Beamformer Design and Power Allocation for Two-Cluster Two-User NOMA System

Carlo Piccoli, Stefano Tomasin
University of Padova, Italy
Email: tomasin@dei.unipd.it

Eduard Jorswieck
Technische Universität Dresden, Germany
Email: Eduard.Jorswieck@tu-dresden.de

***Abstract*—Non-orthogonal multiple-access (NOMA) systems are envisioned as suitable solutions for both next generation cellular and local area networks, in order to support differentiated traffic requirements and a massive number of users. In this paper we consider a scenario with a base station (BS) equipped with multiple antennas serving four users, clustered into two groups of two users each, equipped with a single antenna. This simple but relevant scenario allows to capture the main features of NOMA systems, where in each cluster the user having the best channel applies successive interference cancellation (SIC) to remove the signal intended for the other user in the same cluster, while the other user treats interference as noise. Moreover interference among the two clusters also is treated as additional noise. In this scenario we address the problem of beamformer design and power allocation for the four users with the aim of minimizing the total transmit power while satisfying a minimum rate constraint for each user. Starting from the zero-forcing (ZF) and maximal ratio combining (MRC) beamformers we propose five solutions based on suitable combinations of fixed beamformers to simplify the beamformer design and power allocation. The performance of the proposed solutions is compared in terms of total consumed power and complexity of the various solutions.**

***Index Terms*—Beamforming; Non-orthogonal multiple access; Resource Allocation.**

I. Introduction

Non-orthogonal multiple-access (NOMA) has been shown in [1] to yield significant gains over orthogonal multiple-access (OMA) schemes concerning overall cell throughput, cell-edge user throughput and proportional fairness. On the one hand, these results are achieved by applying successive interference cancellation (SIC) at the receiver and superposition of multiple-user signals. In [2] the NOMA concept is extended to multiple-antennas scenarios (multiple-input-multiple-output (MIMO)-NOMA). On the other hand, SIC still is affected by error propagation [3] and requires additional operations to decode possibly unuseful signals. In [4] it is shown that increasing the number of users in a NOMA cluster does not yield significant gain despite increasing the computational complexity since the strongest user applies SIC multiple times.

The application of MIMO to NOMA is introduced in [2], [3] and extensively studied in [5]–[10]. A survey of single-cell NOMA is presented in [5] and the concept is expanded to a general multi-cell scenario, whose practical challenges are presented. A survey of NOMA for future radio networks is presented in [6] with its various practical implementations, such as code-domain multiplexing. The first ergodic study of NOMA systems is [7], where the MIMO-NOMA capacity with statistical channel state information at the transmitter (CSIT) has been derived. In [8] projection-based hybrid NOMA beamforming and projection-based/inversion-based pairing under properties of quasi-degradation are introduced. A comprehensive investigation of uplink and downlink NOMA systems is carried out in [9], with emphasis on user grouping and power allocation. In [10] a scenario where the total number of user antennas is larger than the number of transmit antennas is investigated. In [11], a base station (BS) with multiple antennas transmits two data streams to two single-antenna users, where both streams are intended for the near user and only the latter for the far user; an iterative beamforming algorithm is proposed. In [12], a multiple-antenna BS serves multiple users and an effective user clustering is applied in conjunction with zero-forcing (ZF) beamforming.

Most of the considered literature focuses on ZF or maximal ratio combining (MRC) beamformers, although these solutions are far from being optimal (from the power or rate point of views) in the clustered scenario with SIC. Therefore in this paper we aim at addressing the issue of beamformer design and power allocation. We consider a scenario with a BS equipped with multiple antennas serving four users, clustered into two groups of two users each, equipped with a single antenna. Beyond cellular systems, this scenario applies also to Wi-Fi networks, e.g., implementing the IEEE 802.11ax standard [13], using a MIMO-NOMA system. This simple but relevant scenario allows to capture the fundamental features of NOMA systems, where in each cluster the user having the best channel applies SIC to remove the signal intended for the weaker user before proceeding with detection. The weaker user suffers the interference from the signal to the stronger user. Interference among the two clusters is treated as additional noise. In this scenario we address the problem of beamformer design and power allocation for the four users with the aim of minimizing the total transmit power while satisfying a minimum rate constraint for each user. Starting from the ZF and MRC beamformers we propose five solutions based on suitable combinations of fixed beamformers to simplify the beamformer design and power allocation. The performance of the proposed solutions is compared in terms of total consumed power and complexity of the various solutions.

In the rest of the paper, Section II introduces the system model, while the proposed beamformers and power allocation strategies are described in Section III. Numerical results are presented in Section IV before conclusions are drawn in Section V.

II. SYSTEM MODEL

We consider the scenario of a BS equipped with N antennas transmitting to four users with a single antenna each. The four users are divided into two clusters, each with two users. In each cluster we denote as *strong* the user having the highest channel gain among the two, while the other is denoted as *weak*. As usual in a NOMA scenario, signals to all users are transmitted over the same time and frequency. The strong user in each cluster performs SIC over the signal of the weak user of the same cluster whereas the weak user treats the strong-user interference as noise. Throughout this work, perfect SIC is considered and channel state information (CSI) is considered available at the BS. Without loss of generality, we consider users 1 and 2 belonging to the first cluster, with user 1 being the strong one, while users 3 and 4 belong to the second cluster, with user 3 being the strong one. As mention earlier, beyond cellular systems, this scenario applies also to the IEEE 802.11ax Wi-Fi networks [13].

The BS transmits using two complex-valued beamforming vectors $\boldsymbol{w}_n \in \mathbb{C}^{N\times 1}$, $n = 1, 2$, with $\|\boldsymbol{w}_n\| = 1$, associated to cluster n so that $\boldsymbol{w}_1$ serves user 1 and 2 in cluster 1 and $\boldsymbol{w}_2$ serves user 3 and 4 in cluster 2. The complex-valued data signal transmitted to user i is denoted $s_i \in \mathbb{C}$, and it has unitary power, zero mean and is independent of other users' signals.

The precoded signals for both clusters are

$$\begin{aligned}\boldsymbol{x}^{(1)} &= (\sqrt{p_1}s_1 + \sqrt{p_2}s_2)\,\boldsymbol{w}_1\\ \boldsymbol{x}^{(2)} &= (\sqrt{p_3}s_3 + \sqrt{p_4}s_4)\,\boldsymbol{w}_2\end{aligned} \tag{1}$$

and the overall precoded signal is $\boldsymbol{x} = \boldsymbol{x}^{(1)} + \boldsymbol{x}^{(2)}$.

The complex-valued vector $\boldsymbol{h}_i \in \mathbb{C}^{1\times N}$ represents the channel from BS to user i and the corresponding allocated power is p_i. The received signal is

$$\boldsymbol{y} = \boldsymbol{h}\boldsymbol{x} + \boldsymbol{n} \tag{2}$$

with $\boldsymbol{h} = \left[\boldsymbol{h}_1^T, \boldsymbol{h}_2^T, \boldsymbol{h}_3^T, \boldsymbol{h}_4^T\right]^T \in \mathbb{C}^{4\times N}$ being the channel matrix of the overall system, $\boldsymbol{y} \in \mathbb{C}^{4\times 1}$ and $\boldsymbol{n} \in \mathbb{C}^{4\times 1}$ the additive Gaussian noise vector with zero mean and power σ^2, equal for all users.

For user 1, i.e., the strong user in the first cluster, we have

$$y_1 = (\sqrt{p_1}s_1 + \sqrt{p_2}s_2)\,\boldsymbol{h}_1\boldsymbol{w}_1 + \boldsymbol{h}_1\boldsymbol{x}^{(2)} + n_1\,, \tag{3}$$

and after SIC (with respect to user 2) we obtain

$$\tilde{y}_1 = \sqrt{p_1}s_1\boldsymbol{h}_1\boldsymbol{w}_1 + (\sqrt{p_3}s_3 + \sqrt{p_4}s_4)\,\boldsymbol{h}_1\boldsymbol{w}_2 + n_1\,. \tag{4}$$

Instead, the weak user of each cluster treats the strong-user interference as noise; it is also needed that the strong user is able to decode the weak-user signal, through SIC, treating it as noise.

ZF beamforming: A first option for the choice of the beamformers is the ZF beamforming, where interference among the strong-user channels is nulled. Let us consider the matrix comprising the two strong channels

$$\boldsymbol{h}_s = \left[\boldsymbol{h}_1^T, \boldsymbol{h}_3^T\right]^T \tag{5}$$

and then construct the ZF beamforming vectors as

$$\boldsymbol{w} = \left[\boldsymbol{I} - \boldsymbol{h}_s^H\left(\boldsymbol{h}_s\boldsymbol{h}_s^H\right)^{-1}\boldsymbol{h}_s\right]\tilde{\boldsymbol{\Lambda}} = \left[\boldsymbol{w}_1^T, \boldsymbol{w}_2^T\right]^T \tag{6}$$

with $\tilde{\boldsymbol{\Lambda}}$ a real-positive-valued diagonal matrix to assure that $\|\boldsymbol{w}_n\| = 1$ for $n = 1, 2$.

MRC beamforming: A second beamforming choice is the MRC combining beamformer, aiming at maximizing the information transfer to strong users and neglecting the interference among them. In this case the beamformers are

$$\boldsymbol{w}_1 = \frac{\boldsymbol{h}_1^H}{\|\boldsymbol{h}_1\|}\,,\quad \boldsymbol{w}_2 = \frac{\boldsymbol{h}_3^H}{\|\boldsymbol{h}_3\|}. \tag{7}$$

This straightforward approach is usually avoided since it leads to high inter-cluster interference.

III. PROPOSED APPROACHES

In this paper we consider the problem of power allocation to each user and the choice of the beamforming vectors, with the objective of a) guaranteeing a minimum rate to each user and b) minimize the total transmit power. In particular, let $\bar{R}_i$ be the minimum rate required for user i, then in formulas, the constrained optimization problem can be written as follows:

$$\min_{\boldsymbol{p},\boldsymbol{w}_1,\boldsymbol{w}_2} \quad p_1 + p_2 + p_3 + p_4 \tag{8a}$$

subject to

$$\log_2\left(1 + \frac{p_1\left|\boldsymbol{h}_1\boldsymbol{w}_1\right|^2}{\sigma^2 + (p_3 + p_4)\left|\boldsymbol{h}_1\boldsymbol{w}_2\right|^2}\right) \ge \bar{R}_1 \tag{8b}$$

$$\min_{i\in\{1,2\}}\left\{\log_2\left(1 + \frac{p_2\left|\boldsymbol{h}_i\boldsymbol{w}_1\right|^2}{\sigma^2 + (p_3 + p_4)\left|\boldsymbol{h}_i\boldsymbol{w}_2\right|^2 + p_1\left|\boldsymbol{h}_i\boldsymbol{w}_1\right|^2}\right)\right\} \ge \bar{R}_2 \tag{8c}$$

$$\log_2\left(1 + \frac{p_3\left|\boldsymbol{h}_3\boldsymbol{w}_2\right|^2}{\sigma^2 + (p_1 + p_2)\left|\boldsymbol{h}_3\boldsymbol{w}_1\right|^2}\right) \ge \bar{R}_3 \tag{8d}$$

$$\min_{i\in\{3,4\}}\left\{\log_2\left(1 + \frac{p_4\left|\boldsymbol{h}_i\boldsymbol{w}_2\right|^2}{\sigma^2 + (p_1 + p_2)\left|\boldsymbol{h}_i\boldsymbol{w}_1\right|^2 + p_3\left|\boldsymbol{h}_i\boldsymbol{w}_2\right|^2}\right)\right\} \ge \bar{R}_4 \tag{8e}$$

$$p_i \ge 0 \qquad i = 1, \ldots, 4 \tag{8f}$$

$$\left\|\boldsymbol{w}_n\right\| = 1 \quad \forall\, n = 1, 2 \tag{8g}$$

where constraints (8b) and (8d) are for strong users, (8c) and (8e) are for weak users. The variables to be optimized are the two beamforming vectors $\boldsymbol{w}_n$ and the allocated powers p_i.

We first observe that problem (8) can be simplified by replacing the minimum rate constraints with equality constraints. In fact, if we had a power $\tilde{p}_1$ that *strictly* satisfies (8b), i.e.,

$$\log_2\left(1+\frac{\tilde{p}_1\left|\boldsymbol{h}_1\boldsymbol{w}_1\right|^2}{\sigma^2+(p_3+p_4)\left|\boldsymbol{h}_1\boldsymbol{w}_2\right|^2}\right)>\bar{R}_1 \tag{9}$$

there exists a lower power $p_1 < \tilde{p}_1$ that satisfies (8b) at equality, thus reducing the sum power, i.e., the objective function. Besides, the choice of p_1 over $\tilde{p}_1$ would also decrease the inter-cluster interference to other users, therefore we conclude that at solution problem (8) satisfies the rate constraints at equality. Note that for constraints (8c) and (8e) the minimum of the two user rates achieves the equality at optimum, while the other rate can be larger than the minimum rate.

Power Allocation for Given Beamformers: For given beamformers, the optimization problem (8) boils down to the optimal power allocation that satisfies the rate constraints at equality. Assuming that the minimum is always satisfied by the weak-user channel the optimization problem becomes a linear system of four equations in the four unknowns $p_1,\ldots,p_4$, as shown in (10) at the top of the page, with

$$A_i=\frac{1}{2^{\bar{R}_i}-1}. \tag{11}$$

If either of the strong users is achieving the minimum rate, we replace the corresponding line in the linear system with the channel of the strong user. In practice, we solve the linear system of (10) and then we compute the resulting rates for each user. If the minimum rate is not achieved by the user considered in the linear system, we solve a new linear system, otherwise we already have the global solution.

In the following we consider various options for the choice of the beamformers. Indeed, we note that the original problem (8) is in general difficult to be solved since the constraints are non-convex functions in the powers as well the beamforming vectors.

A. Inter-cluster Beamforming (INTER)

The first approach that we propose combines the ZF beamforming with the MRC with respect to the strong users. In particular, the beamformers are linear combination of the ZF and MRC beamformers, and are constructed as

$$\begin{aligned}\boldsymbol{w}_1(\gamma_1)&=\frac{\left(\gamma_1\boldsymbol{h}_3^{\perp}+(1-\gamma_1)\boldsymbol{h}_1\right)^H}{\left\|\gamma_1\boldsymbol{h}_3^{\perp}+(1-\gamma_1)\boldsymbol{h}_1\right\|}\\ \boldsymbol{w}_2(\gamma_2)&=\frac{\left(\gamma_2\boldsymbol{h}_1^{\perp}+(1-\gamma_2)\boldsymbol{h}_3\right)^H}{\left\|\gamma_2\boldsymbol{h}_1^{\perp}+(1-\gamma_2)\boldsymbol{h}_3\right\|}\,,\end{aligned} \tag{12}$$

with combining coefficients $0 \leq \gamma_1, \gamma_2 \leq 1$, explicitly indicated here in the notation of the beamformers. The combining beamformer $\boldsymbol{h}_3^{\perp}$ is the projection of $\boldsymbol{h}_1$ onto the space orthogonal to $\boldsymbol{h}_3$, i.e.,

$$\boldsymbol{h}_3^{\perp}=\boldsymbol{h}_1\boldsymbol{\Pi}_{\boldsymbol{h}_3}^{\perp}\,, \tag{13}$$

with

$$\boldsymbol{\Pi}_{\boldsymbol{h}_3}^{\perp}=\boldsymbol{I}-\boldsymbol{h}_3^H\left(\boldsymbol{h}_3\boldsymbol{h}_3^H\right)^{-1}\boldsymbol{h}_3\,, \tag{14}$$

that is the projection matrix over the null space of $\boldsymbol{h}_3$. We denote this solution as *inter-cluster (INTER) beamforming* as it only considers the strong users and through them the interaction among the two clusters.

Note that (12) boils down to ZF beamforming when $\gamma_1 = \gamma_2 = 1$ and to MRC beamforming when $\gamma_1 = \gamma_2 = 0$.

The choice of the beamformers now is reduced to the choice the two real positive numbers γ_1 and γ_2 by numerical methods.

B. Intra-cluster Beamforming (INTRA)

With the second approach we apply the ZF beamforming to *equivalent* channels, obtained as a convex combination of the strong and weak-user channels of each cluster. The resulting solution takes into account also the effects of beamforming design inside each cluster (by properly choosing the combination beamformer) and therefore we denote it as *intra-cluster (INTRA) beamforming*. In formulas, the two equivalent channels are

$$\boldsymbol{c}^{(1)}(\delta_1)=\delta_1\boldsymbol{h}_1+(1-\delta_1)\boldsymbol{h}_2\,,\ \boldsymbol{c}^{(2)}(\delta_2)=\delta_2\boldsymbol{h}_3+(1-\delta_2)\boldsymbol{h}_4 \tag{15}$$

with combining coefficients to be properly chosen, $0 \leq \delta_1, \delta_2 \leq 1$. The beamformers are designed as ZF solutions of the equivalent channels, i.e.,

$$\boldsymbol{w}_1(\delta_1,\delta_2)=\frac{\boldsymbol{c}_{\perp}^{(2)H}(\delta_1,\delta_2)}{\left\|\boldsymbol{c}_{\perp}^{(2)}(\delta_1,\delta_2)\right\|}\,,\quad \boldsymbol{w}_2(\delta_1,\delta_2)=\frac{\boldsymbol{c}_{\perp}^{(1)H}(\delta_1,\delta_2)}{\left\|\boldsymbol{c}_{\perp}^{(1)}(\delta_1,\delta_2)\right\|} \tag{16}$$

with $\boldsymbol{c}_{\perp}^{(1)}(\delta_1,\delta_2)$ being the projection of $\boldsymbol{c}^{(2)}(\delta_2)$ onto the null space of $\boldsymbol{c}^{(1)}(\delta_1)$ and, vice versa, for $\boldsymbol{c}_{\perp}^{(2)}(\delta_1,\delta_2)$. Again, INTRA boils down to ZF beamforming by setting $\delta_1 = \delta_2 = 1$. Also in this case we resort to numerical methods to optimize δ_1 and δ_2.

C. Inter-intra-cluster Beamforming (TER+TRA)

In this third approach, denoted TER+TRA, we combine the previous two approaches by constructing the beamformers as a convex combination of ZF beamforming and MRC with respect to the equivalent channels, i.e.,

$$\begin{aligned}\boldsymbol{w}_1(\gamma_1,\delta_1,\delta_2)&=\frac{\left(\gamma_1\boldsymbol{c}_{\perp}^{(2)}(\delta_1,\delta_2)+(1-\gamma_1)\boldsymbol{c}^{(1)}(\delta_1)\right)^H}{\left\|\gamma_1\boldsymbol{c}_{\perp}^{(2)}(\delta_1,\delta_2)+(1-\gamma_1)\boldsymbol{c}^{(1)}(\delta_1)\right\|}\\ \boldsymbol{w}_2(\gamma_2,\delta_1,\delta_2)&=\frac{\left(\gamma_2\boldsymbol{c}_{\perp}^{(1)}(\delta_1,\delta_2)+(1-\gamma_2)\boldsymbol{c}^{(2)}(\delta_2)\right)^H}{\left\|\gamma_2\boldsymbol{c}_{\perp}^{(1)}(\delta_1,\delta_2)+(1-\gamma_2)\boldsymbol{c}^{(2)}(\delta_2)\right\|}\,,\end{aligned} \tag{17}$$

with combining coefficients $0 \leq \gamma_1, \gamma_2 \leq 1$, which must be optimized by numerical methods.

D. Parallel Channels Approximation (PARA)

We now consider a different approach where the beamformers are not forced to be unit-norm vectors. Therefore $\left\|\boldsymbol{w}_1\right\|^2$ is the power allocated to the first cluster and $\left\|\boldsymbol{w}_2\right\|^2$ the power

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = \begin{bmatrix} A_1 \left|\boldsymbol{h}_1\boldsymbol{w}_1\right|^2 & 0 & -\left|\boldsymbol{h}_1\boldsymbol{w}_2\right|^2 & -\left|\boldsymbol{h}_1\boldsymbol{w}_2\right|^2 \\ -\left|\boldsymbol{h}_2\boldsymbol{w}_1\right|^2 & A_2\left|\boldsymbol{h}_2\boldsymbol{w}_1\right|^2 & -\left|\boldsymbol{h}_2\boldsymbol{w}_2\right|^2 & -\left|\boldsymbol{h}_2\boldsymbol{w}_2\right|^2 \\ -\left|\boldsymbol{h}_3\boldsymbol{w}_1\right|^2 & -\left|\boldsymbol{h}_3\boldsymbol{w}_1\right|^2 & A_3\left|\boldsymbol{h}_3\boldsymbol{w}_2\right|^2 & 0 \\ -\left|\boldsymbol{h}_4\boldsymbol{w}_1\right|^2 & -\left|\boldsymbol{h}_4\boldsymbol{w}_1\right|^2 & -\left|\boldsymbol{h}_4\boldsymbol{w}_2\right|^2 & A_4\left|\boldsymbol{h}_4\boldsymbol{w}_2\right|^2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 \\ \sigma^2 \\ \sigma^2 \\ \sigma^2 \end{bmatrix} \tag{10}$$

allocated to the second cluster. The original problem (8) then becomes

$$\min_{\eta_1,\eta_2,\boldsymbol{w}_1,\boldsymbol{w}_2} \quad \|\boldsymbol{w}_1\|^2 + \|\boldsymbol{w}_2\|^2 \tag{18a}$$

subject to

$$\log_2\left(1 + \frac{\eta_1\left|\boldsymbol{h}_1\boldsymbol{w}_1\right|^2}{\sigma^2 + \left|\boldsymbol{h}_1\boldsymbol{w}_2\right|^2}\right) = \bar{R}_1 \tag{18b}$$

$$\min_{i\in\{1,2\}}\left\{\log_2\left(1 + \frac{(1-\eta_1)\left|\boldsymbol{h}_i\boldsymbol{w}_1\right|^2}{\sigma^2 + \left|\boldsymbol{h}_i\boldsymbol{w}_2\right|^2 + \eta_1\left|\boldsymbol{h}_i\boldsymbol{w}_1\right|^2}\right)\right\} = \bar{R}_2 \tag{18c}$$

$$\log_2\left(1 + \frac{\eta_2\left|\boldsymbol{h}_3\boldsymbol{w}_2\right|^2}{\sigma^2 + \left|\boldsymbol{h}_3\boldsymbol{w}_1\right|^2}\right) = \bar{R}_3 \tag{18d}$$

$$\min_{i\in\{3,4\}}\left\{\log_2\left(1 + \frac{(1-\eta_2)\left|\boldsymbol{h}_i\boldsymbol{w}_2\right|^2}{\sigma^2 + \left|\boldsymbol{h}_i\boldsymbol{w}_1\right|^2 + \eta_2\left|\boldsymbol{h}_i\boldsymbol{w}_2\right|^2}\right)\right\} = \bar{R}_4 \tag{18e}$$

$$0 \le \eta_1, \eta_2 \le 1\,, \tag{18f}$$

where now η_1 is the portion of the power of the first cluster allocated to the strong user, i.e., user 1, whereas $1-\eta_1$ is the remaining portion allocated to the weak user, i.e., user 2; the same holds for η_2.

Consider the unit norm vectors

$$\boldsymbol{u} = \bar{\boldsymbol{h}}_1^{\perp} = \frac{\boldsymbol{\Pi}_{\boldsymbol{h}_1}^{\perp}\boldsymbol{h}_3^H}{\left\|\boldsymbol{\Pi}_{\boldsymbol{h}_1}^{\perp}\boldsymbol{h}_3^H\right\|} \quad \boldsymbol{v} = \bar{\boldsymbol{h}}_3^{\perp} = \frac{\boldsymbol{\Pi}_{\boldsymbol{h}_3}^{\perp}\boldsymbol{h}_1^H}{\left\|\boldsymbol{\Pi}_{\boldsymbol{h}_3}^{\perp}\boldsymbol{h}_1^H\right\|}\,, \tag{19}$$

with $\bar{\boldsymbol{h}}_1^{\perp}$ being the unit norm projection of $\boldsymbol{h}_3$ onto the null space of $\boldsymbol{h}_1$ and vice versa for $\bar{\boldsymbol{h}}_3^{\perp}$. With these constructed vectors, we have that

$$\boldsymbol{h}_1\boldsymbol{u} = \boldsymbol{h}_3\boldsymbol{v} = 0\,. \tag{20}$$

Now, we construct the beamforming vectors as a linear combination of $\boldsymbol{u}$ and $\boldsymbol{v}$ as follows

$$\boldsymbol{w}_1(\alpha,\beta) = \alpha\boldsymbol{u} + \beta\boldsymbol{v} \quad \boldsymbol{w}_2(\gamma,\delta) = \gamma\boldsymbol{u} + \delta\boldsymbol{v}. \tag{21}$$

In order to obtain a closed form solution for the beamformer design we introduce the approximation that channels inside clusters are parallel, i.e., $\boldsymbol{h}_1 \parallel \boldsymbol{h}_2$ and $\boldsymbol{h}_3 \parallel \boldsymbol{h}_4$, thus having also

$$\boldsymbol{h}_2\boldsymbol{u} = \boldsymbol{h}_4\boldsymbol{v} = 0\,. \tag{22}$$

In Appendix A we derive a close-form expression of the optimal beamformers (i.e., optimal values of α, β, γ and δ) as a function of η_1 and η_2. We denote this approach as parallel (PARA) channel approximation solution.

Fixed η_1 and η_2 within bounds: In order to simplify the choice of η_1 and η_2 we first observe that for the PARA solution we have specific bounds on these parameters, i.e.,

$$\begin{aligned} \eta_1^{\mathrm{lb}} &\le \eta_1 \le \eta_1^{\mathrm{ub}} \\ \eta_2^{\mathrm{lb}} &\le \eta_2 \le \eta_2^{\mathrm{ub}}\,, \end{aligned} \tag{23}$$

where the close-form values for the bounds are derived in Appendix A. Then we can express the values of η_1, η_2 with respect to their bounds

$$\begin{aligned} &\eta_1 = k_1\eta_1^{\mathrm{lb}} + (1-k_1)\eta_1^{\mathrm{ub}} \quad \eta_2 = k_2\eta_2^{\mathrm{lb}} + (1-k_2)\eta_2^{\mathrm{ub}} \\ &\quad 0 \le k_1, k_2 \le 1 \end{aligned} \tag{24}$$

and choose fixed values of k_1 and k_2. By doing this, η_1 and η_2 will still adapt to the changing channel conditions, but we just need to compute the bounds (23) and then use the chosen values of k_1 and k_2 in (24) to obtain η_1 and η_2. In Section IV we will optimize the fixed values of k_1 and k_2 so that the average sum power is minimized and then we keep this values to assess the performance of the PARA system.

E. Orthogonal Strong Channels Approximation (ORTH)

A last proposed solution assumes that the strong channels are orthogonal, i.e.,

$$\boldsymbol{h}_1\boldsymbol{h}_3^H = 0 \implies \boldsymbol{u}^H\boldsymbol{v} = 0 \tag{25}$$

so that we can simplify

$$\left|\boldsymbol{h}_2\boldsymbol{w}_1\right|^2 = \alpha^2\left|\boldsymbol{h}_2\boldsymbol{u}\right|^2 + \beta^2\left|\boldsymbol{h}_2\boldsymbol{v}\right|^2 + 2\alpha\beta\Re\left[\boldsymbol{h}_2\boldsymbol{v}\boldsymbol{h}_2^*\boldsymbol{u}^*\right]\,. \tag{26}$$

As reported in Appendix B, also in this case a close-form solution for the beamformers can be obtained, as a solution of a linear system of equation. We denote this solution as orthogonal (ORTH) strong channels approximation.

IV. Numerical Results

We now present numerical results for the various proposed techniques. We consider $N = 2$ antennas at the BS. Channels are assumed to be circularly-symmetric complex Gaussian random vectors with zero mean and independent entries. We set the same rate constraint $R_{\mathrm{s}} = R_1 = R_3 = 0.5$ b/s/Hz to the strong users, i.e., users 1 and 3, and rate constraint $R_{\mathrm{w}} = R_2 = R_4 = 0.1$ b/s/Hz to the weak users, i.e., users 2 and 4. We also set a constraint on the maximum instantaneous sum power $P_{\max} = 10$. Channels are generated as complex Gaussian symmetric with correlation ρ. All the proposed approaches have been numerically solved by the optimization algorithm `Global Search` in MATLAB, with the exception of PARA for which fixed values of k_1 and k_2 are chosen. For comparison purposes we also consider both

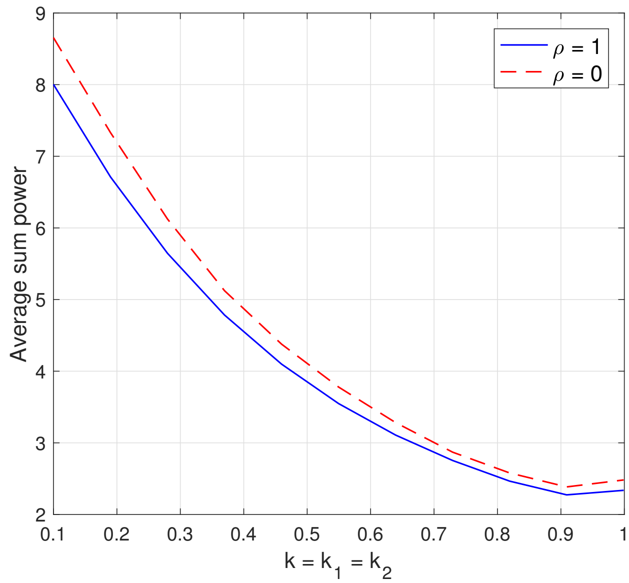

Fig. 1. Average sum power as a function of $k = k_1 = k_2$ for two values of correlation coefficient ρ.

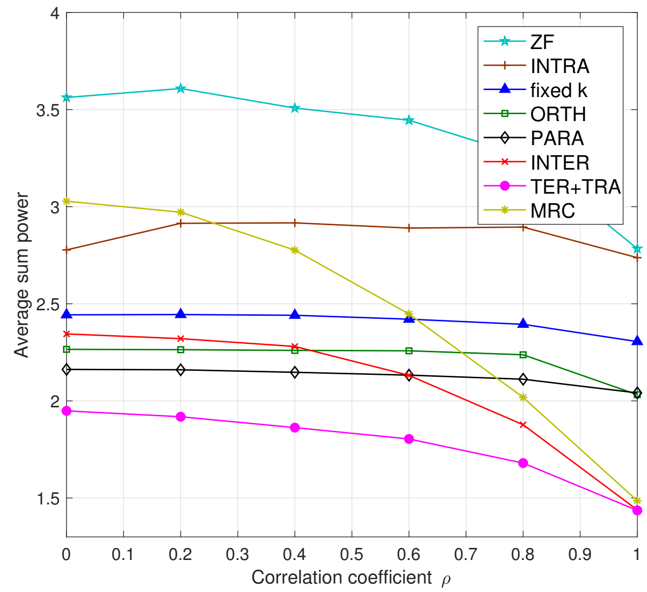

Fig. 2. Average sum power vs. the correlation coefficient.

ZF and MRC solutions, where powers are chosen to satisfy the rate constraints.

In particular for PARA, since the scenario is symmetric with respect to the clusters, we fix $k = k_1 = k_2$ and we evaluate the average sum power with respect to the correlation coefficient ρ. It is important to note that fixing k means that η_n are given by (24) and the problem can be directly solved as in (30) without the need of optimization algorithms. Fig. 1 shows the average sum power as a function of $k = k_1 = k_2$ for two extreme values of the correlation coefficient ρ, namely 0 and 1. Intermediate values of ρ provide an average sum power that is between those obtained with the extreme values. We can clearly see that the minimum is in the proximity of $k = 1$, that means taking η_1, η_2 near their lower bounds, therefore we set $k = 0.95$.

We now compare the various solutions in terms of the average sum power as a function of the correlation coefficient in Fig. 2. We observe that the best performing method providing the minimum average sum power is the TER+TRA method, followed by the INTER method. The ORTH and PARA solutions still have a remarkable performance with respect to the basic ZF solution that requires the largest total average sum power. The MRC solution exhibits a strong variations among the various values of ρ. The total power is also not very sensitive to the correlation coefficients, with a small decay of the total power as the correlation increases.

About complexity, we recall that the TER+TRA approach has four variables to be optimized and it involves the 4×4 matrix inversion to evaluate the objective function. Also INTER and INTRA techniques need the mentioned matrix inversion but also have 2 variables to be optimized. In any case, the INTER approach outperforms the INTRA approach; we also recall that for values of the correlation coefficient near 1 the INTRA technique is equal to the ZF on strong users and the INTER technique is equal to the TER+TRA technique.

V. Conclusions

In this paper we have proposed various techniques for the minimum sum-power allocation and beamformer design in a NOMA system with two clusters and four users, using SIC. In the various approaches the optimization of the beamformers is achieved by suitably combining simple beamformers designed to balance the interference among the clusters and within each cluster. One solution in particular does not require on-line optimization and still outperforms classical solutions based on ZF and MRC. By a comparison among the solutions we conclude that the TER+TRA solution achieves the lowest average sum-power.

Appendix A
PARA Beamformer Design

We now derive the beamformers for the PARA approximation. Let us define

$$\mu_1 = \frac{\|\boldsymbol{h}_2\|^2}{\|\boldsymbol{h}_1\|^2} < 1 \quad \mu_2 = \frac{\|\boldsymbol{h}_4\|^2}{\|\boldsymbol{h}_3\|^2} < 1. \tag{27}$$

This approximation actually underestimates the inter-cluster interference at the weak users hence their rate constraints will not be satisfied precisely. This reduction in weak-user rates has to be accounted for. Let

$$c_1 = |\boldsymbol{h}_1 \boldsymbol{v}|^2 \quad c_3 = |\boldsymbol{h}_3 \boldsymbol{u}|^2 , \quad K = 2\Re\left[\boldsymbol{u}^H \boldsymbol{v}\right] . \tag{28}$$

Now for given η_1 and η_2, the rate constraints provide the following linear system in the four variables $\alpha^2, \beta^2, \gamma^2, \delta^2$

$$\begin{gathered} A_1\eta_1\beta^2 - \delta^2 = \sigma^2/c_1 \\ \left[A_2 - (A_2+1)\eta_1\right]\beta^2 - \delta^2 = \sigma^2/\mu_1 c_1 \\ A_3\eta_2\gamma^2 - \alpha^2 = \sigma^2/c_3 \\ \left[A_4 - (A_4+1)\eta_2\right]\gamma^2 - \alpha^2 = \sigma^2/\mu_2 c_3 \\ \eta_n \in [0,1] \end{gathered} \tag{29}$$

with A_i as in (11). The system solution is

$$\begin{aligned} \alpha^2 &= -\frac{\sigma^2}{\mu_2 c_3}\frac{(A_3 + A_4\mu_2 + \mu_2)\eta_2 - A_4\mu_2}{(1 + A_3 + A_4)\eta_2 - A_4} \geq 0 \\ \beta^2 &= \frac{\sigma^2}{\mu_1 c_1}\frac{\mu_1 - 1}{(1 + A_1 + A_2)\eta_1 - A_2} \geq 0 \\ \gamma^2 &= \frac{\sigma^2}{\mu_2 c_3}\frac{\mu_2 - 1}{(1 + A_3 + A_4)\eta_2 - A_4} \geq 0 \\ \delta^2 &= -\frac{\sigma^2}{\mu_1 c_1}\frac{(A_1 + A_2\mu_1 + \mu_1)\eta_1 - A_2\mu_1}{(1 + A_1 + A_2)\eta_1 - A_2} \geq 0\,. \end{aligned} \tag{30}$$

In order to have real values for variables $\alpha, \beta, \gamma, \delta$ to construct the beamforming vectors as in (21) we need the solution in (30) to be non-negative.

To summarize, for each couple of values η_1 and η_2 we use (30) and then beamformers (21). Again, for the choie of η_1 and η_2 we may resort to numerical methods.

About the bounds on η_1 and η_2, by imposing that real positive solutions of the system (30) we must have

$$\begin{aligned} \eta_1^{\mathrm{lb}} &= \frac{A_2\mu_1}{A_1 + A_2\mu_1 + \mu_1} \leq \eta_1 \leq \frac{A_2}{1 + A_1 + A_2} = \eta_1^{\mathrm{ub}} \\ \eta_2^{\mathrm{lb}} &= \frac{A_4\mu_2}{A_3 + A_4\mu_2 + \mu_2} \leq \eta_2 \leq \frac{A_4}{1 + A_3 + A_4} = \eta_2^{\mathrm{ub}}. \end{aligned} \tag{31}$$

It is interesting to note that the upper bounds are constant, given the users' rate constraints, whereas the lower bounds depend on the single realization of the channels, specifically on μ_1 and μ_2 defined in (27).

Appendix B
ORTH Beamformer Design

For the ORTH solution we first observe

$$\begin{aligned} |\boldsymbol{h}_2\boldsymbol{w}_1|^2 &= \alpha^2\,|\boldsymbol{h}_2\boldsymbol{u}|^2 + \beta^2\,|\boldsymbol{h}_2\boldsymbol{v}|^2 + 2\alpha\beta\Re\left[\boldsymbol{h}_2\boldsymbol{v}\boldsymbol{h}_2^*\boldsymbol{u}^*\right] \\ &= \alpha^2 c_{21} + \beta^2 c_{22} + \alpha\beta c_{23} \\ &\simeq \alpha^2 c_{21} + \beta^2 c_{22} \end{aligned} \tag{32}$$

and $|\boldsymbol{h}_2\boldsymbol{w}_2|^2 \simeq \gamma^2 c_{21} + \delta^2 c_{22}$, $|\boldsymbol{h}_4\boldsymbol{w}_1|^2 \simeq \alpha^2 c_{41} + \beta^2 c_{42}$, $|\boldsymbol{h}_4\boldsymbol{w}_2|^2 \simeq \gamma^2 c_{41} + \delta^2 c_{42}$, thus the system becomes

$$\frac{\eta_1\beta^2 c_1}{\sigma^2 + \delta^2 c_1} = 1/A_1 \tag{33a}$$

$$\frac{(1-\eta_1)[\alpha^2 c_{21} + \beta^2 c_{22}]}{\sigma^2 + [\gamma^2 c_{21} + \delta^2 c_{22}] + \eta_1[\alpha^2 c_{21} + \beta^2 c_{22}]} = 1/A_2 \tag{33b}$$

$$\frac{\eta_2\gamma^2 c_3}{\sigma^2 + \alpha^2 c_3} = 1/A_3 \tag{33c}$$

$$\frac{(1-\eta_2)[\gamma^2 c_{41} + \delta^2 c_{42}]}{\sigma^2 + [\alpha^2 c_{41} + \beta^2 c_{42}] + \eta_2[\gamma^2 c_{41} + \delta^2 c_{42}]} = 1/A_4 \tag{33d}$$

$$0 \leq \eta_1, \eta_2 \leq 1 \tag{33e}$$

a linear system of four equations in $\alpha^2, \beta^2, \gamma^2, \delta^2$.

References

[1] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 611–615.

[2] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pp. 1–5.

[3] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 770–774.

[4] K. Higuchi and Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," *IEICE Transactions on Communications*, pp. 403–414, Mar. 2015.

[5] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, "Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges," vol. 55, no. 10, pp. 176–183.

[6] L. Dai, B. Wang, Y. Yuan, S. Han, C. l I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," vol. 53, no. 9, pp. 74–81.

[7] Q. Sun, S. Han, C. L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," vol. 4, no. 4, pp. 405–408.

[8] Z. Chen, Z. Ding, and X. Dai, "Beamforming for combating inter-cluster and intra-cluster interference in hybrid NOMA systems," vol. 4, pp. 4452–4463.

[9] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) Systems," vol. 4, pp. 6325–6343.

[10] S. Ali, E. Hossain, and D. I. Kim, "Non-orthogonal multiple access (NOMA) for downlink multiuser MIMO systems: User clustering, beamforming, and power allocation," vol. 5, pp. 565–577.

[11] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," vol. 63, no. 3, pp. 791–800.

[12] ——, "A suboptimal approach for minimum transmit power NOMA beamforming," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5.

[13] M. S. Afaqui, E. Garcia-Villegas, and E. Lopez-Aguilera, "Ieee 802.11ax: Challenges and requirements for future high efficiency wifi," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 130–137, June 2017.