average SNRs stacked into the vector $\mathbf{snr}$.

On the second hop, all the rates R_i are considered to be statistically independent random variables and we distinguish two destination shadowing scenarios:

- Weak destination shadowing: The variance in the shadowing is small and the R_i are uniformly drawn from the interval $[4.5, 5.5]$ bps/Hz.
- Strong destination shadowing: The variance in the shadowing is large and the R_i are uniformly drawn from the interval $[0, 10]$ bps/Hz.

Combining the different first hop and second hop scenarios we end up with the four typical operation regimes shown in Table I. For all simulations σ_n^2 is set to 1. For each setup 1000 Monte-Carlo simulations are conducted.

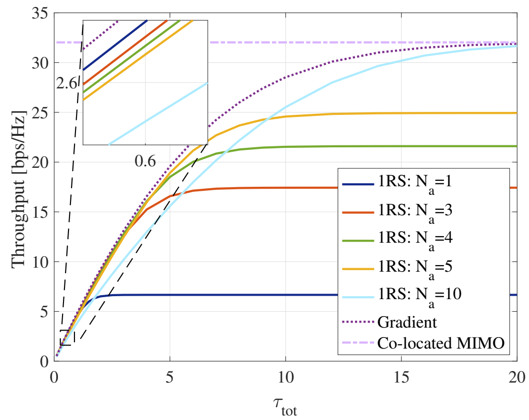

Fig. 3. Average decoding rate R^{QF} in regime II with $\mathbf{snr} = [2, 4, \ldots, 20]$ dB and $R_i \sim \mathcal{U}(4.5, 5.5)$ bps/Hz for the first hop relay selection scheme.

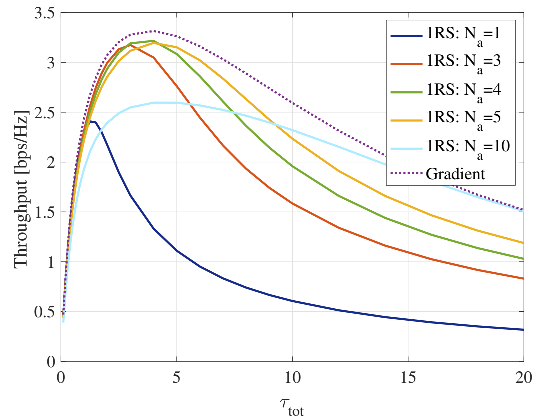

Fig. 4. Total average throughput R_{MS} in regime II with $\mathbf{snr} = [2, 4, \ldots, 20]$ dB and $R_i \sim \mathcal{U}(4.5, 5.5)$ bps/Hz for the first hop relay selection scheme.

A. Results

In **regime I** with similar SNRs and weakly varying backhaul rates, the results and conclusions are similar to regime II, on which we will focus in the sequel.

Regime II features strongly varying SNRs and similar backhaul rates. In Fig. 3 the average achievable decoding rate at the final destination R^{QF} (3), is shown for the first hop relay selection, for the gradient search, as well as for the co-located MIMO case, i.e. without any quantization noise. While huge degeneration of the achievable rates can be observed for all schemes compared to the co-located MIMO case at low τ_{tot} (due to the strong quantization noise), the performance is improving for increasing τ_{tot}. However, the performance of the relay selection scheme is limited by its maximal spatial degrees of freedom (DOF) which are determined by the number of active relays N_{a}. This is clearly visible in Fig. 3. The curves for the different N_{a} only achieve a certain plateau and can not improve anymore for higher τ_{tot}. The $\sigma^2_{q_i}$ are already much smaller than the thermal noise σ^2_n and thus do not affect the performance significantly anymore. Obviously, the higher N_{a}, the higher is the achievable performance if τ_{tot} is large enough, and eventually the co-located MIMO bound is achieved for $N_{\text{a}} = 10$. However, for lower τ_{tot} (i.e. lower resulting SQNRs) a lower N_{a} is beneficial. Hence, it is best to only serve a subset of the relays at low SQNR. As expected, the gradient search leads to the best performance for all τ_{tot}, as it adaptively chooses the optimal number of relays and additionally optimizes the assignment of the τ_i.

The curves for the second hop relay selection are almost identical and therefore not shown. As the variance in the backhaul rates is small, the best relays are mainly determined by the first hop SNR. Hence, for similar backhaul rates, first hop relay selection is sufficient.

The corresponding total average throughput R_{MS} (5) is shown in Fig. 4. The trade-off between high R^{QF} and long backhaul access time can be clearly observed for all curves. However, depending on the number of active relays N_{a}, the peak performance is achieved at a different τ_{tot} and the width of the curves varies strongly. This comes from the fact, that for smaller N_{a} the plateau is achieved at lower τ_{tot}. As soon as the plateau is reached, the performance drops strongly. Eventually, the R_{MS} of all schemes will tend to 0 as τ_{tot} grows large. The best peak performance in this setup is achieved for $N_{\text{a}} = 4$, getting close to the performance of the gradient search. Although a large loss can be observed compared to the co-located MIMO case (due to the quantization noise and the smaller DOF), large gains can be achieved compared to $N_{\text{a}} = 1$ and $N_{\text{a}} = 10$. That is, applying a virtual MIMO scheme based on simple relay selection can strongly increase the performance in such a setup, if τ_{tot} and N_{a} are reasonably chosen.

In **regime III** with strongly varying SNRs ($\mathbf{snr} = [2, 4, \ldots, 20]^{\mathsf{T}}$ dB) and large differences in the backhaul rates ($R_i \sim \mathcal{U}(0, 10)$ bps/Hz), the backhaul rates R_i have a significant impact on the performance. This can be seen in Fig. 5 which shows the total average throughput R_{MS} of both relay selection schemes for their specific best average N_{a} ($N_{\text{a}} = 5$ for 1RS respectively $N_{\text{a}} = 3$ for 2RS), as well as for the reference schemes. Although the first hop relay selection still leads to significant gains with the optimal N_{a} compared to $N_{\text{a}} = 10$, there is a large loss compared to the second hop relay selection. That is, the backhaul rates R_i have a strong impact on the performance. A node with high SNR but low backhaul rate is wasting a lot of resources compared to a node with slightly lower SNR but much higher backhaul rate. Hence, the backhaul rates should be included in the relay selection if the variance among them is large.

Note, that besides the gain in the maximal total average throughput, also large gains considering the delay

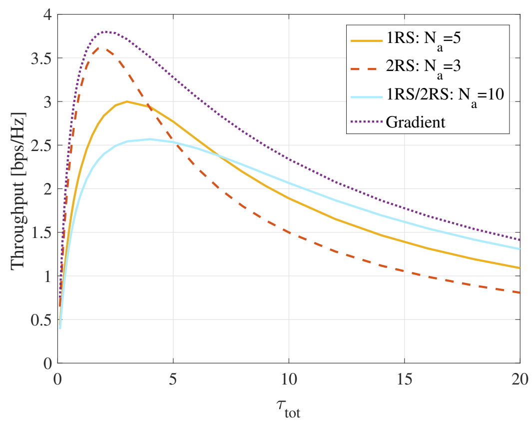

Fig. 5. Total average throughput R_{MS} in regime III with $\mathbf{snr} = [2, 4, \ldots, 20]$ dB and $R_i \sim \mathcal{U}(0, 10)$ bps/Hz for both relay selection schemes.

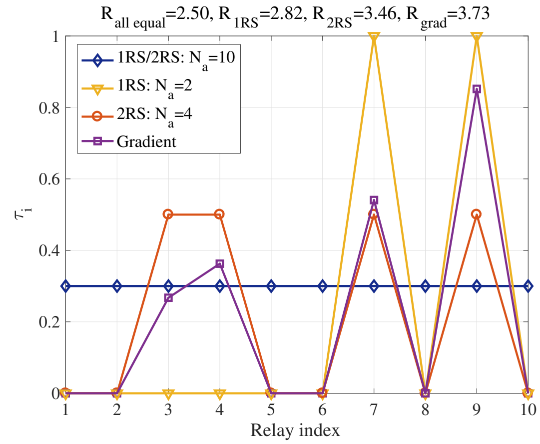

Fig. 6. Instantaneous τ_i at the respective peak performance for one specific channel realization in regime III with $\mathbf{snr} = [2, 4, \ldots, 20]$ dB and $R_i \sim \mathcal{U}(0, 10)$ bps/Hz.

can be achieved at the same performance. In contrast to the first hop relay selection, where the peak performance is achieved at $\tau_{\mathrm{tot}} = 3$, the second hop relay selection only requires $\tau_{\mathrm{tot}} = 0.7$ for the same throughput.

Apparently, the second hop relay selection is very efficient, getting close to the gradient search performance. This gap could even be made smaller by weighting the τ_i of each selected relay. However, this is not considered in this paper. The most important gains are already achieved by choosing the best relays, and only minor additional gains are achieved with the optimal τ_i.

This is also visualized in Fig. 6, where for one specific channel realization the instantaneous τ_i are shown for both relay selection schemes, the gradient search and the reference approach at their respective peak performance. The title of the figure also shows the corresponding instantaneous total throughput. While first hop relay selection achieves its best performance for $N_{\mathrm{a}} = 2$, the second hop relay selection chooses the same $N_{\mathrm{a}} = 4$ relays as the gradient search, achieving large gains compared to the reference approach and the first hop relay selection. The gradient search in the end only achieves minor additional gains by optimizing the τ_i, although especially for τ_3 and τ_9, the differences are quite significant.

Except for the instantaneous total throughput in Fig. 6, only averaged values have been considered so far. Fig. 7 shows the empirical CDF of the instantaneous total throughput of the two relay selection schemes if always their average best τ_{tot} and N_{a} are chosen (i.e. for $\tau_{\mathrm{tot}} = 3$ and $N_{\mathrm{a}} = 5$ for 1RS and $\tau_{\mathrm{tot}} = 2$ and $N_{\mathrm{a}} = 3$ for 2RS). It furthermore shows the empirical CDF of the two relay selection schemes and the gradient search if always the maximal instantaneous total throughput is considered (i.e. for varying τ_{tot} and N_{a}). For the second hop relay selection scheme only minor gains can be achieved by always choosing the optimal τ_{tot} and N_{a}. Hence, considering a fixed number of active relays and a fixed length for the second hop is reasonable if chosen wisely. This could strongly simplify the protocol in practice, as N_{a} and τ_{tot} could be set initially, e.g., based on experience, and then considered constant as long as the setup does not change significantly.

For the first hop relay selection scheme, the achievable gain is much higher. That is, it is more sensitive to N_{a} and τ_{tot}. This comes from the strongly varying backhaul rates, which are not incorporated into the relay selection of the first scheme, but have a high impact on the performance.

Fig. 2 in Sec. I shows the corresponding curves at the instantaneous optimal τ_{tot} for the case where always 1 relay is chosen randomly (i.e. no knowledge at all about the first and second hop is required), for the reference case of $N_{\mathrm{a}} = 10$, as well as for the gradient search. This figure impressively visualizes how valuable it is to apply virtual MIMO, and furthermore, how large the gains with a proper resource allocation can be.

Of course, not only the variance of the backhaul rates strongly affects the performance, but also their level. For high backhaul rates, much better quantization can be achieved with equal delay (τ_{tot}) and thus, the throughput is strongly improved. This can be observed in Fig. 8.

In the upper part of the figure the achievable throughput is shown for the gradient search as well as for the second hop relay selection scheme for $N_{\mathrm{a}} = 1$ and for the optimum average N_{a} for random backhaul rates drawn from $\mathcal{U}(0, 100)$. Such large backhaul rates could be achieved by using a much larger bandwidth on the second hop relative to the first hop (e.g. second hop in the 60 GHz band as proposed in [2] for civil applications). Due to the high backhaul rates, the optimal N_{a} increases to 6, i.e. the spatial degrees of freedom are strongly increased. This leads to large gains compared to $N_{\mathrm{a}} = 1$ (i.e. the benefit of virtual MIMO is large).

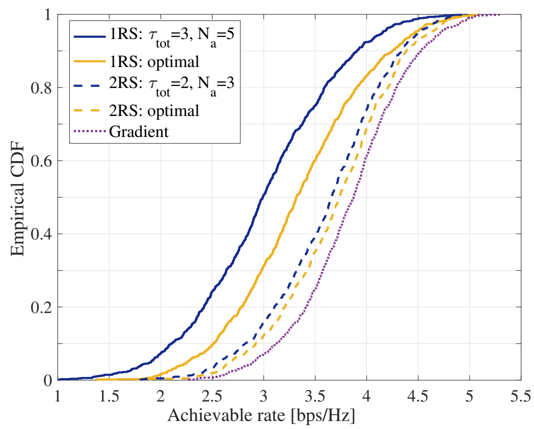

Fig. 7. Instantaneous total throughput at the instantaneous optimal $\tau_{\rm tot}$ and $N_{\rm a}$, and at the average optimal $\tau_{\rm tot}$ and $N_{\rm a}$ in regime III with $\mathbf{snr} = [2, 4, \ldots, 20]$ dB and $R_i \sim \mathcal{U}(0, 10)$ bps/Hz.

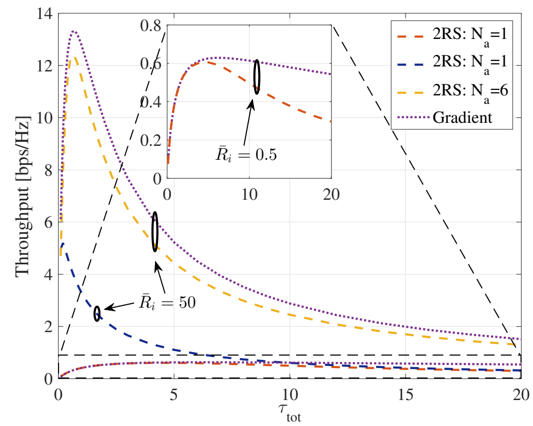

Fig. 8. Total average throughput $R_{\rm MS}$ for $\mathbf{snr} = [2, 4, \ldots, 20]$ dB and $R_i \sim \mathcal{U}(0, 1)$ with average rate $\bar{R}_i = 0.5$ bps/Hz, and $R_i \sim \mathcal{U}(0, 100)$ with average rate $\bar{R}_i = 50$ bps/Hz.

In contrast to these large gains, the lower part of the figure shows the achievable throughput for random backhaul rates drawn from $\mathcal{U}(0, 1)$. As the resulting SQNR is very low, the optimum number of relays is $N_{\rm a} = 1$. As to be expected, the lower the backhaul rates, the lower is the benefit of virtual MIMO. Still, the choice of the optimal relays is crucial.

Regime IV: As the main influence factor for the relay selection is again the backhaul rates, the characteristics of the results of regime IV are very similar to the results of regime III and are thus omitted as well.

V. Conclusions

In this paper, we considered the resource allocation problem of quantize and forward virtual MIMO receivers with a shared backhaul and individual rate constraints. Due to the shared backhaul, only one relay can forward its observation at the time. Thus, a proper resource allocation in such a setup is crucial. In this context, we thoroughly evaluated two different simple relay selection schemes and compared their performance to a gradient search resource allocation, and the reference approach of equal resources for all relays. It has been shown, that depending on the backhaul rates (i.e. the resulting signal to quantization noise ratio), it is best to only serve a subset of relays. While the most significant gains are achieved by selecting the optimal relays and assigning equal resources to them, only minor additional gains can be achieved by also optimizing the distribution of the backhaul resources among them. Hence, a simple relay selection scheme with equal resource allocation leads to very good performance at low computational complexity and limited required CSI. Depending on the variance of the backhaul rates, already the received SNR is sufficient for the relay selection (at low variance), or the SQNR should be considered (at high variance). Due to their robustness and the ability to be implemented on existing hardware (half-duplex single antenna nodes, requiring low computational complexity), they are well suited for military applications, where such a scheme could be of high practical relevance, e.g. for the simultaneous communication of multiple independent nodes to one destination.

VI. Acknowledgment

This work was partially funded by the Swiss federal office *armasuisse Science and Technology*.

References

[1] J. E. Kleider *et al.*, “Achieving MIMO performance with single-antenna radios,” in *MILCOM 2010*, Oct 2010, pp. 814–819.

[2] T. Rüegg, Y. Hassan, and A. Wittneben, “User cooperation enabled traffic offloading in urban hotspots,” in *IEEE PIMRC 2016*, September 2016.

[3] S. K. Jayaweera, “Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 984–989, May 2006.

[4] *Personal communications*. Swiss federal office armasuisse Science and Technology, March 2017.

[5] A. Sanderovich, S. Shamai, and Y. Steinberg, “Distributed MIMO receiver - achievable rates and upper bounds,” *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4419–4438, Oct 2009.

[6] R. Karasik, O. Simeone, and S. S. Shitz, “Robust uplink communications over fading channels with variable backhaul connectivity,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5788–5799, November 2013.

[7] L. Zhou and W. Yu, “Uplink multicell processing with limited backhaul via per-base-station successive interference cancellation,” *IEEE Journal on Sel. Areas in Com.*, vol. 31, no. 10, pp. 1981–1993, October 2013.

[8] Y. Zhou and W. Yu, “Optimized backhaul compression for uplink cloud radio access network,” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1295–1307, June 2014.

[9] Y. Zhou, Y. Xu, W. Yu, and J. Chen, “On the optimal fronthaul compression and decoding strategies for uplink cloud radio access networks,” *IEEE Trans. on Inf. Theory*, vol. 62, no. 12, pp. 7402–7418, Dec 2016.

[10] T. M. Cover and J. A. Thomas, *Elements of Information theory*, 2nd ed. Wiley Interscience, 2006.