Weighted Sum Rate Maximization for DMA-Based MU-MIMO System

Seraphin F. Kimaryo
Department of Electrical and Information Engineering
Seoul National University of Science and Technology
Seoul, Republic of Korea
email:seraphin@seoultech.ac.kr

Kyungchun Lee
Department of Electrical and Information Engineering and
the Research Center for Electrical and Information Technology
Seoul National University of Science and Technology
Seoul, Republic of Korea
email:kclee@seoultech.ac.kr

***Abstract*—Advancements and emergencies in consumer technologies such as 6G, virtual reality, augmented reality, and now artificial intelligence put more demands on wireless communication networks to be able to accommodate the resulting large data and low latency transmissions. In addition to the several ongoing efforts to realize this goal, this work proposes the combination of multiple-input multiple-output (MIMO), dynamic metasurface antennas (DMA), and millimeter wave (mmWave) technologies to optimize the downlink weighted sum rate (WSR) of the MIMO system operating at the mmWave band, and whose users are all equipped with DMA. We formulate the system model and the associated WSR maximization problem, which is later solved by the proposed algorithm designed to tackle the optimization of each variable using novel techniques that guarantee fast convergence. The numerical results demonstrate the effectiveness of the proposed algorithm in achieving higher performance than the conventional MIMO system with the same number of radio frequency chains.**

***Index Terms*—Metamaterials, dynamic metasurface antenna (DMA), downlink MIMO communication, weighted sum rate (WSR), 6G.**

I. Introduction

Artificial intelligence (AI) is taking the world to the new unprecedented technological shift that is envisioned to greatly simplify several aspects of daily human activities. Just like the industrial revolution in the 18th century, which was pivotal in transitioning most procedural tasks into faster, efficient, and reliable machines, the 21st century's AI is aimed at further simplification of challenging cognitive tasks that the former machines could not handle [1], [2]. As the name suggests, the core of AI technology is to give these machines, through a technical process that we call in this work "model training", some human-like (artificial) intelligence such that they can reason and take action based on the context under consideration.

Looking at this technological shift through the lens of wireless communication and signal processing reveals more stringent demands in terms of large data transfer and low-latency connectivity between the consuming user equipment (UE) such as mobile phones, laptops, cars, etc., and the cloud servers hosting the trained models[1]. Luckily, multiple-input multiple-output (MIMO), a technology that uses multiple antennas at the transmitter and/or the receiver [3], was proposed in the early 2010s to enable wireless communication systems to become a large data traffic bearer [4]. Essentially, MIMO exploits the rich-scattering property of the wireless channel to spatially multiplex multiple data streams and transmit them simultaneously. In scenarios of poor-scattering channels, the MIMO's multiple antennas can be used to beamform the signals in the desired direction, thereby creating a relatively stronger signal at the receiver, thus allowing the use of high-order modulation schemes that carry more bits per symbol, hence achieving the same goal of higher data rate transmission. The realization of multiple antennas at most UEs is, however, still very challenging due to their small sizes and limited power budgets, limiting them from accommodating the large numbers of bulky and power-hungry radio frequency (RF) chains needed to drive each attached antenna. To overcome this issue, some previous efforts have resorted to hybrid systems that employ fewer RF chains and connect them through a network of phase shifters to a large number of antennas [5].

Noticing the use of extra circuitry of phase shifters in hybrid systems, which does not greatly simplify the hardware bulkiness of transceivers, this work takes a different approach of replacing the traditional metallic antennas with the metamaterial-inspired type of antenna called dynamic metasurface antennas (DMA) [6] that aggregates multiple transmitting/receiving elements within a waveguide structure called microstrip and connects each microstrip to its own RF chain [7]. Being derived from metamaterials, DMA brings several advantages. First, its elements are known for their low power consumption and small sizes, thus allowing them to be packed in large amounts within space-limited UEs [8]. Secondly, it provides reconfigurability options for each element [9], increasing the overall degree of freedom of the system [10]. Lastly, it naturally achieves RF chain reduction without

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2019R1A6A1A03032119 and in part by the NRF Grant funded by the Korean Government (MSIT) under Grant NRF-2022R1A2C1006566.

[1]It is understood that some of these models can be deployed in the UEs themselves; however, due to the sheer size of most of the models, the complexity of their execution, and their frequent updating requirement, they necessitate being hosted in large-storage and computationally powerful cloud servers.

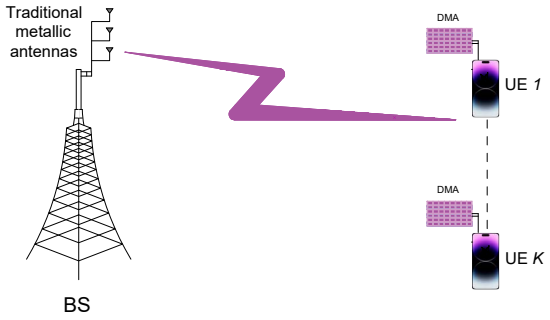

Fig. 1: DMA-Based MU-MIMO communication system.

needing extra hardware—this greatly reduces the bulkiness and power consumption of the resulting transceivers [11].

Contributions: There are various ways through which the aforementioned advantages can be exploited. This work, however, focuses on the downlink weighted sum rate (WSR) maximization of the multiuser (MU) MIMO system, where all the users are equipped with DMA. We formulate the associated system model and then the WSR maximization problem. Finally, we propose efficient algorithms for optimizing each user's transmit precoder and DMA weights.

Notations: This work adopts the following notations. Scalars, vectors, and matrices are respectively denoted by lowercase italic letters, boldface lowercase letters, and boldface uppercase letters. The Euclidean norm and the i-th element of vector $\boldsymbol{a}$ are denoted by $\|\boldsymbol{a}\|$ and $[\boldsymbol{a}]_i$, respectively. The element at the i-th row and j-th column of the matrix $\boldsymbol{A}$ is denoted by $[\boldsymbol{A}]_{i,j}$, whereas its determinant, inverse, trace, and Frobenius norm are denoted by $|\boldsymbol{A}|$, $\boldsymbol{A}^{-1}$, $\mathrm{Tr}\,(\boldsymbol{A})$, and $\|\boldsymbol{A}\|_F$, respectively. An $N \times N$ diagonal matrix with diagonal elements $x_1, \cdots, x_N$ is denoted by $\mathrm{diag}\,(x_1, \cdots, x_N)$. The indicator function and the operator that rounds its argument to the nearest larger integer are denoted by $\mathbb{1}_{(\cdot)}$ and $\lceil \cdot \rceil$, respectively. $vec(\boldsymbol{A})$ vectorizes the matrix $\boldsymbol{A}$ by stacking $[\boldsymbol{A}]_i$ above $[\boldsymbol{A}]_{i+1}$.

II. System Model and Problem Formulation

In this work, we consider a communication setup depicted in Fig. 1, where a BS equipped with N_B metallic antennas sends signals to K users who use DMA as their receiving antennas. The DMA of user u contains a total of $N_u = M_u \times L_u$ receiving antennas distributed across M_u microstrips, each containing L_u elements. The DMA microstrips normally attenuate the signal as it travels through them. This is, unfortunately, a natural phenomenon that also occurs in many other mediums like wireless channels. Therefore, a signal reaching element l located at a distance $\partial_{i,l}$ from the input port of microstrip i experiences the following amount of attenuation

$$h_{i,l} = e^{-\partial_{i,l}(\alpha_i + j\beta_i)}, \tag{1}$$

where α_i and β_i denote the attenuation constant and the wavenumber of the given microstrip, respectively. Another characteristic property of DMA is the ability of its elements to reconfigure their permeability and permittivity properties, which, as a result, changes the amplitude and phase of the interacting signal. This reconfiguration is governed by the Lorentzian form, which, in the case of the element considered in (1), is given by

$$q_{i,l} = \left\{ \frac{j + e^{j\theta_{i,l}}}{2} \middle| \theta_{i,l} \in [0, 2\pi] \right\}, \tag{2}$$

with $\theta_{i,l}$ being its tunable phase shift. The parameter $q_{i,l}$ is commonly referred to in the literature as the DMA configurable weight of element l of the i-th microstrip; thus, we call it the same hereinafter. The DMA configurable weights and the attenuation experienced by all N_u elements are conveniently collected in matrices

$$[\boldsymbol{Q}_u]_{m,n} \in \mathbb{C}^{M_u \times N_u} = \begin{cases} q_{m,\mathrm{mod}(n-1,L_u)+1}, & (m-1)\,L_u + 1 \le n \le mL_u, \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

and

$$[\boldsymbol{H}_u]_{(i-1)L_u+l,(i-1)L_u+l} \in \mathbb{C}^{N_u \times N_u} = h_{i,l}, \tag{4}$$

respectively, where $m, i \in [1, \cdots, M_u]$, $n \in [1, \cdots, N_u]$, $l \in [1, \cdots, L_u]$, and $\mathrm{mod}(\cdot,\cdot)$ is the modulus operator. The $\Pi_u = \min(N_B, M_u)$ unit-power data streams, $\boldsymbol{s}_u \in \mathbb{C}^{\Pi_u \times 1}$, for user u are firstly precoded at the BS by a precoder $\boldsymbol{F}_u \in \mathbb{C}^{N_B \times \Pi_u}$ to create the transmitted symbol

$$\boldsymbol{x} \in \mathbb{C}^{N_B \times 1} = \sum_{u=1}^{K} \boldsymbol{F}_u \boldsymbol{s}_u, \tag{5}$$

which is then sent to all users with a power P. The signal arriving at user u via the channel $\boldsymbol{G}_u \in \mathbb{C}^{N_u \times N_B}$, which is assumed to be known after being estimated by techniques such as [12], [13], is weighted by the weight of each element and then guided through the attenuating microstrips to their connected RF chains. The signal observed by the RF chains of this user can be represented as

$$\boldsymbol{y}_u \in \mathbb{C}^{M_u \times 1} = \boldsymbol{Q}_u \boldsymbol{H}_u \boldsymbol{G}_u \boldsymbol{x} + \boldsymbol{Q}_u \boldsymbol{H}_u \boldsymbol{z}_u, \tag{6}$$

where $\boldsymbol{z}_u \in \mathbb{C}^{N_u \times 1}$ is the additive circularly symmetric white Gaussian noise with zero mean and variance σ_u^2. On top of $\boldsymbol{z}_u$, in the perspective of user u, the signal $\boldsymbol{Q}_u \boldsymbol{H}_u \boldsymbol{G}_u \sum_{i \ne 1}^{K} \boldsymbol{F}_i \boldsymbol{s}_i$ in (6) is interference and thus undesired. With such consideration, the rate for this user is given as

$$R_u = \log_2 \left| \boldsymbol{I} + \boldsymbol{C}_u^{-1} \boldsymbol{Q}_u \boldsymbol{H}_u \boldsymbol{G}_u \boldsymbol{K}_u \boldsymbol{G}_u^H \boldsymbol{H}_u^H \boldsymbol{Q}_u^H \right|, \tag{7}$$

where $\boldsymbol{K}_u = \boldsymbol{F}_u \boldsymbol{F}_u^H$, $\boldsymbol{C}_u = \sigma_u^2 \boldsymbol{Q}_u \boldsymbol{H}_u \boldsymbol{H}_u^H \boldsymbol{Q}_u^H + \boldsymbol{Q}_u \boldsymbol{H}_u \boldsymbol{G}_u \sum_{i \ne u}^{K} \boldsymbol{K}_i \boldsymbol{G}_u^H \boldsymbol{H}_u^H \boldsymbol{Q}_u^H$, and $\boldsymbol{I}$ is an identity matrix. In this work, we aim to find the transmit precoders and the element weights of each user that maximize the WSR of the system while ensuring the transmit power budget is not

exceeded and all element weights follow (2). This problem is mathematically formulated as

$$\mathcal{P}(1): \quad \max_{\boldsymbol{F}_u, \boldsymbol{Q}_u, \forall u} \quad \sum_{u=1}^{K} \omega_u R_u \tag{8a}$$

$$\text{s.t.} \quad \sum_{u=1}^{K} \mathrm{Tr}\left(\boldsymbol{K}_u\right) \leq P, \tag{8b}$$

$$(2), \tag{8c}$$

where ω_u is the priority of user u. In the next section, we lay out the procedures for solving this problem.

III. Proposed Solution to $\mathcal{P}(1)$

Notice that $\mathcal{P}(1)$ requires the optimization of two variables, which can be seen in the objective function (7), to be tightly coupled to one another. This coupling generally complicates their joint optimization; thus, we propose to solve for one variable at a time while fixing the other in its previous update state. This procedure is described next.

A. Optimization of $\boldsymbol{F}_u, u \in [1, \cdots, K]$ for fixed $\boldsymbol{Q}_u$

The subproblem resulting from the fixation of $\boldsymbol{Q}_u$ takes the following simplified form:

$$\mathcal{P}(1\text{-}\boldsymbol{F}_u): \quad \max_{\boldsymbol{F}_u, \forall u} \quad \sum_{u=1}^{K} \omega_u R_u \tag{9a}$$

$$\text{s.t.} \quad \sum_{u=1}^{K} \mathrm{Tr}\left(\boldsymbol{K}_u\right) \leq P. \tag{9b}$$

Next, we quote in Lemma 1 below a brief summary of Lemma 4.1 of [14], which is used to initiate the solution to problem $\mathcal{P}(1\text{-}\boldsymbol{F}_u)$.

Lemma 1. For any arbitrary matrices $\boldsymbol{M} \in \mathbb{C}^{m \times n}$ and $\boldsymbol{N} \in \mathbb{C}^{n \times l}$, and a positive definite matrix $\boldsymbol{O} \in \mathbb{C}^{m \times m}$, the following holds:

$$\log\left|\boldsymbol{I} + \boldsymbol{M}\boldsymbol{N}\boldsymbol{N}^H\boldsymbol{M}^H\boldsymbol{O}^{-1}\right| = \max_{\boldsymbol{L}, \boldsymbol{R}} \quad \log|\boldsymbol{R}| + l$$
$$-\mathrm{Tr}\left(\boldsymbol{R}\left(\boldsymbol{I} - \boldsymbol{L}^H\boldsymbol{M}\boldsymbol{N}\right)\left(\boldsymbol{I} - \boldsymbol{L}^H\boldsymbol{M}\boldsymbol{N}\right)^H + \boldsymbol{R}\boldsymbol{L}^H\boldsymbol{O}\boldsymbol{L}\right), \tag{10}$$

where $\boldsymbol{L}$ and a positive semidefinite matrix, $\boldsymbol{R}$, are auxiliary variables, whose solutions are given by

$$\tilde{\boldsymbol{L}} = \left(\boldsymbol{O} + \boldsymbol{M}\boldsymbol{N}\boldsymbol{N}^H\boldsymbol{M}^H\right)^{-1}\boldsymbol{M}\boldsymbol{N} \tag{11}$$

and

$$\tilde{\boldsymbol{R}} = \left(\boldsymbol{I} - \tilde{\boldsymbol{L}}^H\boldsymbol{M}\boldsymbol{N}\right)^{-1}, \tag{12}$$

respectively.

Note that the left-hand side of (10) and right-hand side of (7) exhibit the structural similarity, which suggests that the results of Lemma 1 can be used to recast $\mathcal{P}(1\text{-}\boldsymbol{F}_u)$ as follows:

$$\mathcal{P}(1\text{-}\boldsymbol{F}_u): \quad \min_{\boldsymbol{F}_u, \boldsymbol{V}_u, \boldsymbol{W}_u} \quad f_1\left(\boldsymbol{F}_u, \boldsymbol{V}_u, \boldsymbol{W}_u\right) \tag{13a}$$

$$\text{s.t.} \quad \sum_{u=1}^{K} \mathrm{Tr}\left(\boldsymbol{K}_u\right) \leq P, \tag{13b}$$

where

$$f_1\left(\boldsymbol{F}_u, \boldsymbol{V}_u, \boldsymbol{W}_u\right)$$
$$= \sum_{u=1}^{K} \omega_u \,\mathrm{Tr}\left(\boldsymbol{W}_u\left(\boldsymbol{I} - \boldsymbol{V}_u^H\bar{\boldsymbol{G}}_u\boldsymbol{F}_u\right)\left(\boldsymbol{I} - \boldsymbol{V}_u^H\bar{\boldsymbol{G}}_u\boldsymbol{F}_u\right)^H\right)$$
$$+ \sum_{u=1}^{K} \omega_u \,\mathrm{Tr}\left(\boldsymbol{W}_u\boldsymbol{V}_u^H\boldsymbol{C}_u\boldsymbol{V}_u\right) - \sum_{u=1}^{K} \omega_u \log_2|\boldsymbol{W}_u| \tag{14}$$

is the new objective function derived from the right-hand side of (10), and $\bar{\boldsymbol{G}}_u = \boldsymbol{Q}_u\boldsymbol{H}_u\boldsymbol{G}_u$; $\boldsymbol{V}_u$ and $\boldsymbol{W}_u$ are the introduced auxiliary variables, whose closed-form solutions given by

$$\boldsymbol{V}_u = \left(\boldsymbol{C}_u + \bar{\boldsymbol{G}}_u\boldsymbol{K}_u\bar{\boldsymbol{G}}_u^H\right)^{-1}\bar{\boldsymbol{G}}_u\boldsymbol{F}_u \tag{15}$$

and

$$\boldsymbol{W}_u = \left(\boldsymbol{I} - \boldsymbol{V}_u^H\bar{\boldsymbol{G}}_u\boldsymbol{F}_u\right)^{-1}, \tag{16}$$

follow from (11) and (12), respectively. Unfortunately, the solution of $\boldsymbol{F}_u$ is still not that straightforward. To obtain its tractable solution, we first fix $\boldsymbol{V}_u$ and $\boldsymbol{W}_u$ and apply the Lagrange multiplier method to the resulting problem, i.e.,

$$\frac{\partial}{\partial \boldsymbol{F}_u} f_1\left(\boldsymbol{F}_u, \boldsymbol{V}_u, \boldsymbol{W}_u\right) + \mu\frac{\partial}{\partial \boldsymbol{F}_u}\left(\sum_{u=1}^{K}\mathrm{Tr}\left(\boldsymbol{K}_u\right) - P\right) = 0, \tag{17}$$

and solve for $\boldsymbol{F}_u$, which yields

$$\boldsymbol{F}_u = \left(\mu\boldsymbol{I} + \boldsymbol{X}\right)^{-1}\omega_u\bar{\boldsymbol{G}}_u^H\boldsymbol{V}_u\boldsymbol{W}_u, \tag{18}$$

where $\frac{\partial}{\partial x}g(x)$ denotes the derivative of $g(x)$ with respect to x and $\boldsymbol{X} = \sum_{u=1}^{K}\omega_u\bar{\boldsymbol{G}}_u^H\boldsymbol{V}_u\boldsymbol{W}_u\boldsymbol{V}_u^H\bar{\boldsymbol{G}}_u$. Here, the parameter μ known as the Lagrange multiplier can be computed using one-dimensional search techniques, e.g., bisection method, such that

$$\sum_{u=1}^{K}\sum_{i=1}^{N_B}\frac{\left[\omega_u^2\boldsymbol{\Psi}^H\bar{\boldsymbol{G}}_u^H\boldsymbol{V}_u\boldsymbol{W}_u^2\boldsymbol{V}_u^H\bar{\boldsymbol{G}}_u\boldsymbol{\Psi}\right]_{i,i}}{\left([\boldsymbol{\Sigma}]_{i,i} + \mu\right)^2} = P, \tag{19}$$

which is an expanded and equality form of the power constraint (8b), is satisfied. In (19), the unitary matrix, $\boldsymbol{\Psi}$, and the diagonal matrix, $\boldsymbol{\Sigma}$, are from the eigenvalue decomposition (EVD) of $\boldsymbol{X}$, i.e., $\boldsymbol{X} = \boldsymbol{\Psi}\boldsymbol{\Sigma}\boldsymbol{\Psi}^H$. The procedure for optimization of $\boldsymbol{F}_u$ is organized in steps 3–7 of Algorithm 1.

B. Optimization of $\boldsymbol{Q}_u, u \in [1, \cdots, K]$ for fixed $\boldsymbol{F}_u$

Interestingly, the solution for $\mathcal{P}(2\text{-}\boldsymbol{Q}_u)$ after fixing $\boldsymbol{F}_u$ is still kick-started by the result of Lemma 1.

$$\mathcal{P}(2\text{-}\boldsymbol{Q}_u): \quad \max_{\boldsymbol{Q}_u, \forall u} \quad \sum_{u=1}^{K} \omega_u R_u \tag{20a}$$

$$\text{s.t.} \quad (2). \tag{20b}$$

Specifically, by letting $\boldsymbol{A}_u$ and $\boldsymbol{B}_u$ be the auxiliary variables that respectively correspond to $\boldsymbol{L}$ and $\boldsymbol{R}$ in (10), $\mathcal{P}(2\text{-}\boldsymbol{Q}_u)$ can be transformed to its equivalent form

$$\mathcal{P}(2\text{-}\boldsymbol{Q}_u): \quad \min_{\boldsymbol{Q}_u, \boldsymbol{A}_u, \boldsymbol{B}_u, \forall u} \quad f_2\left(\boldsymbol{Q}_u, \boldsymbol{A}_u, \boldsymbol{B}_u\right) \tag{21a}$$

$$\text{s.t.} \quad (2), \tag{21b}$$

where

$$\begin{aligned} & f_2\left(\boldsymbol{Q}_u, \boldsymbol{A}_u, \boldsymbol{B}_u\right) \\ & \overset{(a)}{=} \sum_{u=1}^{K} \omega_u \operatorname{Tr}\left(\boldsymbol{B}_u \left(\boldsymbol{I} - \boldsymbol{Y}_u \bar{\boldsymbol{Q}}_u \boldsymbol{\Gamma}_{u,u}\right)\left(\boldsymbol{I} - \boldsymbol{Y}_u \bar{\boldsymbol{Q}}_u \boldsymbol{\Gamma}_{u,u}\right)^H\right) \\ & + \sum_{u=1}^{K} \omega_u \operatorname{Tr}\left(\boldsymbol{B}_u \boldsymbol{Y}_u \bar{\boldsymbol{Q}}_u \left(\sigma_u^2 \boldsymbol{I} + \sum_{i \neq u}^{K} \boldsymbol{\Gamma}_{u,i} \boldsymbol{\Gamma}_{u,i}^H\right) \bar{\boldsymbol{Q}}_u^H \boldsymbol{Y}_u^H\right) \\ & \qquad - \sum_{u=1}^{K} \omega_u \log_2 \left|\boldsymbol{B}_u\right|, \end{aligned} \tag{22}$$

$\boldsymbol{Y}_u = \boldsymbol{A}_u^H \bar{\boldsymbol{H}}_u$, and $\boldsymbol{\Gamma}_{u,i} = \boldsymbol{G}_u \boldsymbol{F}_i$. Moreover, $\left[\bar{\boldsymbol{Q}}_u\right]_{i,i} = \left[\boldsymbol{Q}_u\right]_{\lceil \frac{i}{L_u} \rceil, i}$, $\left[\bar{\boldsymbol{H}}_u\right]_{\lceil \frac{i}{L_u} \rceil, i} = \left[\boldsymbol{H}_u\right]_{i,i}$ for $i \in \{1, \cdots, N_u\}$, and (a) is due to the fact that $\bar{\boldsymbol{H}}_u \bar{\boldsymbol{Q}}_u = \boldsymbol{Q}_u \boldsymbol{H}_u$. Similarly, as we did earlier, the closed-form solutions for $\boldsymbol{A}_u$ and $\boldsymbol{B}_u$ given by

$$\boldsymbol{A}_u = \left(\boldsymbol{C}_u + \bar{\boldsymbol{H}}_u \bar{\boldsymbol{Q}}_u \boldsymbol{G}_u \boldsymbol{K}_u \boldsymbol{G}_u^H \bar{\boldsymbol{Q}}_u^H \bar{\boldsymbol{H}}_u^H\right)^{-1} \bar{\boldsymbol{H}}_u \bar{\boldsymbol{Q}}_u \boldsymbol{G}_u \boldsymbol{F}_u \tag{23}$$

and

$$\boldsymbol{B}_u = \left(\boldsymbol{I} - \boldsymbol{Y}_u \bar{\boldsymbol{Q}}_u \boldsymbol{\Gamma}_{u,u}\right)^{-1}, \tag{24}$$

follow from (11) and (12), respectively. To find $\boldsymbol{Q}_u$, we start by examining the objective function (22) and notice two important insights that can greatly simplify the subsequent analysis. First, its subtrahend is constant with respect to $\boldsymbol{Q}_u$; hence it can be dropped. Secondly, because DMA is used for reception in this work, the weight $\boldsymbol{Q}_u$ only affects the signal of user u, which suggests that we can independently optimize the weights of different users without impacting the overall system capacity. Leveraging on the aforementioned insights, we proceed by representing the following matrices in their EVD form, i.e., $\boldsymbol{B}_u = \boldsymbol{U}_u \boldsymbol{\Lambda}_u^{1/2} \boldsymbol{\Lambda}_u^{1/2} \boldsymbol{U}_u^H$ and $\sigma_u^2 \boldsymbol{I} + \sum_{i \neq u}^{K} \boldsymbol{\Gamma}_{u,i} \boldsymbol{\Gamma}_{u,i}^H = \boldsymbol{P}_u \boldsymbol{D}_u^{1/2} \boldsymbol{D}_u^{1/2} \boldsymbol{P}_u^H$. Next, we let $\boldsymbol{\Omega}_u = \boldsymbol{P}_u \boldsymbol{D}_u^{1/2}$ and apply the splitting technique on $\bar{\boldsymbol{Q}}_u$ such that $2\bar{\boldsymbol{Q}}_u = \boldsymbol{S}_u + \boldsymbol{E}_u$, where $\boldsymbol{S}_u = j\boldsymbol{I}$, $\boldsymbol{E}_u = \operatorname{diag}\left(\bar{q}_1, \bar{q}_2, \cdots, \bar{q}_{N_u}\right)$, $\bar{q}_i = e^{j\varphi_i}$, $\varphi_i = \theta_{\lceil \frac{i}{L_u} \rceil, \max\left(\vartheta_i, L_u \mathbb{1}_{\vartheta_i = 0}\right)}$, and $\vartheta_i = \operatorname{mod}\left(i, L_u\right)$ [7]. These decompositions help to transform (22) into the following tractable form

$$\begin{aligned} & f_2'\left(\boldsymbol{E}_u\right) \\ & = \omega_u \left\| \boldsymbol{\Lambda}_u^{1/2} \boldsymbol{U}_u^H \left(\boldsymbol{I} - 0.5 \boldsymbol{Y}_u \left(\boldsymbol{S}_u + \boldsymbol{E}_u\right) \boldsymbol{\Gamma}_{u,u}\right) \right\|_F^2 \\ & \qquad + \omega_u \left\| 0.5 \boldsymbol{\Lambda}_u^{1/2} \boldsymbol{U}_u^H \boldsymbol{Y}_u \left(\boldsymbol{S}_u + \boldsymbol{E}_u\right) \boldsymbol{\Omega}_u \right\|_F^2, \end{aligned} \tag{25a}$$

$$= \omega_u \left\| \boldsymbol{C}_{2u} - \boldsymbol{J}_u \boldsymbol{E}_u \boldsymbol{\Gamma}_{u,u} \right\|_F^2 + \omega_u \left\| \boldsymbol{C}_{3u} + \boldsymbol{J}_u \boldsymbol{E}_u \boldsymbol{\Omega}_u \right\|_F^2, \tag{25b}$$

where $\boldsymbol{C}_{2u} = \boldsymbol{\Lambda}_u^{1/2} \boldsymbol{U}_u^H - \boldsymbol{J}_u \boldsymbol{S}_u \boldsymbol{\Gamma}_{u,u}$, $\boldsymbol{C}_{3u} = \boldsymbol{J}_u \boldsymbol{S}_u \boldsymbol{\Omega}_u$, and $\boldsymbol{J}_u = 0.5 \boldsymbol{\Lambda}_u^{1/2} \boldsymbol{U}_u^H \boldsymbol{Y}_u$. To further simplify $f_2'\left(\boldsymbol{E}_u\right)$ we convert the Forbenius norm in (25b) to Euclidean norm by letting $\boldsymbol{T}_{1u} = \left[\operatorname{diag}\left(\left[\boldsymbol{\Gamma}_{u,u}\right]_1\right) \boldsymbol{J}_u^T, \cdots, \operatorname{diag}\left(\left[\boldsymbol{\Gamma}_{u,u}\right]_{\Pi_u}\right) \boldsymbol{J}_u^T\right]^T$, $\boldsymbol{T}_{2u} = \left[\operatorname{diag}\left(\left[\boldsymbol{\Omega}_u\right]_1\right) \boldsymbol{J}_u^T, \cdots, \operatorname{diag}\left(\left[\boldsymbol{\Omega}_u\right]_{N_u}\right) \boldsymbol{J}_u^T\right]^T$, $\boldsymbol{v}_{2u} = vec\left(\boldsymbol{C}_{2u}\right)$, $\boldsymbol{v}_{3u} = vec\left(\boldsymbol{C}_{3u}\right)$, and vectorizing $\boldsymbol{E}_u$ into $\boldsymbol{q}_u \in \mathbb{C}^{N_u \times 1}$ such that $\left[\boldsymbol{q}_u\right]_n = \left[\boldsymbol{E}_u\right]_{n,n}$. This procedure yields

$$f_2'\left(\boldsymbol{q}_u\right) = \omega_u \left\| \boldsymbol{v}_{2u} - \boldsymbol{T}_{1u} \boldsymbol{q}_u \right\|^2 + \omega_u \left\| \boldsymbol{v}_{3u} + \boldsymbol{T}_{2u} \boldsymbol{q}_u \right\|^2, \tag{26}$$

which finally transforms $\mathcal{P}(2\text{-}\boldsymbol{Q}_u)$ to

$$\mathcal{P}(2\text{-}\boldsymbol{q}_u): \quad \min_{\boldsymbol{q}_u} \quad f_2'\left(\boldsymbol{q}_u\right) \tag{27a}$$

$$\text{s.t.} \quad \left|\left[\boldsymbol{q}_u\right]_n\right|^2 = 1, \forall n. \tag{27b}$$

Because $f_2'\left(\boldsymbol{q}_u\right)$ is continuous and differentiable over the domain of $\boldsymbol{q}_u$, i.e.,

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{q}_u} f_2'\left(\boldsymbol{q}_u\right) = & 2\omega_u \left(\boldsymbol{T}_{1u}^H \boldsymbol{T}_{1u} + \boldsymbol{T}_{2u}^H \boldsymbol{T}_{2u}\right) \boldsymbol{q}_u \\ & + 2\omega_u \left(\boldsymbol{T}_{2u}^H \boldsymbol{v}_{3u} - \boldsymbol{T}_{1u}^H \boldsymbol{v}_{2u}\right), \end{aligned} \tag{28}$$

problem $\mathcal{P}(2\text{-}\boldsymbol{q}_u)$ can then be solved by a variety of techniques such as manifold optimization [7] or gradient projection [15]. The summary of the optimization of $\boldsymbol{Q}_u$ is given in steps 8–15 of Algorithm 1. The algorithm starts with a random initialization of $\boldsymbol{Q}_u$ and $\boldsymbol{F}_u$, $\forall u$, based on its inputs, $\boldsymbol{G}_u$, $\boldsymbol{H}_u$, transmit power, P, user priority, ω_u, and user's noise variance, σ_u^2. Finally, when it converges, it returns each user's transmit precoder and DMA weights.

IV. Numerical Results

To demonstrate the effectiveness of the proposed DMA-based system, we present its WSR performance in this section. We compare this performance with that of the conventional MIMO (Conv. MIMO) system that has N_B and M_u antennas at the BS and each user, respectively, connected with their dedicated RF chains. We also adopt a random weight scheme denoted by "Rand. $\boldsymbol{Q}_u$" that randomly chooses the DMA weight of each user to demarcate the low-bound performance.

Algorithm 1: Proposed DMA scheme (Prop. DMA)

Input : $\boldsymbol{G}_u$, P, $\boldsymbol{H}_u$, ω_u, and $\sigma_u^2, \forall u$

Initialization: Randomly generate $\boldsymbol{Q}_u$ according to (3) and $\boldsymbol{F}_u$'s such that $\sum_{u=1}^{K} \mathrm{Tr}\left(\boldsymbol{K}_u\right) \leq P$.

repeat

  while $\sum_{u=1}^{K} \omega_u \log_2 |\boldsymbol{W}_u|$ has not converged **do**

    Update $\boldsymbol{V}_u$'s according to (15).

    Update $\boldsymbol{W}_u$'s according to (16).

    Update $\boldsymbol{F}_u$ according to (18).

  end

  while (22) has not converged **do**

    Update $\boldsymbol{A}_u$'s according to (23).

    Update $\boldsymbol{B}_u$'s according to (24).

    for $u = 1;\ u \leq K;\ u = u + 1$ **do**

      Solve $\mathcal{P}(2\text{-}\boldsymbol{q}_u)$ to obtain $\boldsymbol{q}_u$.

      $\bar{\boldsymbol{Q}}_u = 0.5\left(\boldsymbol{S}_u + \mathrm{diag}\left(\boldsymbol{q}_u\right)\right)$.

    end

  end

until $\sum_{u=1}^{K} \omega_u R_u$ converges

Output: $\boldsymbol{F}_u$ and $\bar{\boldsymbol{Q}}_u$, $\forall u$

A. Simulation settings

We consider a system where $K = 4$ users, who are equipped with DMA having $M_1 = 2$, $M_2 = 4$, $M_3 = 6$, and $M_4 = 8$ microstrips, respectively, receive a signal through a single-cluster 20-ray mmWave channel [5], sent with a power $P = 23$ dBm from a $D_u = 1000$ m distant BS. The microstrip of each user, whose attenuation constant and wavenumber are given by $\alpha_i = 0.6$ m^{-1} and $\beta_i = 827.67$ m^{-1}, respectively, are equipped with $L_u = 4$ elements spaced at $d = \lambda/2$, with λ being the wavelength of the 28 GHz carrier frequency. The noise variance, priority, and the path loss of user u are given by $\sigma_u^2 = -80$ dBm, $\omega_u = \frac{\psi_u}{\sum_{j=1}^{K} \psi_j}$, and

$$PL\left(D_u\right)[\mathrm{dB}] = 35.6 + 22\log\left(D_u\right), \tag{29}$$

respectively, where $\psi_j = 10^{0.1PL(D_j)}$.

B. Performance analysis:

We start by presenting in Fig. 2 the WSR performance of various schemes against the number of transmit antennas at the BS. First, we observe that the performance of all the schemes improves with the number of BS transmit antennas, mainly due to the increased number of supported streams and array gain. Interestingly, the Prop. DMA algorithm is observed to attain the highest performance of all the compared schemes. Quantitatively, we note that Conv. MIMO scheme needs twice the number of RF chains to achieve the performance of Prop. DMA algorithm. This remarkable performance of the DMA system over Conv. MIMO system is due to two factors. First, the capability of the DMA to pack a large number of elements within its microstrip greatly elevates the array gain past that of Conv. MIMO system. The second reason is obviously the efficiency of the proposed DMA scheme in effectively

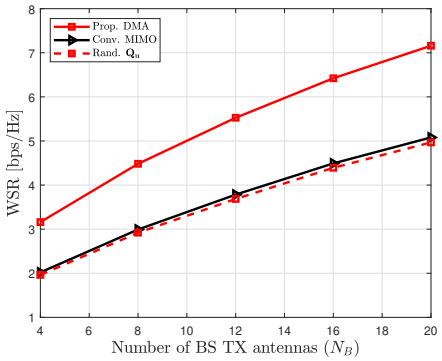

Fig. 2: WSR vs. N_B for $K = 4$ and $P = 23$ dBm.

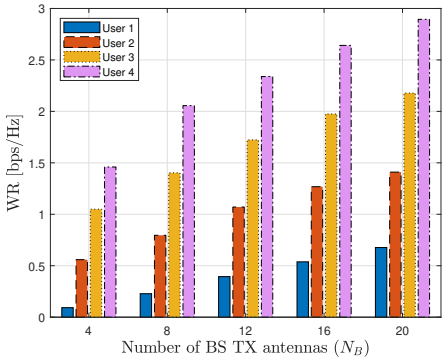

Fig. 3: User WR of the Prop. DMA scheme for $K = 4$ and $P = 23$ dBm.

materializing the aforementioned gain by properly optimizing the transmit precoder and the DMA weights. This fact can be supported by looking at the performance of the Rand. $\boldsymbol{Q}_u$ scheme. Specifically, despite using DMA, its performance is lowest because the DMA weights are not properly tuned.

Next, in Fig. 3, we unpack the performance of the Prop. DMA system by examining the individual weighted rate (WR) of each user. As expected, the WR of user 1 is the lowest because it has the fewest number of microstrip and RF chains, whereas user 4 has the highest WR due to its highest number of RF chains that can support many streams. Moreover, observe that as the number of BS transmit antennas increases, so does the performance of each user. This is due to the improved array gain and increased number of streams for users 2–4.

Finally, we provide in Fig. 4 the per-user WR comparison between the conventional MIMO and the proposed DMA-based systems when the BS is equipped with 20 antennas. Due to the same reasons stated for Fig. 2, the WR performance of the Prop. DMA algorithm is observed to be higher than that of Conv. MIMO scheme for each user. Similarly, as mentioned in the analysis of Fig. 3, the performance of user 4 for all the schemes is highest because of its highest number of supported streams derived from its large number of RF chains, i.e., $M_4 = 8$, whereas the performance of user 1,

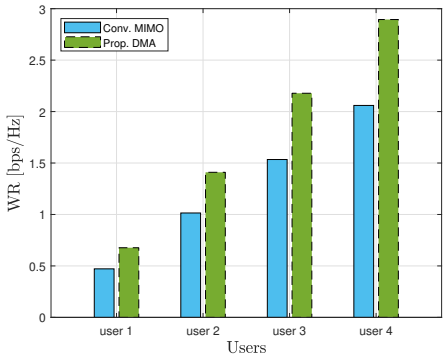

Fig. 4: User WR comparison between Conv. MIMO and Prop. DMA schemes for $N_B = 20$ and $P = 23$ dBm.

which has only $M_1 = 2$ RF chains, is the lowest. The superlative performance of the proposed DMA-based system and its proposed procedures for optimizing the DMA weights and transmit precoders clearly demonstrate the feasibility of using DMA to realize high-capacity wireless communication systems.

V. Conclusion

This work studied the downlink MU-MIMO system, where users are all equipped with DMA. Noting the continuous surge in the demands for high-capacity links between different nodes in the wireless networks, we focused on the downlink WSR maximization. Specifically, we started with the layout of the system model, followed by the formulation of the WSR maximization problem, and lastly, we proposed efficient techniques for optimizing the transmit precoders and the DMA weights of each user. The presented simulation results demonstrated the remarkable performance gains of the proposed DMA-based system and its proposed algorithm in attaining higher performance with respect to the benchmark schemes.

References

[1] Y.-S. Ong and A. Gupta, "AIR5: Five pillars of artificial intelligence research," *IEEE Trans. Emerging Top. Comput. Intell.*, vol. 3, no. 5, pp. 411–415, 2019.

[2] C. Huang, Z. Zhang, B. Mao, and X. Yao, "An overview of artificial intelligence ethics," *IEEE Trans. Artif. Intell.*, vol. 4, no. 4, pp. 799–819, 2023.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.

[4] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, 2014.

[5] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.

[6] S. F. Kimaryo and K. Lee, "Uplink and downlink capacity maximization of a P2P DMA-based communication system," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2024.

[7] S. F. Kimaryo and K. Lee, "Downlink beamforming for dynamic metasurface antennas," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4745–4755, 2023.

[8] N. Shlezinger, G. C. Alexandropoulos, M. F. Imani, Y. C. Eldar, and D. R. Smith, "Dynamic metasurface antennas for 6G extreme massive MIMO communications," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 106–113, 2021.

[9] D. R. Smith, J. B. Pendry, and M. C. K. Wiltshire, "Metamaterials and negative refractive index," *Science*, vol. 305, no. 5685, pp. 788–792, 2004.

[10] D. R. Smith, O. Yurduseven, L. P. Mancera, P. Bowen, and N. B. Kundtz, "Analysis of a waveguide-fed metasurface antenna," *Phys. Rev. Appl.*, vol. 8, no. 5, p. 054 048, Nov. 2017.

[11] N. Shlezinger, O. Dicker, Y. C. Eldar, I. Yoo, M. F. Imani, and D. R. Smith, "Dynamic metasurface antennas for uplink massive MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6829–6843, 2019.

[12] S. Liu and X. Huang, "Sparsity-aware channel estimation for mmWave massive MIMO: A deep CNN-based approach," *China Commun.*, vol. 18, no. 6, pp. 162–171, 2021.

[13] Y. Wang, W. Xu, H. Zhang, and X. You, "Wideband mmWave channel estimation for hybrid massive MIMO with low-precision ADCs," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 285–288, 2019.

[14] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841–2853, 2015.

[15] J. Tranter, N. D. Sidiropoulos, X. Fu, and A. Swami, "Fast unit-modulus least squares with applications in beamforming," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2875–2887, 2017.