NOMA Beamforming in SDMA Networks: Riding on Existing Beams or Forming New Ones?

Zhiguo Ding, *Fellow, IEEE*

***Abstract*—In this letter, the design of non-orthogonal multiple access (NOMA) beamforming is investigated in a spatial division multiple access (SDMA) legacy system. In particular, two popular beamforming strategies in the NOMA literature, one to use existing SDMA beams and the other to form new beams, are adopted and compared. The studies carried out in the letter show that the two strategies realize different tradeoffs between system performance and complexity. For example, riding on existing beams offers a significant reduction in computational complexity, at the price of a slight performance loss. Furthermore, this simple strategy can realize the optimal performance when the users' channels are structured.**

***Index Terms*—Non-orthogonal multiple access (NOMA), spatial division multiple access (SDMA), beamforming.**

I. Introduction

The design of beamforming is one of the most studied topics in the research area of non-orthogonal multiple access (NOMA) [1]. In principle, there are two types of NOMA beamforming strategies. One is to encourage multiple users to share the same beamformer, whose rationale is to treat orthogonal spatial directions as a type of bandwidth resources and try to serve as many users as possible on each spatial direction [2], [3]. The other is to generate many non-orthogonal beamformers, where a single user is served on each of these non-orthogonal beams [4], [5]. Both principles have been shown superior to orthogonal multiple access (OMA), and applicable to various communication scenarios [6]–[8].

This letter is to consider the NOMA beamforming design in a legacy system based on spatial division multiple access (SDMA), i.e., multiple primary users have been served via SDMA beamforming and additional secondary users are to be served via NOMA. Following the beamforming strategies in the NOMA literature, two straightforward designs are to either use the existing SDMA beams or generate new ones dedicated for the secondary users. The studies carried out in the letter show that riding on existing beams offers a significant reduction in computational complexity, at the price of a slight performance loss. Furthermore, this simple strategy can realize the optimal performance when the users' channels are structured, which opens up promising applications of reconfigurable intelligent surfaces (RISs) [9].

Manuscript received January 7, 2022; accepted January 22, 2022. Date of publication January 26, 2022; date of current version April 11, 2022. The work of Zhiguo Ding was supported by the UK EPSRC under grant number EP/P009719/2 and by H2020-MSCA-RISE-2020 under grant number 101006411. The associate editor coordinating the review of this letter and approving it for publication was H. Tabassum.

The author is with the School of Electrical and Electronic Engineering, The University of Manchester, Manchester M13 9PL, U.K. (e-mail: zhiguo.ding@manchester.ac.uk).

Digital Object Identifier 10.1109/LCOMM.2022.3146583

II. System Model

Consider an SDMA legacy system with K single-antenna primary users, denoted by U_i, $1 \leq i \leq K$, and one base station equipped with N antennas, where the primary users are served by using the following zero-forcing (ZF) precoder: $\mathbf{P} = c\mathbf{H}\left(\mathbf{H}^H\mathbf{H}\right)^{-1}$, where $\mathbf{H} = \left[\mathbf{h}_1 \cdots \mathbf{h}_K\right]$, $\mathbf{h}_i$ denotes U_i's channel vector, c is the power normalization parameter, i.e., $c^2 = \frac{P_{\mathrm{SDMA}}}{tr\left(\mathbf{H}(\mathbf{H}^H\mathbf{H})^{-2}\mathbf{H}^H\right)}$, $tr(\cdot)$ denotes the trace operation, and P_{SDMA} denotes the primary users' transmit power. Perfect channel state information (CSI) is assumed to be available at the base station for the implementation of SDMA.

By using NOMA, additional secondary users can be served together with the primary users. For illustration purposes, this letter focuses on serving a single secondary user, denoted by U_0. Denote $\mathbf{x}$ by the signal sent by the base station, which can be written as follows: $\mathbf{x} = \mathbf{P}\mathbf{s} + \mathbf{w}s_0$, where $\mathbf{s}$ and s_0 are the signals sent to the primary users and U_0, respectively, and $\mathbf{w}$ denotes U_0's beamforming vector.

In this letter, it is assumed that the secondary user directly decodes its own signal by treating the primary users' signals as noise, because of its limited decoding capability. Depending on the primary users' SIC strategies, U_0's data rate can be expressed differently as follows.

- If none of the primary users carries out SIC, i.e., they decode their own signals directly, U_0's data rate is given by
$$R_0 = \log\left(1 + \frac{|\mathbf{g}^H\mathbf{w}|^2}{|\mathbf{g}^H\mathbf{P}|^2 + \sigma^2}\right), \tag{1}$$
where $\mathbf{g}$ denotes U_0's channel vector and σ^2 denotes the noise power. In order to strictly guarantee the primary users' QoS requirements, the following constraints are required:
$$\log\left(1 + \frac{|\mathbf{h}_i^H\mathbf{p}_i|^2}{|\mathbf{h}_i^H\mathbf{w}|^2 + \sigma^2}\right) \geq R_i, \quad 1 \leq i \leq K, \tag{2}$$
where $\mathbf{p}_i$ denotes the i-th column of $\mathbf{P}$, and R_i denotes U_i's target data rate. Due to the use of ZF, the interference term, $\mathbf{h}_i^H\mathbf{p}_j$, $i \neq j$, can be discarded in (2).
- Consider that some primary users decode the secondary user's signal first before decoding their owns. Denote the set containing these users by $\mathcal{S}$, and its complementary set by $\mathcal{S}^c$. The secondary user's data rate is given by
$$\tilde{R}_0 = \min\left\{R_0, \log\left(1 + \frac{|\mathbf{h_i}^H\mathbf{w}|^2}{|\mathbf{h}_i^H\mathbf{P}|^2 + \sigma^2}\right), i \in \mathcal{S}\right\}. \tag{3}$$
The constraint shown in (2) is still needed for U_j, $j \in \mathcal{S}^c$, but not for U_i, $i \in \mathcal{S}$, since U_i, $i \in \mathcal{S}$, is able to

remove s_0 successfully and hence experiences the same performance as in conventional SDMA.

III. Two Strategies of Beamforming

This letter is to investigate the use of the two following types of beamforming, termed Strategies I and II.

A. Strategy I - Riding on an Existing Beam

The secondary user can simply use one of the existing beams, i.e., $\mathbf{w} = \sqrt{\alpha}\mathbf{p}_i$, where $\mathbf{p}_i$ is U_i's beam, α is to meet the power constraint, i.e., $\alpha \leq \frac{P_0}{|\mathbf{p}_i|^2}$, and P_0 denotes the transmit power budget for U_0.

- If U_i chooses to decode its own signal directly, i.e., U_i is in $\mathcal{S}^c$, by using (1) and (2), the optimization problem of interest is given by

$$\max_{\alpha} \ \log\left(1+\frac{\alpha|\mathbf{g}^H\mathbf{p}_i|^2}{|\mathbf{g}^H\mathbf{P}|^2+\sigma^2}\right) \tag{P1a}$$

$$\text{s.t. } \log\left(1+\frac{|\mathbf{h}_{\mathrm{i}}^{\mathrm{H}}\mathbf{p}_{\mathrm{i}}|^2}{\alpha|\mathbf{h}_{\mathrm{i}}^{\mathrm{H}}\mathbf{p}_{\mathrm{i}}|^2+\sigma^2}\right) \geq \mathrm{R}_{\mathrm{i}}, \quad \alpha \leq \frac{\mathrm{P}_0}{|\mathbf{p}_{\mathrm{i}}|^2}. \tag{P1b}$$

Note that U_0 does not cause interference to U_j, $j \neq i$, due to the use of ZF. With some algebraic manipulations, the optimal solution is obtained as follows:

$$\alpha_{I,i}^* = \min\left\{\frac{P_0}{|\mathbf{p}_i|^2}, \frac{\tau_i}{|\mathbf{h}_i^H\mathbf{p}_i|^2}\right\}, \tag{4}$$

for $\mathbf{w} = \sqrt{\alpha}\mathbf{p}_i$, where $\tau_i = \max\left\{0, \frac{|\mathbf{h}_i^H\mathbf{p}_i|^2}{2^{R_i}-1} - \sigma^2\right\}$.

- U_i can also decode U_0's signal first before decoding its own, i.e., U_i is in $\mathcal{S}$. In this case, by using (2) and (3), the optimization problem of interest is given by

$$\max_{\alpha} \ \min\left\{\log\left(1+\frac{\alpha|\mathbf{g}^H\mathbf{p}_i|^2}{|\mathbf{g}^H\mathbf{P}|^2+\sigma^2}\right),\ \log\left(1+\frac{\alpha|\mathbf{h_i}^H\mathbf{p}_i|^2}{|\mathbf{h}_i^H\mathbf{P}|^2+\sigma^2}\right)\right\} \quad \text{s.t. } \alpha \leq \frac{\mathrm{P}_0}{|\mathbf{p}_{\mathrm{i}}|^2}, \tag{P2a}$$

whose solution is given by $\alpha_{II,i}^* = \frac{P_0}{|\mathbf{p}_i|^2}$.

By comparing the data rates realized by $\alpha_{II,i}^*$ and $\alpha_{I,i}^*$, the optimal solution for Strategy I can be found straightforwardly.

B. Strategy II - Forming a New Beam

Forming a new beam will be more complicated than using an existing SDMA beam, as shown in the following.

- If all the primary users decode their own signals directly, by using (1) and (2), the optimization problem of interest is given by

$$\max_{\mathbf{w}} \ \log\left(1+\frac{|\mathbf{g}^H\mathbf{w}|^2}{|\mathbf{g}^H\mathbf{P}|^2+\sigma^2}\right) \quad \text{s.t. } |\mathbf{w}|^2 \leq \mathrm{P}_0,\ \log\left(1+\frac{|\mathbf{h}_i^H\mathbf{p}_i|^2}{|\mathbf{h}_i^H\mathbf{w}|^2+\sigma^2}\right) \geq R_i, \quad 1 \leq i \leq K, \tag{P3a}$$

which can be recast as the following equivalent form:

$$\max_{\mathbf{w}} \ |\mathbf{g}^H\mathbf{w}|^2 \tag{P4a}$$

$$\text{s.t. } |\mathbf{h}_{\mathrm{i}}^{\mathrm{H}}\mathbf{w}|^2 \leq \tau_{\mathrm{i}}, \quad 1 \leq \mathrm{i} \leq \mathrm{K}, \ |\mathbf{w}|^2 \leq \mathrm{P}_0. \tag{P4b}$$

Problem P4 is a nonconvex quadratically constrained quadratic program and can be solved by applying SDR [10], i.e., problem P4 can be relaxed as follows:

$$\max_{\mathbf{W}} \ tr\left(\mathbf{GW}\right) \tag{P5a}$$

$$\text{s.t. } \mathrm{tr}\left(\mathbf{H}_{\mathrm{i}}\mathbf{W}\right) \leq \tau_{\mathrm{i}}, \quad 1 \leq \mathrm{i} \leq \mathrm{K}, \ \mathrm{tr}\left(\mathbf{W}\right) \leq \mathrm{P}_0, \tag{P5b}$$

where the rank-one constraint is omitted, $\mathbf{G} = \mathbf{g}\mathbf{g}^H$, and $\mathbf{H}_i = \mathbf{h}_i\mathbf{h}_i^H$. Problem P5 can be straightforwardly solved by applying optimization solvers.

- If U_i, $i \in \mathcal{S}$, decodes the secondary user's signal first, by using (2) and (3), the optimization problem of interest is given by

$$\max_{\mathbf{w}} \ \min\left\{R_0, \log\left(1+\frac{|\mathbf{h_i}^H\mathbf{w}|^2}{|\mathbf{h}_i^H\mathbf{P}|^2+\sigma^2}\right), i \in \mathcal{S}\right\}$$
$$\text{s.t. } \log\left(1+\frac{|\mathbf{h}_{\mathrm{j}}^{\mathrm{H}}\mathbf{p}_{\mathrm{j}}|^2}{|\mathbf{h}_{\mathrm{j}}^{\mathrm{H}}\mathbf{w}|^2+\sigma^2}\right) \geq \mathrm{R}_{\mathrm{j}}, \quad \mathrm{j} \in \mathcal{S}^c,\quad |\mathbf{w}|^2 \leq P_0. \tag{P6a}$$

Problem P6 can be recast equivalently as follows:

$$\max_{t,\mathbf{w}} \ t \tag{P7a}$$

$$\text{s.t. } \log\left(1+\frac{|\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2}{|\mathbf{g}^{\mathrm{H}}\mathbf{P}|^2+\sigma^2}\right) \geq \mathrm{t} \tag{P7b}$$

$$\log\left(1+\frac{|\mathbf{h_i}^H\mathbf{w}|^2}{|\mathbf{h}_i^H\mathbf{P}|^2+\sigma^2}\right) \geq t, i \in \mathcal{S} \tag{P7c}$$

$$\log\left(1+\frac{|\mathbf{h}_j^H\mathbf{p}_j|^2}{|\mathbf{h}_j^H\mathbf{w}|^2+\sigma^2}\right) \geq R_j, \quad j \in \mathcal{S}^c, \tag{P7d}$$

$$|\mathbf{w}|^2 \leq P_0, \tag{P7e}$$

which can be further simplified as follows:

$$\max_{z,\mathbf{w}} \ z \tag{P8a}$$

$$\text{s.t. } |\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2 \geq \mathrm{a}_0 \ \mathrm{z}, \quad |\mathbf{h_i}^{\mathrm{H}}\mathbf{w}|^2 \geq \mathrm{a}_{\mathrm{i}}\mathrm{z}, \ \mathrm{i} \in \mathcal{S} \tag{P8b}$$

$$|\mathbf{h}_j^H\mathbf{w}|^2 \leq \tau_j, \quad j \in \mathcal{S}^c, \ |\mathbf{w}|^2 \leq P_0, \tag{P8c}$$

where $z = 2^t - 1$, $a_0 = |\mathbf{g}^H\mathbf{P}|^2 + \sigma^2$, and $a_i = |\mathbf{h}_i^H\mathbf{P}|^2 + \sigma^2$. Again, problem P8 can be solved by applying SDR as shown previously.

- If all the primary users decode the secondary user's signal first before decoding their own, by using (2), the optimization problem of interest is given by

$$\max_{\mathbf{w}} \ \min\left\{R_0, \log\left(1+\frac{|\mathbf{h_i}^H\mathbf{w}|^2}{|\mathbf{h}_i^H\mathbf{P}|^2+\sigma^2}\right), 1 \leq i \leq K\right\} \tag{P9a}$$

$$\text{s.t. } |\mathbf{w}|^2 \leq \mathrm{P}_0, \tag{P9b}$$

which can be recast equivalently as follows:

$$\max_{z,\mathbf{w}} \ z \tag{P10a}$$

$$\text{s.t. } |\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2 \geq \mathrm{a}_0\mathrm{z} \tag{P10b}$$

$$|\mathbf{h_i}^H\mathbf{w}|^2 \geq a_i z, \quad 1 \leq i \leq K, \ |\mathbf{w}|^2 \leq P_0. \tag{P10c}$$

Similar to problem P4, problem P10 can be solved by applying SDR.

By comparing the solutions obtained for problems P3, P6 and P9, a solution can be obtained for Strategy II.

C. Optimality of the Obtained SDR Solutions

Because the SDR solutions have been obtained by removing the rank-one constraint, they are not guaranteed to be optimal. The optimality of the SDR solutions for the two-user special case can be established as follows. Note that this conclusion can also be proved by applying Theorem 3.2 in [11], but the steps shown in the proof will be useful for the follow-up discussions about the general case with K users.

Proposition 1: For the two-user special case with random realizations of complex-valued channel coefficients, where the base station is equipped with two antennas, the obtained solutions via SDR are optimal.

Proof: The proposition can be proved by simply showing that the obtained SDR solutions are always rank-one. Without loss of generality, problem P9 is focused, where the proofs for the other cases can be obtained straightforwardly. After applying SDR, problem P9 can be expressed as follows:

$$\min_{z,\mathbf{w}} \ -z \tag{P11a}$$
$$\text{s.t. } \mathrm{tr}\left(\mathbf{G}\mathbf{W}\right) \geq \mathrm{a}_0 \mathrm{z} \tag{P11b}$$
$$tr\left(\mathbf{H}_i\mathbf{W}\right) \geq a_i z, \quad 1 \leq i \leq 2 \tag{P11c}$$
$$tr\left(\mathbf{W}\right) \leq P_0, \quad \mathbf{W} \succeq 0, \tag{P11d}$$

where the rank-one constraint is ignored. The corresponding Lagrange can be written as follows [12]:

$$\begin{aligned} &L(z, \mathbf{W}, \lambda_1, \cdots, \lambda_4, \boldsymbol{\lambda}) \\ &\quad = -z + \lambda_1\left(a_0 z - tr\left(\mathbf{G}\mathbf{W}\right)\right) \\ &\qquad + \sum_{i=2}^{3} \lambda_i \left(a_{i-1} z - tr\left(\mathbf{H}_{i-1}\mathbf{W}\right)\right) \\ &\qquad + \lambda_4\left(tr\left(\mathbf{W}\right) - P_0\right) - tr\left(\boldsymbol{\lambda}\mathbf{W}\right), \end{aligned} \tag{5}$$

where λ_i denotes the Lagrange multipliers. Because problem P11 is convex, the use of the Karush-Kuhn-Tucker (KKT) conditions is applicable and leads to the following:

$$-\lambda_1\mathbf{G} - \sum_{i=2}^{3} \lambda_i \mathbf{H}_{i-1} + \lambda_4 \mathbf{I} - \boldsymbol{\lambda} = 0. \tag{6}$$

and

$$\boldsymbol{\lambda}\mathbf{W} = 0. \tag{7}$$

In the remainder of the proof, we focus on the case with $\mathbf{W} \neq 0$ to avoid the trivial all-zero solution.

If $\boldsymbol{\lambda} = 0$, the use of (6) yields the following:

$$\lambda_4 \mathbf{I} = \lambda_1 \mathbf{G} + \sum_{i=2}^{3} \lambda_i \mathbf{H}_{i-1}. \tag{8}$$

Note that $\lambda_1\mathbf{G} + \sum_{i=2}^{3} \lambda_i \mathbf{H}_{i-1}$ is a 2×2 complex-valued matrix. For a fixed λ_4, (8) represents a set of linear equations with 3 unknown variables $(\lambda_1, \ldots, \lambda_3)$ and 8 equations. Because the channel coefficients are assumed to be random variables, no solution exists for this overdetermined set of linear equations, and hence $\boldsymbol{\lambda} \neq 0$. By using (7) and the fact that $\boldsymbol{\lambda} \neq 0$, one can conclude that $\mathbf{W}$ is not full-rank. For the two-user case, if $\mathbf{W}$ is not full-rank, the rank of a feasible non-zero $\mathbf{W}$ has to be one, which proves the proposition. □

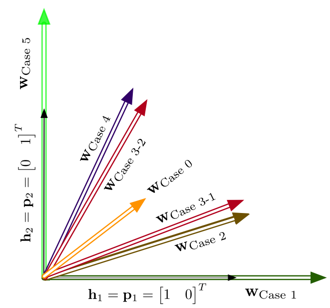

Fig. 1. Illustration of optimal beamforming, where $\mathbf{g} = [\sin(\theta) \quad \cos(\theta)]^T$ and θ is chosen as shown in Table I. For illustration purposes, large scale path loss is omitted, $P_{\text{SDMA}} = 30$ dBm, $N = K$, $R_i = 1$ bit per channel use (BPCU), and $\sigma^2 = -10$ dBm. In particular, $\mathbf{w}_{\text{Case 1}} = 1.50[1 \quad 0]^T$, $\mathbf{w}_{\text{Case 2}} = [1.18 \quad 0.32]^T$, $\mathbf{w}_{\text{Case 3-1}} = [1.13 \quad 0.47]^T$, $\mathbf{w}_{\text{Case 3-2}} = [0.47 \quad 1.13]^T$, $\mathbf{w}_{\text{Case 4}} = [0.32 \quad 1.18]^T$, $\mathbf{w}_{\text{Case 5}} = 1.50[0 \quad 1]^T$, $\mathbf{w}_{\text{Case 0}} = 0.50[1 \quad 1]^T$.

Note that there are a few situations, where SDR solutions are not rank-one, as explained in the following.

1) The Two-User Case With $h_1 \perp h_2$: If $\mathbf{h}_1 \perp \mathbf{h}_2$, by choosing $\lambda_1 = 0$, and $\lambda_{i-1} = \frac{1}{|\mathbf{h}_i|^2}$ for $2 \leq i \leq 3$, $\sum_{i=2}^{3} \lambda_i \mathbf{H}_i$ is a product of a unitary matrix and itself, and hence becomes an identity matrix. As a result, the equality in (8) becomes possible, i.e., $\boldsymbol{\lambda} = 0$ becomes possible, which can lead to a situation that $\mathbf{W}$ is full-rank.

2) The Two-User Case With Real-Valued Channel Coefficients: In this case, (8) can be viewed as a set of 3 linear equations, one to ensure the diagonal elements of $\lambda_1\mathbf{G} + \sum_{i=2}^{3} \lambda_i \mathbf{H}_i$ to be the same, the other two to ensure the off-diagonal elements to be zero. As a result, it is possible to find λ_i, $1 \leq i \leq 3$, to satisfy (8), i.e., $\boldsymbol{\lambda} = 0$ becomes possible and $\mathbf{W}$ is not necessarily rank-one.

3) The General Case With $K > 2$: For the general case, to establish the rank-one conclusion, it is not sufficient to just show $\mathbf{W}$ rank-deficient. The key step to establish the rank-one conclusion is to rewrite (6) as follows:

$$\boldsymbol{\lambda} = \lambda_{K+2}\mathbf{I} - \lambda_1 \mathbf{G} - \sum_{i=2}^{K+1} \lambda_i \mathbf{H}_{i-1}. \tag{9}$$

Assuming that $\lambda_1\mathbf{G} + \sum_{i=2}^{K+1} \lambda_i \mathbf{H}_{i-1}$ has two different largest eigenvalues, denoted by λ_1^* and λ_2^*, $\lambda_1^* > \lambda_2^*$, $\mathbf{W}$ can be proved to be rank-one, by using the following steps:

- If $\lambda_1^* < \lambda_{K+2}$, $\boldsymbol{\lambda}$ is full-rank, which is not possible due to the constraint $\boldsymbol{\lambda}\mathbf{W} = 0$ shown in (7).
- If $\lambda_1^* = \lambda_{K+2}$, and $\lambda_2^* < \lambda_{K+2}$, the dimension of the null space of $\boldsymbol{\lambda}$ is one, and hence $\mathbf{W}$ is rank-one, due to the constraint $\boldsymbol{\lambda}\mathbf{W} = 0$.
- If $\lambda_2^* \geq \lambda_{K+2}$, $\boldsymbol{\lambda}$ has at least one negative eigenvalue, which is not possible, since $\boldsymbol{\lambda}$ is positive semi-definite.

Unfortunately, our carried out simulation results indicate that $\lambda_1\mathbf{G} + \sum_{i=2}^{K+1} \lambda_i \mathbf{H}_{i-1}$ can have two repeated eigenvalues. When this situation happens, the rank of $\mathbf{W}$ becomes two, and the use of Gaussian randomization procedure is needed.

TABLE I

THE DETERMINISTIC CASES USED TO GENERATE FIG. 1

	θ	P_0 in dBm	Adopted Beamforming
Case 1	$\frac{\pi}{2}$	30	Strategies I and II (SIC at U_1)
Case 2	$\frac{\pi}{3}$	30	Strategy II (SIC at U_1)
Case 3	$\frac{\pi}{4}$	30	Strategy II (SIC at U_1 or U_2)
Case 4	$\frac{\pi}{6}$	30	Strategy II (SIC at U_2)
Case 5	0	30	Strategies I and II (SIC at U_2)
Case 0	$\frac{\pi}{4}$	27	Strategy II (no SIC at U_1 & U_2)

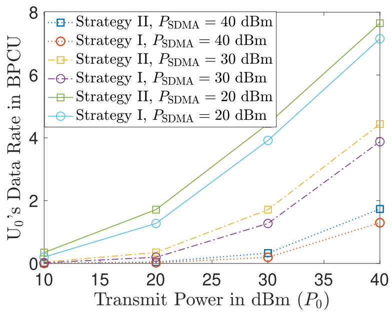

Fig. 2. Performance comparison of the two strategies with different P_0.

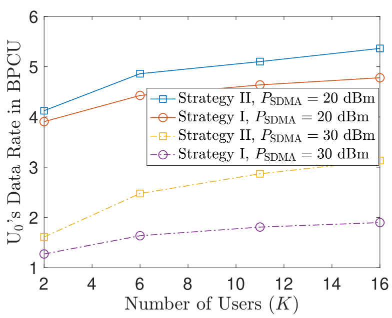

Fig. 3. Performance of the strategies with different K. $P_0 = 30$ dBm.

IV. SIMULATION

In this section, the computer simulation results are used to evaluate the performance of the two beamforming strategies. A deterministic two-user case is first focused on in Fig. 1, where Strategy II yields the optimal solution[1], as indicated by Proposition 1. The aim is to investigate whether the use of Strategy I can also lead to the optimal performance. As can be seen from the figure, for the case where the primary users' channels are orthogonal to each other and the secondary user's channel is aligned with one primary user's, the use of Strategy I can yield the optimal performance. Fig. 1 also shows that the optimal choice of $\mathbf{w}$ is depending on $\mathbf{g}$ and P_0.

In Figs. 2 and 3, both path loss and small scale fading are considered. In particular, the K primary users are randomly located in a square with edge 6 m and the base station located at its center, where the location of the secondary user is fixed at $(0, 1)$ m. The path loss exponent is set as 3, the noise power is $\sigma^2 = -94$ dBm, $N = K$, and $R_i = 1$ bit per channel use (BPCU). Fig. 2 focuses on the two-user case, and shows that Strategy II outperforms Strategy I, at the price of high computational complexity. In Fig. 3, a general multi-user case is considered, where a simplified Strategy II is adopted to reduce complexity by focusing on the following two cases, one to allow all the primary users to decode their own signals directly and the other to allow the primary user with the largest channel vector norm to carry out SIC. As shown in Fig. 3, the performance gain of Strategy II over Strategy I is increased by increasing K; however, it is important to point out that the complexity of SDR also grows by increasing K.

[1]To avoid the sub-optimality issue discussed in Section III, the users' real-valued channels are generated with small complex-valued perturbations.

V. CONCLUSION

In this letter, the design of NOMA beamforming has been investigated in an SDMA legacy system. In particular, two popular beamforming designs in the NOMA literature have been adopted and shown to realize different tradeoffs between system performance and complexity. In particular, Strategy I can be implemented with low-complexity, since its solution can be obtained in a closed form. Strategy II can yield a significant performance gain over Strategy I, but at a price of high computational complexity. An important direction for future research is to investigate how the two beamforming strategies can be extended to the scenario with multiple secondary users by applying sophisticated resource allocation algorithms [13].

ACKNOWLEDGMENT

The author thanks Drs. Xiaofang Sun, Fang Fang, Jingjing Cui, and Hong Xing for useful discussions related to SDR.

REFERENCES

[1] M. Vaezi *et al.*, *Multiple Access Techniques for 5G Wireless Networks and Beyond*. Springer, 2019.

[2] Z. Ding *et al.*, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, Jan. 2016.

[3] W. Shin *et al.*, "Coordinated beamforming for multi-cell MIMO-NOMA," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 84–87, Jan. 2017.

[4] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, Mar. 2015.

[5] M. S. Ali *et al.*, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[6] Z. Sun and Y. Jing, "Transmission and clustering designs for multi-antenna NOMA based on average transmit power," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3412–3427, Apr. 2021.

[7] Z. Shang *et al.*, "Secure transmission for NOMA-based cognitive radio networks with imperfect CSI," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2517–2521, Aug. 2021.

[8] J. Zhu *et al.*, "Enhanced user grouping and power allocation for hybrid mmWave MIMO-NOMA systems," *IEEE Trans. Wireless Commun.*, early access, Sep. 8, 2021, doi: 10.1109/TWC.2021.3109053.

[9] M. D. Renzo *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 1–20, May 2019.

[10] Z.-Q. Luo *et al.*, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[11] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[12] X. Sun *et al.*, "Joint beamforming and power allocation in downlink NOMA multiuser MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5367–5381, Aug. 2018.

[13] Y. Sun *et al.*, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.