Multi-way massive MIMO relay networks with maximum-ratio processing

https://doi.org/10.1109/SIGTELCOM.2017.7849808

Published in:
2017 International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom)

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© 2017 IEEE.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Multi-way Massive MIMO Relay Networks with Maximum-Ratio Processing

Chung Duc Ho*, Hien Quoc Ngo†, Michail Matthaiou*, and Trung Q. Duong*

*School of Electronics, Electrical Engineering and Computer Science, Queen’s University Belfast, BT7 1NN, Belfast, U.K.
†Department of Electrical Engineering (ISY), Linköping University, 581 83 Linköping, Sweden
Email: {choduc01, m.matthaiou, trung.q.duong}@qub.ac.uk, hien.ngo@liu.se

Abstract—This paper considers a multi-way massive multiple-input multiple-output (MIMO) relaying system. The bearing-information is exchanged among multiple users with the help of a multiple-antenna relay (the base station). The maximum-ratio (MR) processing is applied at the relay under the assumption of perfect channel state information. The spectral efficiency and the asymptotic results for the signal-to-interference-plus-noise ratio (when the number of relay antennas becomes large) are derived. By using a massive number of antennas, the transmit power at both user side and/or relay can be made inversely proportional to the number of relay antennas without degradation in the system performance.

Index Terms—Massive MIMO, maximum-ratio processing, multi-user MIMO, multi-way relay networks, zero-forcing processing.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) technology has attracted a lot of research interest recently for its ability to improve significantly the throughput, energy efficiency, and communication reliability of wireless systems [1]–[4]. In massive MIMO, each base station (BS) is equipped with very large antenna arrays and serves many tens of active users at the same time-frequency resource. Simple linear processing schemes such as zero-forcing (ZF), maximum-ratio (MR) or minimum mean-square error (MMSE), are used and nearly optimal. Furthermore, massive MIMO can be constructed with low-cost, low-power hardware components [5].

On a parallel avenue, during the past years, there has been a great deal of interest in multi-way relay networks [6], [7]. In multi-way relay networks, multiple users located in a large area exchange their information with the help of a relay. These systems will play an increasingly important role in future communication systems, since they provide robustness against channel variations, to transfer and to amplify the versions of the transmitted signal to all users without increasing system complexity [7]. Multi-way networks become very relevant when users are distributed within a large geographical area and cannot exchange information directly with each other. Thus, multi-way relay networks is a bridge to transfer (decode-and-forward, amplify-and-forward, and compress-and-forward) bearing-data among geographically separated users. For all these reasons, there is a plethora of applications of multi-way networks, including wireless conference, low-complexity data exchange between sensor nodes and data fusion centers in wireless sensor nodes, and power control in heterogeneous cellular networks.

Multi-way massive MIMO relay networks can reap all benefits achieved from both multi-way and massive MIMO technologies, and hence, they are expected to offer a huge spectral and energy efficiency. However, the majority of existing works in the literature considers multi-way systems with few antennas at the relay [7], [8] or multi-pair systems with massive relay antenna arrays [9], [10]. There is only little work on multi-way networks with massive antenna arrays at the relay node [11], [12]. Expressions for the spectral efficiency as well as its asymptotic result (in terms of the number of antennas at the relay) for multi-way massive MIMO systems with ZF processing at the relay are derived in [11]. Moreover, [12] investigates multi-way massive MIMO relay networks with simultaneous wireless information and power transfer techniques. Recall that the ZF scheme offers a good performance due to its ability for canceling out the interference. However, it involves a matrix inversion, and hence, has high computational complexity, especially when the numbers of antennas and users are large. By contrast, MR processing is very simple and is nearly optimal in the large-number-of-antenna regimes and can be implemented in a distributed manner [4]. To the best of the authors’ knowledge, there is no work on multi-way massive MIMO systems with MR processing in the literature.

In this paper, we consider a multi-way relaying system with massive antenna arrays at the relay and MR processing, and analyze the spectral efficiency. In particular, we investigate the potential for transmit power savings of the considered system. By using large antenna arrays, the transmit powers at users or/and relay can be reduced significantly (i.e. reduced \( M \) times, where \( M \) is the number of antennas), while maintaining a fixed quality of service. Furthermore, we compare the performance of MR and ZF processing. Numerical results show that MR performs as well as ZF when the number of relay antennas is large and the transmit power at relay or/and users is scaled down with substantial complexity reduction.

Notation: Boldface upper and lower case letters stand for matrices and vectors, respectively. The trace, transpose, and Hermitian operators are denoted by \( \text{Tr}(\cdot) \), \( (\cdot)^T \), and \( (\cdot)^H \), respectively. The \( (m,k) \)-th element of a matrix \( A \) is denoted by \( [A]_{m,k} \). Finally, we use \( \mathbb{E}\{\cdot\} \) and \( \|\cdot\| \) to denote the expectation and the Euclidean norm, respectively.
II. System Model

In this section, we present the system model of multi-way massive relay networks with massive antenna arrays. The system includes one relay equipped with $M$ antennas, and $K$ single-antenna users, where each of them operates in half-duplex mode. In particular, we focus on the scenario that both $M$ and $K$ are large, but we assume that $M$ is much larger than $K (1 \ll K \ll M)$. Each user wants to detect all $K-1$ signals transmitted from other users. We assume that the relay and users have perfect channel state information (CSI). The transmission is divided into two phases: multiple-access and broadcast phases.

A. Multiple-Access Phase

During this phase all users transmit their data to the relay in the same time-frequency resource. Let $s_k$, where $\mathbb{E}\{ |s_k|^2 \} = 1$, be the signal transmitted from the $k$-th user. Then, the signal vector received at the relay is given by

$$y_t = \sqrt{P_u}G s_t + n_t,$$

where $G \in \mathbb{C}^{M \times K}$ represents the channel matrix between the relay and the $K$ users, $s_t \triangleq [s_1, s_2, \ldots, s_K]^T$, $w_t \triangleq [w_1, w_2, \ldots, w_K]^T$ is the additive noise vector of independent and identically distributed (i.i.d.) Gaussian random variables (RVs) with zero-mean and unit variance, and $P_u$ is the transmit power of each user.

The channel coefficient between the $m$-th antenna of the relay and the $k$-th user is modeled as follows:

$$g_{mk} \triangleq \langle G \rangle_{mk} = h_{mk}\sqrt{\beta_k},$$

where $h_{mk}$ represents the small-scale fading and $\beta_k$ models the large-scale fading. We assume that $h_{mk} \sim \mathcal{CN}(0,1)$, and denote

$$G = HD^{1/2},$$

where $H$ is an $M \times K$ matrix of small-scale fading between the $K$ users and the relay with $[H]_{mk} = h_{mk}$, and $D$ is a $K \times K$ diagonal matrix, where $[D]_{kk} = \beta_k$.

After receiving the signal, the relay uses the MR combining scheme in the following manner:

$$\hat{y}_{t} = G^H y_{t}.$$  

B. Broadcast Phase

To transmit all signals to all users, the relay needs to spend $K-1$ time slots. At each time slot the relay uses the maximum-ratio precoding technique to broadcast a permuted version of $\hat{y}_t$. More precisely, the transmit signal vector at the relay for the $t$-th time slot is

$$x^{(t)}_t = \sqrt{\alpha^{(t)}}G \Pi^{(t)} \hat{y}_t,$$

where $\Pi^{(t)} \in \mathbb{C}^{K \times K}$ is the permutation matrix for the $t$-th time slot given by [7]

$$\Pi^{(t)} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

and $\alpha^{(t)}$ is the normalization factor, chosen to meet the power constrain at the relay

$$\mathbb{E}\{ \| x^{(t)}_t \|^2 \} = P_s.$$

From (1), (4), (5), and (7), we obtain (8) shown at the top of the next page. Then, the $K \times 1$ received signal vector at the $K$ users in the $t$-th time slot can be written as follows:

$$y^{(t)}_u = G^H x^{(t)}_t + n^{(t)}.$$

Substituting (1), (4), and (5) into (9), we have

$$y^{(t)}_u = \sqrt{\alpha^{(t)}}P_u G^H A^{(t)} s_t + \sqrt{\alpha^{(t)}}G^H B^{(t)} w_t + n^{(t)},$$

where

$$A^{(t)} = G \Pi^{(t)} G^H,$$

$$B^{(t)} = G \Pi^{(t)} G^H.$$

The received signal at the $k$-th user is written as

$$y^{(t)}_{uk} = \sqrt{\alpha^{(t)}}P_u \sum_{i=1}^{K} g_{ki} a^{(t)}_i s_i + \sqrt{\alpha^{(t)}}\sum_{m=1}^{M} b_{km}^H w_m + n^{(t)}_k,$$

where $g_{ki}, a^{(t)}_i$, and $b_{km}$ are the $k$-th columns of $G, A$ and $B$, respectively.

III. Spectral Efficiency Analysis and Asymptotic $(M \rightarrow \infty)$ Performance

In this section, we will focus on the spectral efficiency and asymptotic performance of the considered system. Without loss of generality, we consider the performance in the first time slot. The analysis will be the same for other time slots. In the first time slot, user $k$ wants to detect signal $s_{k+1}$ transmitted from user $k+1$. Note that, hereafter, we set $k+1$ as 1 if $k = K$. 

Fig. 1. Schematic illustration of a multi-way massive MIMO system.
\[ \alpha^{(t)} = \frac{P_t}{P_u \text{Tr} \left( G \Pi^{(t)} G^H (G \Pi^{(t)} G^H)^H \right) + \text{Tr} \left( G \Pi^{(t)} G^H (G \Pi^{(t)} G^H)^H \right)}. \] (8)

\[ M^2 \alpha^{(1)} = \frac{P_t}{E_u \text{Tr} \left( \frac{G^H G}{M} \right) \left( \Pi^{(1)} \right)^T \left( \frac{G^H G}{M} \right) \Pi^{(1)}} + \text{Tr} \left( \Pi^{(1)} \left( \frac{G^H G}{M} \right) \left( \Pi^{(1)} \right)^T \left( \frac{G^H G}{M} \right) \right) \] \[ \xrightarrow{a.s.} \frac{P_t}{E_u \text{Tr} \left( D^2 (\Pi^{(1)})^T D \Pi^{(1)} \right) + \text{Tr} \left( \Pi^{(1)} D (\Pi^{(1)})^T D \right)} \triangleq c_1, \quad M \to \infty. \] (18)

A. Spectral Efficiency Analysis

The received signal at first time slot for the \( k \)-th user is given by

\[ y_{u_k}(1) = \sqrt{\alpha^{(1)} P_u} \mathbf{g}_k^H \mathbf{a}_{k+1} s_{k+1} + \sqrt{\alpha^{(1)} P_u} \sum_{i=1}^{K} \mathbf{g}_i^H \mathbf{a}_i s_i + \sqrt{\alpha^{(1)} P_u} \sum_{m=1}^{M} \mathbf{g}_m^H \mathbf{b}_m w_m + n_k^{(1)}. \] (13)

Since user \( k \) knows its transmitted signal \( s_k \), it can subtract the self-interference term (which includes \( s_k \)) before detecting the desired signal \( s_{k+1} \) as follows:

\[ y_{u_k}^{(1)} = y_{u_k}^{(1)} - \sqrt{\alpha^{(1)} P_u} \mathbf{g}_k^H \mathbf{a}_{k+1} s_k \]
\[ = \sqrt{\alpha^{(1)} P_u} \mathbf{g}_k^H \mathbf{a}_{k+1} s_{k+1} + \sqrt{\alpha^{(1)} P_u} \sum_{i=1}^{K} \mathbf{g}_i^H \mathbf{a}_i s_i \]
\[ + \sqrt{\alpha^{(1)} P_u} \sum_{m=1}^{M} \mathbf{g}_m^H \mathbf{b}_m w_m + n_k^{(1)}. \] (14)

From (14), the signal-to-interference-plus-noise ratio (SINR) in the first time slot for the \( k \)-th user is derived as:

\[ \gamma_k^{(1)} = \frac{\alpha^{(1)} P_u \mathbf{g}_k^H \mathbf{a}_{k+1}^2}{\alpha^{(1)} P_u \sum_{i=1}^{K} \left| \mathbf{g}_i^H \mathbf{a}_i \right|^2 + \alpha^{(1)} M \sum_{m=1}^{M} \left| \mathbf{g}_m^H \mathbf{b}_m \right|^2 + 1}. \] (15)

Then, the spectral efficiency of the \( k \)-th user is

\[ \text{SE}_{k}^{(1)} = \frac{K - 1}{K} E \left\{ \log_2 \left( 1 + \gamma_k^{(1)} \right) \right\}. \] (16)

The pre-log factor \( \frac{K - 1}{K} \) accounts for the fact that we spend \( K \) time slots to transmit \( K - 1 \) signals to a given user.

B. Asymptotic \((M \to \infty)\) Performance

The elements of \( \mathbf{H} \) are i.i.d. RVs with zero mean and unit variance. Then, by using the results on very long random vectors [2], we have

\[ \frac{G^H G}{M} \xrightarrow{a.s.} D^{1/2} \frac{H^H H}{M} D^{1/2}, \quad M \to \infty \] (17a)
\[ \frac{G^H w}{\sqrt{M}} \xrightarrow{d} \tilde{w} \sim \mathcal{CN}(0, D), \quad M \to \infty. \] (17b)

where \( a.s. \) and \( d \) denote the almost sure convergence and the convergence in distribution, respectively.

1) Transmit power of each user is scaled with \( 1/M \): In this case, we assume that the transmit power of each user is scaled with \( 1/M \) according to \( P_u = E_u/M, \) and \( M \) increases without bound, where \( E_u \) is fixed. By substituting \( P_u = E_u/M \) into (8), and using the result (17a), we get (18) at the top of the page. From (10), we have

\[ \frac{y_{u}^{(1)}}{\sqrt{M}} = \sqrt{\alpha^{(1)} M^2 E_u \left( \frac{G^H G}{M} \right) \Pi^{(1)} \left( \frac{G^H G}{M} \right)} \mathbf{s} \]
\[ + \sqrt{\alpha^{(1)} M^2 \left( \frac{G^H G}{M} \right) \Pi^{(1)} \left( \frac{G^H w}{\sqrt{M}} \right)} + n_k^{(1)}. \] (19)

Using (18) together with (17a) and (17b), we obtain

\[ \frac{y_{u_k}^{(1)}}{\sqrt{M}} \to \sqrt{c_1 E_u D \Pi^{(1)} D} + \sqrt{c_1 D \Pi^{(1)} \tilde{w}} \to \infty, \quad M \to \infty, \] (20)

which leads to

\[ \frac{y_{u_k}^{(1)}}{\sqrt{M}} \to \sqrt{c_1 E_u \beta_k \beta_{k+1} \mathbf{s}_{k+1} + \sqrt{c_1 \beta_k \tilde{w}_{k+1}}} \to \infty. \] (21)

Therefore, the received SINR converges to \( \text{SINR}_k^{\infty} \) as \( M \to \infty \), given by

\[ \text{SINR}_k^{\infty} = \frac{E_u \beta_k^2 \beta_{k+1}^2}{\beta_k^2 \beta_{k+1}} = E_u \beta_k + 1. \] (22)

Result (22) implies that by using very large number of relay antennas \( M \), we can cut the transmit power at each user proportionally to \( 1/M \) without performance loss. Since \( P_t \) is fixed, when \( M \) goes to infinity, the interference and noise effects of the transmission link from the relay and user \( k \).
disappear. As a result, the asymptotic performance at user \( k \) depends only on the large scale fading between user \( k+1 \) and the relay.

2) Transmit power of the relay is scaled with \( 1/M \): In this case the transmit power of the relay is scaled with \( 1/M \) according to \( P_r = E_r/M \), and \( M \) grows without bound, where \( E_r \) is fixed.

Similarly, by using (3), and substituting \( P_r = E_r/M \) into (8), we have (23) at the top of the page. Thus,

\[
y_u^{(1)} = \sqrt{\alpha(1)M^4P_u\left(\frac{G^H G}{M}\right)\Pi(1)\left(\frac{G^H G}{M}\right)s} + \sqrt{\alpha(1)M^4\left(\frac{G^H G}{M}\right)\Pi(1)\left(\frac{G^H w}{M}\right)} + n^{(4)}
\]

\[
\rightarrow \sqrt{c_2P_uD\Pi(1)Ds + n^{(4)}}, \quad M \rightarrow \infty.
\] (24)

As a result, we obtain

\[
y_{u_k}^{(1)} = \sqrt{c_2P_u\beta_k\beta_{k+1}s_{k+1} + n_k^{(1)}}, \quad M \rightarrow \infty.
\] (25)

From (25), as \( M \rightarrow \infty \), the SINR converges to

\[
\text{SINR}_k^\infty = c_2P_u\beta_k\beta_{k+1}^2,
\] (26)

which implies that we can reduce the transmit power at the relay proportionally to \( 1/M \) when \( M \) is large.

3) Transmit power of each user and the relay are scaled with \( 1/M \): For this case, the transmit powers of each user and of the relay are scaled with \( 1/M \). More precisely \( P_u = E_u/M \) and \( P_r = E_r/M \), where both \( E_u \) and \( E_r \) are fixed regardless of \( M \).

Following the same methodology as in previous parts, we obtain (27) shown at the top of the page. Using this result and (17a)–(17b), we have

\[
y_u^{(1)} = \sqrt{\alpha(1)M^3E_u\left(\frac{G^H G}{M}\right)\Pi(1)\left(\frac{G^H G}{M}\right)s} + \sqrt{\alpha(1)M^3\left(\frac{G^H G}{M}\right)\Pi(1)\left(\frac{G^H w}{\sqrt{M}}\right)} + n^{(1)}
\]

\[
\rightarrow \sqrt{c_3E_uD\Pi(1)Ds} + \sqrt{c_3D\Pi(1)\hat{w} + n^{(1)}}.
\] (28)

The received signal at the \( k \)-th user is

\[
y_{u_k}^{(1)} = \sqrt{c_3E_u\beta_k\beta_{k+1}s_{k+1}} + \sqrt{c_3\beta_k\hat{w}_{k+1} + n_k^{(1)}}, \quad M \rightarrow \infty.
\] (29)

Then, the asymptotic SINR is given by

\[
\text{SINR}_k^\infty = \frac{c_3E_u\beta_k^2\beta_{k+1}}{c_3\beta_k^2\beta_{k+1} + 1}.
\] (30)

By deploying large antenna array at the relay, we can scale down the transmit power at both users and relay \( M \) times, without degradation of the system performance.

IV. NUMERICAL RESULTS

In this section, we provide numerical results to verify our analysis. Our performance metric is the sum spectral efficiency defined as

\[
\text{SE}_{\text{sum}} = \sum_{k=1}^{K} \text{SE}_k^{(1)}.
\] (31)

In all examples, we assume that \( \beta_k = 1 \). Furthermore we define \( \text{SNR} = E_u = E_r \).

Figure 2 shows the sum spectral efficiency as a function of SNR for ZF and MR processing with \( M = 200, K = 10 \). In this figure, case-1 corresponds to the case of \( P_u = E_u/M, P_r = E_r \), case-2 corresponds to the case of \( P_u = E_u/M, P_a = E_u \), and case-3 corresponds to the case of \( P_r = E_r/M, P_a = E_u/M \). It can be seen from the figure that the performances of MR and ZF are comparable. As anticipated, at low SNR, MR performs better than ZF. Note that, the implementation complexity of MR is much simpler than that of ZF, since ZF involves matrix inversion.

Next, we examine the asymptotic performance of MR processing. Figure 3 illustrates the sum spectral efficiency versus the number of relay antennas for the three cases as in Figure 2. Here, we choose \( \text{SNR} = 10 \text{ dB} \) and \( K = 10 \). As expected, for all cases, the sum spectral efficiency converges to its asymptotic value when \( M \) grows large. This verifies our asymptotic analysis in Section III-B.
Finally, we study the effect of $K$ on the sum spectral efficiency (see Fig. 4). We can see that, with $M = 100$, when the number of users increases, the sum spectral efficiency increases, reaches the maximum point, and then decreases. This is due to the fact that when $K$ increases, the multiplexing gain increases, but at the same time the system suffers more inter-user interference. With larger $M$, i.e., $M = 200$, the inter-user interference can be (nearly) canceled out, and hence, the sum spectral efficiency increases when $K$ increases from 10 to 100.

V. CONCLUSION

We have analyzed the performance of MR processing in multi-way relaying systems with massive antenna arrays at the relay. The asymptotic spectral efficiency when the number of antennas goes to infinity was derived. We have shown that by using massive MIMO technology we can reduce the transmit power at both user and/or relay proportionally to the inverse of the number of antennas. Furthermore, we showed that the performances of MR and ZF are comparable in the energy efficiency regime.

ACKNOWLEDGMENT

This work was supported by project no. 3811/QD-UBND, Binh Duong government, Vietnam. The work of H. Q. Ngo was supported by the Swedish Research Council (VR) and ELLIIT.

REFERENCES