2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)

A New Transceiver Architecture for Multi-User MIMO Communication Based on Mixture of Linear and Non-Linear Reception

Junyeong Seo, *Student Member, IEEE* and Youngchul Sung, *Senior Member, IEEE*

***Abstract*—In this paper, a new transceiver architecture for K-user multiple-input single-output (MISO) broadcast channels (BCs) based on linear and non-linear mixture reception is proposed as an alternative to conventional fully linear zero-forcing (ZF) downlink beamforming. In the new transceiver architecture, two closely-aligned users are paired as a group, and superposition coding and non-linear successive cancellation (SIC) reception based on Pareto-optimal design is applied to each closely-aligned two-user group, while ZF beamforming is maintained across roughly-orthogonal groups. Numerical results show that the proposed new architecture yields non-trivial gain over conventional full ZF beamforming by mitigating the performance degradation of full ZF beamforming caused by closely-aligned channel vectors.**

***Index Terms*—multi-user MIMO, Transceiver, Pareto-optimality, SIC**

I. Introduction

Multi-user multiple-input multiple-output (MU-MIMO) is one of the key technologies for current and future wireless systems [2]. In MU-MIMO, precisely speaking, MU-MISO, the base station (BS) equipped with multiple transmit antennas transmits data to spatially-separated multiple users each with a single receive antenna. Thus, even in line-of-sight (LoS) radio propagation environments, MU-MISO can provide multiplexing gains for high data rates. However, the capacity of a MU-MISO broadcast channel (BC) is achieved by highly non-linear transmitter processing of dirty-paper coding (DPC) [3], whereas the capacity of single-user (SU) MIMO can simply be achieved by linear eigen-beamforming [4]. To circumvent the burden of DPC, researchers resorted to MU diversity and user scheduling to yield good performance only with linear downlink beamforming for the past decade [5], [6]. That is, when the number of users in the cell is sufficiently large for a given (relatively small) number N_t of transmit antennas, the BS can choose N_t users with nearly-orthogonal channel vectors so that linear ZF or minimum mean-square error (MMSE) downlink beamforming is sufficient due to the near orthogonality among the scheduled users' channels. Unfortunately, such nearly-orthogonal user selection is very difficult and user scheduling based on linear downlink beamforming is ineffective under rich scattering environments when the number of transmit antenna is large [7], [8]. Thus, in MU-MISO with a large number of transmit antennas it was proposed that the BS selects the served users arbitrarily and uses linear ZF beamforming [7]. However, in such cases, the channel vectors of some of the selected users would be closely aligned so that the rough orthogonality among selected users' channels is destroyed and the performance of ZF beamforming is degraded.

In this paper, hinted by the fact that non-linear processing is required for optimal performance for MU-MISO BCs and inspired by the usefulness of superposition and SIC decoding in non-orthogonal multiple access [9] and rate splitting for limited feedback MU-MIMO [10], we propose a new transceiver architecture for MU-MISO BCs to enhance the performance of fully linear downlink beamforming, based on mixture of linear and non-linear reception. The basic idea of the proposed architecture is as follows: For a given set of channel vectors in a MU-MISO BC, if the channel vectors of some users are closely aligned, the performance of the closely-aligned users is significantly degraded with linear ZF or MMSE beamforming. However, if we group the closely-aligned users and apply *superposition coding and non-linear SIC reception* for the closely-aligned user group while ZF beamforming is maintained across roughly-orthogonal other users, the performance degradation by fully ZF beamforming can be alleviated. In this paper, as an initial step towards this research direction, we consider grouping of two users and propose a relevant design procedure for this new architecture.

II. System Model

We consider a K-user Gaussian MU-MISO BC with a BS equipped with N_t transmit antennas and K single-antenna users. We group the K users into N_g groups, where each group consists of one or two users. So, we have $N_g \in \{K/2 \text{ (maximum paring)}, \cdots, K \text{ (no paring)}\}$. (User grouping will be discussed in Section IV.) After user grouping, we assume ZF beamforming across groups to control inter-group interference and assume superposition coding and SIC decoding for each two-user group. Under these assumptions, the transmit signal $\mathbf{x}$ of the BS can be expressed as

$$\mathbf{x} = \sum_{j=1}^{N_g} \mathbf{\Pi}^{(j)} \sum_{i\in\mathcal{G}_j} \sqrt{p_i^{(j)}}\mathbf{w}_i^{(j)} s_i^{(j)}, \quad |\mathcal{G}_j| = 1 \text{ or } 2, \qquad (1)$$

The authors are with the Dept. of Electrical Engineering, KAIST, Daejeon 305-701, South Korea. E-mail:jyseo@kaist.ac.kr, ysung@ee.kaist.ac.kr. This work was supported in part by 'The Cross-Ministry Giga KOREA Project' of The Ministry of Science, ICT and Future Planning, Korea [GK14N0100, 5G mobile communication system development based on mmWave]. Some part of this paper is from [1].

where $s_i^{(j)}$ is the transmit symbol from zero-mean unit-variance Gaussian distribution $\mathcal{CN}(0,1)$ for user i in group $\mathcal{G}_j$, $\mathbf{w}_i^{(j)}$ is the $N_t \times 1$ beamforming vector for user i in group $\mathcal{G}_j$ out of the feasible set $\mathcal{W} := \{\mathbf{w} \mid \|\mathbf{w}\|^2 \le 1\}$, $p_i^{(j)}$ is the power assigned to user i in group $\mathcal{G}_j$, and $\mathbf{\Pi}^{(j)}$ is the inter-group $N_t \times N_t$ ZF projection matrix for group $\mathcal{G}_j$. The total BS transmit power P_t is divided as $2P_t/K$ for each group with two users and P_t/K for each group with one user. The received signal at user k is given by the inner product of user k's channel vector $\mathbf{g}_k$ and the BS transmit signal vector $\mathbf{x}$ corrupted by zero-mean additive white Gaussian noise (AWGN).

III. Two-user MISO BC with SIC: Pareto-optimal design

In the proposed architecture, two closely-aligned users are served with superposition coding and SIC reception to avoid the penalty of fully linear ZF beamforming. Hence, we first investigate the optimal design and corresponding performance of a two-user MISO BC with SIC, to provide a comparison basis with full ZF beamforming performance in the next section. Our optimal design criterion is Pareto-optimality widely used in MU-MISO beam design.

With the group index omitted, the two-user model can be written as

$$y_i = \mathbf{h}_i^H(\sqrt{p_i}\mathbf{w}_i s_i + \sqrt{p_j}\mathbf{w}_j s_j) + n_i \quad i,j \in \{1,2\}, \;\; j \neq i,$$

where y_i is the received signal of user i, $\mathbf{h}_i$ is the *effective* channel vector of user i incorporating the $N_t \times N_t$ ZF projection matrix (i.e., $\mathbf{h}_i = \mathbf{\Pi}^{(j)}\mathbf{g}_i$), $n_i \sim \mathcal{CN}(0,\sigma_i^2)$ is AWGN, and $p_1 + p_2 \le P = 2P_t/K$ with P as the total power allocated to the group. We assume that user 1 has better channel condition than user 2, i.e., $\|\mathbf{h}_1\|^2/\sigma_1^2 > \|\mathbf{h}_2\|^2/\sigma_2^2$ and that user 1 decodes the message of user 2 and subtracts it by SIC before decoding its own data while user 2 treats the interference as noise. With this assumption, the rates of the two users are given by

$$\begin{aligned} R_1(\mathbf{w}_1,p_1) &= \log_2\left(1+\frac{s_1(\mathbf{w}_1,p_1)}{\sigma_1^2}\right), \\ R_2(\mathbf{w}_1,\mathbf{w}_2,p_1,p_2) & \\ = \log_2\Bigg(1+\min&\left\{\frac{r_1(\mathbf{w}_2,p_2)}{s_1(\mathbf{w}_1,p_1)+\sigma_1^2},\frac{s_2(\mathbf{w}_2,p_2)}{r_2(\mathbf{w}_1,p_1)+\sigma_2^2}\right\}\Bigg), \end{aligned} \tag{2}$$

where the signal and interference powers are given by

$$s_i(\mathbf{w}_i,p_i) := p_i|\mathbf{h}_i^H\mathbf{w}_i|^2 \quad \text{and} \quad r_i(\mathbf{w}_j,p_j) := p_j|\mathbf{h}_i^H\mathbf{w}_j|^2. \tag{3}$$

Note in (2) that for the rate of user 1 the interference from user 2 is not incorporated due to SIC and the rate of user 2 is determined by not only the signal-to-interference-plus-noise ratio (SINR) of user 2 at user 2 but also the required 'SINR' for user 1 to decode the message of user 2 for SIC. Then, for given channel vectors $(\mathbf{h}_1,\mathbf{h}_2)$, the achievable rate region $\mathcal{R}$ of the two-user MISO BC with SIC is defined as the union of all the rate-tuples that can be achieved by feasible beam vectors and power allocation:

$$\mathcal{R} := \bigcup_{\substack{(\mathbf{w}_1,\mathbf{w}_2)\in\mathcal{W}^2 \\ p_1,p_2:\; p_1,p_2\ge 0, \\ p_1+p_2=P}} (R_1(\mathbf{w}_1,p_1), R_2(\mathbf{w}_1,\mathbf{w}_2,p_1,p_2)). \tag{4}$$

The Pareto boundary of $\mathcal{R}$ is the outer boundary of $\mathcal{R}$ for which the rate of any one user cannot be increased without sacrificing the rate of the other user and Pareto-optimality has been used widely as an optimality criterion for MU-MISO beam design because the rate operating point can be set optimally [11], [12]. It is known that the Pareto-boundary can be obtained by solving the following optimization problem with feasible R_1^* swept [11], [12]:

$$\begin{aligned} \max_{\substack{(\mathbf{w}_1,\mathbf{w}_2)\in\mathcal{W}^2 \\ p_1,p_2:\; p_1,p_2\ge 0,\; p_1+p_2=P}} \quad & R_2(\mathbf{w}_1,\mathbf{w}_2,p_1,p_2) \\ \text{subject to} \quad & R_1(\mathbf{w}_1,p_i) = R_1^*. \end{aligned} \tag{5}$$

The problem (5) can be rewritten by using (2) in terms of SINR as

$$\begin{aligned} \max_{\substack{(\mathbf{w}_1,\mathbf{w}_2)\in\mathcal{W}^2 \\ p_1,p_2:\; p_1,p_2\ge 0, \\ p_1+p_2=P}} \gamma_2 &:= \min\left\{\frac{r_1(\mathbf{w}_2,p_2)}{s_1(\mathbf{w}_1,p_1)+\sigma_1^2},\frac{s_2(\mathbf{w}_2,p_2)}{r_2(\mathbf{w}_1,p_1)+\sigma_2^2}\right\} \\ \text{subject to} \quad & s_1(\mathbf{w}_1,p_1)/\sigma_1^2 = \gamma_1^*, \end{aligned} \tag{6}$$

where γ_1^* is a given feasible target SINR for user 1. The problem (6) can be solved by an efficient parameterization of the beam vectors $\mathbf{w}_1$ and $\mathbf{w}_2$. It is known that the Pareto-optimal beam vectors for the problem (6) can be parameterized as [13]

$$\begin{aligned} \mathbf{w}_1(\alpha_1,\beta_1) &= \alpha_1\frac{\Pi_{\mathbf{h}_2}\mathbf{h}_1}{\|\Pi_{\mathbf{h}_2}\mathbf{h}_1\|} + \beta_1\frac{\Pi_{\mathbf{h}_2}^{\perp}\mathbf{h}_1}{\|\Pi_{\mathbf{h}_2}^{\perp}\mathbf{h}_1\|}, \\ \mathbf{w}_2(\alpha_2) &= \alpha_2\frac{\Pi_{\mathbf{h}_1}\mathbf{h}_2}{\|\Pi_{\mathbf{h}_1}\mathbf{h}_2\|} + \sqrt{1-\alpha_2^2}\frac{\Pi_{\mathbf{h}_1}^{\perp}\mathbf{h}_2}{\|\Pi_{\mathbf{h}_1}^{\perp}\mathbf{h}_2\|}, \end{aligned} \tag{7}$$

where $(\alpha_1,\beta_1) \in \mathcal{F} := \{(\alpha,\beta),\alpha,\beta \ge 0, \alpha^2+\beta^2 \le 1\}$ and $\alpha_2 \in [0,1]$. Unlike the conventional parametrization without SIC in which both users use full power, the parameterization (7) shows that user 1 may not use full power whereas user 2 uses full power. Substituting (7) into (3), substituting the resulting $s_i(\mathbf{w}_i,p_i)$ and $r_i(\mathbf{w}_j,p_j)$ into the problem (6) and taking square-root operation, we can reexpress (6) as (10) and (11) shown at the top of the next page. Here, the parameter θ regarding the angle between two channel vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ is defined as $\theta := \frac{|\mathbf{h}_1^H\mathbf{h}_2|^2}{\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2} \in [0,\ 1]$. We further define the following quantities related to the channel magnitude and the target SINR for user 1:

$$\lambda_i := ||\mathbf{h}_i||^2/\sigma_i^2, \; i=1,2, \tag{8}$$

$$\Gamma := \gamma_1^*/\lambda_1. \tag{9}$$

Note that feasible range for Γ is $\Gamma \in [0,P]$, where the maximum value P occurs when $\mathbf{w}_1 = \mathbf{h}_1/||\mathbf{h}_1||$ and $p_1 = P$ since $\gamma_1 = s_1(\mathbf{w}_1,p_1)/\sigma_1^2 = p_1|\mathbf{h}_1^H\mathbf{w}_1|^2/\sigma_1^2$.

$$\max_{\substack{(\alpha_1,\beta_1)\in\mathcal{F}\\ \alpha_2\in[0,1]\\ 0\le p_1\le P}} \quad \gamma_2 = \min\left\{\frac{\sqrt{P-p_1}\|\mathbf{h}_1\|\alpha_2}{\sqrt{\sigma_1^2(1+\gamma_1^*)}}, \frac{\sqrt{P-p_1}\|\mathbf{h}_2\|(\sqrt{\theta}\alpha_2+\sqrt{1-\theta}\sqrt{1-\alpha_2^2})}{\sqrt{p_1\|\mathbf{h}_2\|^2\alpha_1^2+\sigma_2^2}}\right\} \tag{10}$$

$$\text{subject to} \quad \sqrt{p_1}\|\mathbf{h}_1\|(\sqrt{\theta}\alpha_1+\sqrt{1-\theta}\beta_1) = \sqrt{\gamma_1^*\sigma_1^2}. \tag{11}$$

A. Pareto-optimal beam design and power allocation

The optimization problem (10) - (11) was considered in [13] with $p_1 = p_2 = 1$ under the framework of a two-user MISO interference channel. In the MISO interference channel case, two transmitters neither cooperate nor share transmit power and hence the two transmit power values p_1 and p_2 are fixed. On the other hand, p_1 and p_2 are design variables under the constraint $p_1 + p_2 = P$ in the considered MISO-BC case. Based on the result from [13], we here solve the joint power allocation and beam design problem (10) - (11). To solve the problem (10) - (11), for given target SINR γ_1^* for user 1 we first express the maximum SINR γ_2 for user 2 in terms of p_1 for given p_1 ($p_2 = P - p_1$) by using the result from [13] and then optimize p_1 for maximum γ_2.

First, based on the results in [13], the optimal solution to (10) - (11) can be expressed as a function of p_1 [13]: $\gamma_2^*(p_1) =$

$$\begin{cases} \gamma_2^{*(1)}(p_1) = (P-p_1)\frac{\|\mathbf{h}_1\|^2}{\sigma_1^2(1+\gamma_1^*)} & \text{if } p_1 \in \mathcal{P}_1,\\ \gamma_2^{*(2)}(p_1) = (P-p_1)\frac{\|\mathbf{h}_1\|^2}{\sigma_1^2(1+\gamma_1^*)}[\alpha_2^*(p_1)]^2 & \text{if } p_1 \in \mathcal{P}_2,\\ \gamma_2^{*(3)}(p_1) = (P-p_1)\frac{\|\mathbf{h}_2\|^2}{\|\mathbf{h}_2\|^2 p_1[\alpha_1^*(p_1)]^2+\sigma_2^2} & \text{if } p_1 \in \mathcal{P}_3,\end{cases} \tag{12}$$

where

$$\alpha_2^*(p_1) = \begin{cases} 1 & \text{if } p_1 \in \mathcal{P}_1,\\ \frac{c(p_1)}{\sqrt{c^2(p_1)+[a(p_1)-b(p_1)]^2}} & \text{if } p_1 \in \mathcal{P}_2,\\ \frac{b(p_1)}{\sqrt{b^2(p_1)+c^2(p_1)}} = \sqrt{\theta} & \text{if } p_1 \in \mathcal{P}_3,\end{cases} \tag{13}$$

and $\mathcal{P}_1 := \{p_1 | a(p_1) \le b(p_1)\}$, $\mathcal{P}_2 := \{p_1 | b(p_1) < a(p_1) \le b(p_1) + c^2(p_1)/b(p_1)\}$, and $\mathcal{P}_3 := \{p_1 | a(p_1) > b(p_1) + c^2(p_1)/b(p_1)\}$.

Then, the original problem (10) - (11) reduces to

$$\max_{0\le p_1\le P} \gamma_2^*(p_1) \tag{14}$$

where $\gamma_2^*(p_1)$ is given by (12). One way to solve the problem (14) is that we sweep p_1 from 0 to P, determine which $\mathcal{P}_i$ each p_1 belongs to, compute and store the corresponding value $\gamma_2^*(p_1)$, and select p_1 that yields maximum $\gamma_2^*(p_1)$ after the sweeping is done. However, a more efficient solution can be found by the following proposition:

Proposition 1: For given $\mathbf{h}_1$, $\mathbf{h}_2$, σ_1^2, σ_2^2, P and γ_1^*, the set $\mathcal{P}_i$ to which the optimal solution p_1^{opt} to the problem (14) belongs can be identified *a priori*. If $\theta\Gamma < \tau$ or if $\theta\Gamma \ge \tau \ge 0$ and $P \ge \Gamma + \frac{1}{1-\theta}(\sqrt{\theta\Gamma} - \sqrt{\tau})\left(\sqrt{\theta\Gamma} + \frac{1}{\lambda_2\sqrt{\tau}}\right)$, then $p_1^{opt} \in \mathcal{P}_2$. Otherwise, $p_1^{opt} \in \mathcal{P}_3$. Here, $\tau := \theta^{-1}\left(\lambda_1^{-1} + \Gamma\right) - \lambda_2^{-1}$, and Γ and λ_i are defined in (8).

Proof: Proof is available in [1]. ■

TABLE I

Pareto-optimal design for 2-user MISO-BC with SIC: $[\sqrt{p_1}\mathbf{w}_1, \sqrt{p_2}\mathbf{w}_2] = \mathcal{D}(\mathbf{h}_1, \mathbf{h}_2, \sigma_1^2, \sigma_2^2, \gamma_1^*, P)$
Input: channel vectors $\mathbf{h}_1, \mathbf{h}_2$, noise power σ_1^2, σ_2^2, target SINR of user 1 γ_1^*, and total power P.
Initialization: $\lambda_1 = \|\mathbf{h}_1\|^2/\sigma_1^2$, $\lambda_2 = \|\mathbf{h}_2\|^2/\sigma_2^2$, $\theta = \frac{\|\mathbf{h}_1^H\mathbf{h}_2\|^2}{\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2}$, $\Gamma = \gamma_1^*/\lambda_1$, and $\tau = \theta^{-1}\left(\lambda_1^{-1}+\Gamma\right) - \lambda_2^{-1}$
if $\theta\Gamma < \tau$
obtain p_1^{opt} maximizing $\gamma_2^{*(2)}$
elseif $\tau \ge 0$ and $P \ge \Gamma + \frac{1}{1-\theta}(\sqrt{\theta\Gamma}-\sqrt{\tau})\left(\sqrt{\theta\Gamma}+\frac{1}{\lambda_2\sqrt{\tau}}\right)$
obtain p_1^{opt} maximizing $\gamma_2^{*(2)}$
else
obtain p_1^{opt} maximizing $\gamma_2^{*(3)}$
endif
Obtain [13] $\alpha_1^*(p_1) = \begin{cases} 0 & \text{if } \Gamma \le p_1(1-\theta),\\ \sqrt{\theta\Gamma/p_1} - \sqrt{(1-\theta)(1-\Gamma/p_1)} & \text{if } \Gamma > p_1(1-\theta).\end{cases}$
and obtain β_1^* and α_2^* using (11) and (13) with $p_1 = p_1^{opt}$, and obtain $\mathbf{w}_1$ and $\mathbf{w}_2$ from (7) with α_1^*, β_1^* and α_2^*.
Output: $\sqrt{p_1^{opt}}\mathbf{w}_1$ and $\sqrt{P-p_1^{opt}}\mathbf{w}_2$

Due to Proposition 1 we know which of the three cases in (12) is applicable to the given combination of $\mathbf{h}_1$, $\mathbf{h}_2$, σ_1^2, σ_2^2, P, and γ_1^*. Once the set $\mathcal{P}_i$ to which p_1^{opt} belongs is determined, optimal p_1^{opt} can be found by maximizing the corresponding $\gamma_2^{*(i)}(p_1)$ in (12) with respect to p_1. The proposed algorithm for optimal power allocation and beam design for two-user MISO BCs with SIC is summarized in Table I.

IV. User Grouping and Overall Design

Now let us consider user grouping and overall design for the linear and SIC mixture reception architecture. For given two users with channel vectors $\mathbf{h}_1$ and $\mathbf{h}_2$, if ZF beamforming with equal power allocation is used, the rates of the two users are given by

$$R_1^{ZF} = \log_2\left(1 + \frac{P}{2}\lambda_1(1-\theta)\right),\ R_2^{ZF} = \log_2\left(1 + \frac{P}{2}\lambda_2(1-\theta)\right),$$

where $\lambda_1 = \|\mathbf{h}_1\|^2/\sigma_1^2, \lambda_2 = \|\mathbf{h}_2\|^2/\sigma_2^2$, $\theta = \frac{|\mathbf{h}_1^H\mathbf{h}_2|^2}{\|\mathbf{h}_1\|^2\|\mathbf{h}_1\|^2}$ and P is the total power for the two users. Note that the sum rate significantly decreases as θ increases, i.e., the two channel vectors become aligned. On the other hand, if we pair the two users and apply the Pareto-optimal beam design and SIC decoding described in Section III, a low bound on the corresponding sum rate is given in the following proposition:

Proposition 2: Under the simple power allocation strategy that assigns minimum power $p_{1,min}(= \gamma_1^*/\lambda_1 = \Gamma)$ to achieve the target SINR γ_1^* to user 1 and the rest of power P to user

2, for sufficiently large $\lambda_1\Gamma$ and $\lambda_2\Gamma$ with $\theta \in (0,1]$, the sum rate is lower bounded by

$$R_1^{SIC} + R_2^{SIC} \geq \log_2(1+\lambda_1 P) - \epsilon, \tag{15}$$

where $\epsilon > 0$ is arbitrarily small.

Proof: For the simple power allocation strategy, from (12), the SINR γ_2^* of user 2 can be obtained as

$$\gamma_2^* = \begin{cases} \frac{P-\Gamma}{\Gamma}\frac{1}{\lambda_1^{-1}\Gamma^{-1}+1}\left[1+\frac{\theta}{1-\theta}\left(\sqrt{\frac{1+\lambda_2^{-1}\Gamma^{-1}\theta^{-1}}{1+\lambda_1^{-1}\Gamma^{-1}}}-1\right)^2\right]^{-1} & \text{if } \theta \leq \theta_1 \\ \frac{P-\Gamma}{\Gamma}\frac{1}{\theta+\lambda_2^{-1}\Gamma^{-1}} & \text{if } \theta > \theta_1, \end{cases} \tag{16}$$

where $\theta_1 := \frac{1}{2}\left[-\lambda_2^{-1}\Gamma^{-1} + \sqrt{\lambda_2^{-2}\Gamma^{-2} + 4(\lambda_1^{-1}\Gamma^{-1}+1)}\right]$. For sufficiently large $\lambda_1\Gamma$ and $\lambda_2\Gamma$, θ_1 approaches to one, and hence the condition $\theta \leq \theta_1$ is satisfied and γ_2^* is given by the upper formula in (16). Then, it can be shown that, as $\lambda_2\Gamma, \lambda_1\Gamma \to \infty$, γ_2^* converges to

$$\lim_{\substack{\lambda_1\Gamma\to\infty\\ \lambda_2\Gamma\to\infty}} \frac{P-\Gamma}{\Gamma}\frac{1}{\lambda_1^{-1}\Gamma^{-1}+1}\left[1+\frac{\theta}{1-\theta}\left(\sqrt{\frac{1+\lambda_2^{-1}\Gamma^{-1}\theta^{-1}}{1+\lambda_2^{-1}\Gamma^{-1}}}-1\right)^2\right]^{-1} = \frac{P-\Gamma}{\Gamma}.$$

From the fact that $\gamma_1^* = \lambda_1\Gamma$, we have for sufficiently large $\lambda_1\Gamma$ and $\lambda_2\Gamma$

$$\begin{aligned} R_1 + R_2 &\geq \log_2(1+\lambda_1\Gamma) + \log_2\left(1+\frac{P-\Gamma}{\Gamma}\right) - \epsilon \\ &= \log_2[(1+\lambda_1\Gamma)\frac{P}{\Gamma}] \overset{(a)}{\geq} \log_2[(\frac{\Gamma}{P}+\lambda_1\Gamma)\frac{P}{\Gamma}] \\ &= \log_2(1+\lambda_1 P) - \epsilon, \end{aligned}$$

where (a) is from the fact $\Gamma/P \leq 1$ by the definition of Γ. ■

Since the optimal power allocation obtained in Section III outperforms the simple power allocation considered in Proposition 2, the sum rate in Proposition 2 is a lower bound on the optimal sum rate. Hence, a sufficient condition for $R_1^{SIC} + R_2^{SIC} \geq R_1^{ZF} + R_2^{ZF}$ is obtained as

$$\theta \geq 1 - \frac{1}{P}\left(-\left(\lambda_1^{-1}+\lambda_2^{-1}\right) + \sqrt{\left(\lambda_1^{-1}+\lambda_2^{-1}\right)^2 + \lambda_2^{-1}P}\right),$$

where we denote the RHS of the above equation by θ_τ. Thus, if the two users' channels with sufficiently large $\lambda_1\Gamma$ and $\lambda_2\Gamma$ have θ larger than θ_τ, the Pareto-optimally designed superposition coding and SIC decoding outperforms the ZF beamforming. Therefore, by smartly pairing two closely aligned users as a group and applying Pareto-optimal beam design and SIC to each two-user group, while applying ZF beamforming across groups, we can enhance the system performance over full ZF beamforming.

The proposed overall design procedure for the K-user MISO BC based on linear and non-linear SIC mixture reception is presented in the below:

Step 0) Given information: K users' $N_t \times 1$ actual channel vectors $\mathbf{g}_1, \cdots, \mathbf{g}_K$ and thermal noise variances $\sigma_1^2, \cdots, \sigma_K^2$, total power P_t shared by all K users

Step 1) Initialization: $\mathcal{U} = \{1,\ldots,K\}$, $\mathbf{G}_{\mathcal{U}} = [\mathbf{g}_1, \cdots, \mathbf{g}_K]$, $i_{count} = 1$

Step 2) Select user κ_1 as $\kappa_1 = \arg\max_{u\in\mathcal{U}} \|\mathbf{g}_u\|^2$.

Step 3) Construct the semi-aligned user set $\mathcal{A}$ for κ_1 as follows: For each $u \in \mathcal{U}\setminus\{\kappa_1\}$, compute the angle between the effective channels of user κ_1 and user u, and the angle threshold as follows:

$$\mathbf{h}_{\kappa_1|u} = \mathbf{\Pi}^{\perp}_{\mathbf{G}_{\mathcal{U}\setminus\{\kappa_1,u\}}}\mathbf{g}_{\kappa_1}, \quad \mathbf{h}_{u|\kappa_1} = \mathbf{\Pi}^{\perp}_{\mathbf{G}_{\mathcal{U}\setminus\{\kappa_1,u\}}}\mathbf{g}_u,$$

$$\theta_u = \frac{|\mathbf{h}_{\kappa_1|u}^H\mathbf{h}_{u|\kappa_1}|^2}{\|\mathbf{h}_{\kappa_1|u}\|^2\|\mathbf{h}_{u|\kappa_1}\|^2}, \quad \text{and} \quad \theta_{\tau,u} =$$

$$1 - \frac{1}{P}\left(-\left(\lambda_{\kappa_1|u}^{-1}+\lambda_{u|\kappa_1}^{-1}\right) + \sqrt{\left(\lambda_{\kappa_1|u}^{-1}+\lambda_{u|\kappa_1}^{-1}\right)^2 + \lambda_{u|\kappa_1}^{-1}P}\right)$$

where $\mathbf{\Pi}^{\perp}_{\mathbf{A}}$ is the projection matrix onto the orthogonal complement of the column space of matrix $\mathbf{A}$, $\mathbf{G}_{\mathcal{U}\setminus\{\kappa_1,u\}}$ is the matrix composed the columns of $\mathbf{G}_{\mathcal{U}}$ except the columns corresponding users κ_1 and u, $\lambda_{\kappa_1|u} = \|\mathbf{h}_{\kappa_1|u}\|^2/\sigma_{\kappa_1}^2$, $\lambda_{u|\kappa_1} = \|\mathbf{h}_{u|\kappa_1}\|^2/\sigma_u^2$, and $P = 2P_t/K$. Then, set $\mathcal{A} = \{u \in \mathcal{U}\setminus\{\kappa_1\} \mid \theta_u \geq \theta_{\tau,u}\}$.

Step 5) Decide whether to pair a user with κ_1 or not, and design the beam vector(s) correspondingly:

If $\mathcal{A} = \emptyset$, then $\mathcal{G}_{i_{count}} = \{\kappa_1\}$, $\mathcal{U} \leftarrow \mathcal{U}\setminus\{\kappa_1\}$, and design the beam vector for user κ_1 in group $\mathcal{G}_{i_{count}}$ as

$$\mathbf{w}_1^{\mathcal{G}_{i_{count}}} = \sqrt{P/2}\mathbf{\Pi}^{\perp}_{\mathbf{G}_{\mathcal{U}\setminus\{\kappa_1\}}}\mathbf{g}_{\kappa_1}/||\mathbf{\Pi}^{\perp}_{\mathbf{G}_{\mathcal{U}\setminus\{\kappa_1\}}}\mathbf{g}_{\kappa_1}||. \tag{17}$$

else

$$\kappa_2 = \arg\min_{u\in\mathcal{A}} \|\mathbf{h}_{u|\kappa_1}\|, \tag{18}$$

$$\mathcal{G}_{i_{count}} = \{\kappa_1,\kappa_2\},\ \mathcal{U} \leftarrow \mathcal{U}\setminus\{\kappa_1,\kappa_2\},\ \text{and} \tag{19}$$

design the beam vectors $\mathbf{w}_1^{\mathcal{G}_{i_{count}}}$ and $\mathbf{w}_2^{\mathcal{G}_{i_{count}}}$ for users κ_1 and κ_2 in group $\mathcal{G}_{i_{count}}$ by applying the Pareto-optimal beam design algorithm in Table I with the projected effective channel vectors $\mathbf{\Pi}^{\perp}_{\mathbf{G}_{\mathcal{U}\setminus\{\kappa_1,\kappa_2\}}}\mathbf{g}_{\kappa_1}$ and $\mathbf{\Pi}^{\perp}_{\mathbf{G}_{\mathcal{U}\setminus\{\kappa_1,\kappa_2\}}}\mathbf{g}_{\kappa_2}$ as the two algorithm input channel vectors and with power P.

Step 6) If $\mathcal{U} \neq \emptyset$, $i_{count} = i_{count}+1$ and repeat steps 2) to 5). If $\mathcal{U} = \emptyset$, set the number N_g of the designed groups as $N_g = i_{count}$ and stop.

In the proposed scheme, the number of users in each group $\mathcal{G}_j$ (which can be either one or two) and the number N_g of groups are automatically determined, and $\mathbf{\Pi}^{(j)}$, $p_i^{(j)}$ and $\mathbf{w}_i^{(j)}$ in (1) are fully determined. In the proposed mixture method, if a user's channel is sufficiently misaligned from all other users' channels so that there is no gain in pairing for SIC, the user is served alone by the ZF beam vector with power P_t/K. On the other hand, if a user's channel has a sufficiently aligned another user's channel, the two users are paired and the two-user Pareto-optimal superposition beam design and SIC are applied to the two users. In the pairing, the user with minimum channel norm is selected in (18). This is because if a user with weak channel is unpaired in the end, the user will be served alone by a ZF beam vector with power P_t/K and the resulting rate is small since the channel norm is small. Hence, by pairing such a user with a compatible

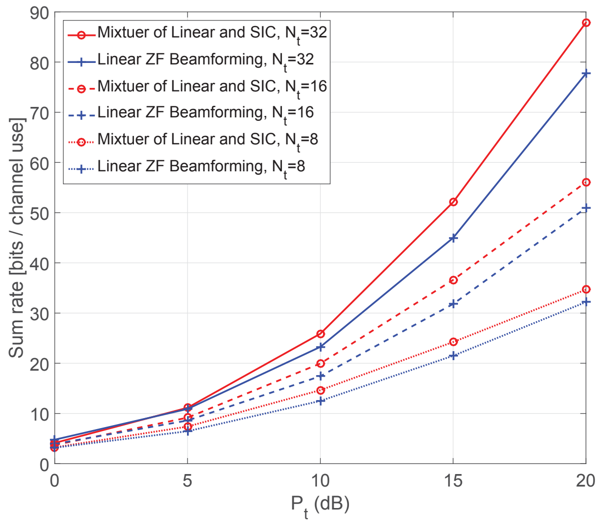

Fig. 1. Sum rate: The proposed method versus the full ZF method

strong user and applying two-user SIC decoding, both of the paired users can be benefitted from pairing. Note that in the proposed mixture scheme, ZF is applied across the designed groups $\mathcal{G}_1, \cdots, \mathcal{G}_{N_g}$ and the performance degradation of the full ZF beamforming is mitigated by proper pairing of two semi-aligned users and applying superposition coding and SIC to the constructed pairs.

A. Feedback and Complexity

To implement the proposed mixture scheme, the channel vectors of K users should be fed back to the BS as in full ZF downlink beamforming. Furthermore, the strong user in each two-user group should be notified to perform SIC. This notification can be done by some downlink control channel parallel to downlink data channel. Note that the weak user in each two-user group treats the interference as noise and thus the weak user need no additional information. Furthermore, there is no inter-group interference since inter-group interference is taken care of by downlink ZF beamforming across groups. Thus, the only additional feature of the proposed scheme is notification to the strong user in each two-user group and SIC operation at the strong user in each two-user group.

V. Numerical Results

Here, we provide some numerical result to evaluate the proposed method. We considered a K-user MISO BC with $K = N_t$, where K channel vectors $\mathbf{g}_1, \cdots, \mathbf{g}_K$ were generated independently from zero-mean complex Gaussian distribution with the identity covariance matrix $\mathcal{CN}(\mathbf{0}, \mathbf{I})$. The sum rates of the proposed scheme and the full ZF scheme were computed for each channel realization and the average sum rates were obtained over 1000 independent channel realizations. The result is shown in Fig. 1. It is seen that the proposed method yields non-trivial gain over conventional full ZF downlink beamforming. It is also seen that the gain over full ZF beamforming is large at the high SNR range where the performance is limited not by noise but by interference and ZF beamforming is equivalent to MMSE beamforming. It is seen that in the case of $N_t = 32$, at low SNR, the full ZF performance is slightly better than the proposed mixture scheme although the performance gap is insignificant. This is because the sum rate lower bound in Proposition 2 is based on the assumption of sufficiently large $\lambda_1\Gamma$ and $\lambda_2\Gamma$, i.e., high operating SNR. Even with this assumption, the method works well except for very low SNR.

VI. Conclusion

In this paper, we have proposed a new framework for precoder-and-decoder design for MU-MISO downlink using linear and non-linear SIC mixture reception at the receiver side. For the proposed transceiver architecture the required feedback information at the BS is the same as that for conventional ZF beamforming, but receivers should have the capability of SIC and certain receivers need to be informed to perform SIC by some forward-link control channel. Therefore, when receivers are equipped with SIC capability in coming mobile networks, the proposed transceiver architecture provides an effective alternative to conventional fully linear downlink beamforming for MU-MISO downlink.

References

[1] J. Seo and Y. Sung, "On the operation of non-orthogonal multiple access with multiple transmit antennas," *to be submitted to IEEE Trans. Signal Process.*, May 2017.

[2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3590–3600, Nov. 2010.

[3] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3936–3964, Sep. 2009.

[4] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[5] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. on Sel. Areas in Commun.*, vol. 24, pp. 528–541, Mar. 2006.

[6] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, pp. 506–522, Feb. 2005.

[7] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Trans. on Inf. Theory*, vol. 58, pp. 2911–2934, Dec. 2012.

[8] G. Lee, Y. Sung, and M. Kountouris, "On the performance of random beamforming in sparse millimeter wave channels," *IEEE J. of Sel. Topics in Signal Process.*, vol. 10, pp. 560–575, Apr. 2016.

[9] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE VTC*, pp. 1–5, Jun. 2013.

[10] B. Clerckx, H. Joudeh, C. Hao, M. Dai, and B. Rassouli, "Rate splitting for MIMO wireless networks: A promising PHY-layer strategy for LTE evolution," *IEEE Commun. Mag.*, vol. 54, pp. 98–105, May 2016.

[11] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Trans. on Signal Process.*, vol. 56, pp. 5292–5296, Jul. 2008.

[12] E. A. Jorswieck and E. G. Larsson, "Linear precoding in multiple antenna broadcast channels: Efficient computation of the achievable rate region," in *Proc.Int. Workshop on Smart Antennas*, pp. 21–28, Feb. 2008.

[13] J. Lindblom, E. Karipidis, and E. G. Larsson, "Efficient computation of pareto optimal beamforming vectors for the miso interference channel with successive interference cancellation," *IEEE Trans. on Signal Process.*, vol. 61, pp. 4782–4795, Jul. 2013.