Dynamic Metasurfaces for Massive MIMO Networks

Nir Shlezinger, Or Dicker, Yonina C. Eldar, Mohammadreza F. Imani, and David R. Smith

***Abstract*—Massive multiple-input multiple-output (MIMO) communications are the focus of considerable interest in recent years. While theoretical gains of such massive MIMO have been established, implementing MIMO systems with large-scale antenna arrays in practice is challenging. Among the practical difficulties associated with massive MIMO implementations are increased cost, power consumption, and physical size. In this work we study the implementation of massive MIMO antenna arrays using dynamic metasurface antennas (DMAs), an emerging technology which inherently handles the aforementioned challenges. DMAs realize planar large-scale arrays of tunable antenna elements, and can adaptively incorporate compression and analog combining in the physical antenna structure, thus reducing cost and power consumption. We first propose a mathematical model for massive MIMO systems with DMAs and discuss their constraints compared to ideal antenna arrays. Then, we characterize the fundamental limits of the resulting systems, and propose an algorithm for designing practical DMAs to approach these limits. Our numerical results indicate that the performance of practical DMA-based massive MIMO systems is comparable with ideal antenna arrays.**

***Index terms*— Massive MIMO, metasurface antennas.**

I. INTRODUCTION

Future wireless systems are required to support an increasing number of end-users with growing throughput demands. Recent years have witnessed a growing interest in massive multiple-input multiple-output (MIMO) systems, in which the base station (BS) is equipped with a large antenna array, as a method for meeting these demands and increasing the spectral efficiency (SE). In particular, such systems can increase the throughput in a manner which is scalable with the number of BS antennas [1].

The theoretical benefits of massive MIMO systems in terms of SE are well-established [2]–[4]. However, implementing a massive MIMO BS with a standard antenna array, capable of achieving these benefits, is still a very challenging task. In particular, some of the difficulties which arise when realizing large-scale antenna arrays include high cost [5], [6], increased power consumption [7], and constrained physical size and shape [8], [9]. Several signal processing methods have been studied, aimed at tackling these difficulties. The proposed approaches include introducing analog combining to reduce the size and cost of the system [5], [10]; implementing low-resolution quantization and/or antenna selection to mitigate the power consumption [7], [11]–[15]; and utilizing efficient power amplifiers operating at reduced peak-to-average-power ratio [16], [17]. Nonetheless, all these approaches assume a fixed optimal antenna array, and attempt to tackle the difficulties which arise from this antenna array architecture from a signal processing perspective.

In parallel to the ongoing efforts to make massive MIMO feasible using signal processing techniques, a large body of research has focused on designing practical antenna arrays for massive MIMO systems [8], [9], [18]. An emerging technology for realizing large-scale antenna arrays of small physical size uses metamaterial radiators instead of conventional antenna elements. Metamaterial antennas consist of arrays of sub-wavelength radiators placed close to each other and are able to realize desired radiation patterns [19]. While the resulting antenna arrays typically exhibit mutual coupling and frequency selectivity, a larger number of tunable metamaterial antenna elements can be packed in the same physical area. For example, an antenna array of 144 metamaterial radiators operating at 60 GHz can be realized at a physical size which is less than 1% of that required when using conventional geometries [9]. Most previous works on metamaterial antennas for MIMO communications focus on designing the physical antenna structure and metamaterial substrate to satisfy desired requirements, such as gain, bandwidth, efficiency, and level of mutual correlation [18]–[21]. Thus, the resulting antenna structure is fixed and independent of the processing which the signals undergo.

Recently, dynamic metasurface antennas (DMAs) have been proposed as a method for electrically tunning the physical characteristics of metamaterial antennas [22], [23]. DMAs inherently implement signal processing techniques such as beamforming, analog combining, compression, and antenna selection, without additional hardware. The application of DMAs was shown to yield simple, robust, and low-power systems for microwave imaging [24]–[26], radar systems [27]–[29], and satellite communications [30]. Despite the potential of DMAs in combining signal processing and antenna design, their application for massive MIMO systems has not yet been studied.

In this work we aim to fill this gap by studying large-scale multi-user MIMO networks utilizing DMAs. In particular, we derive the achievable performance on the uplink, namely, when data is transmitted from the user terminals (UTs) to the BS. The application of DMAs yields a physically small large-scale antenna array which inherently implements signal processing techniques such as analog combining, subject to specific constraints induced by the physics of the metasurfaces. Such structures can be used for realizing small size, reduced cost, and spectrally-efficient massive MIMO BSs. Unlike standard analog combining, e.g., [5], [10], DMAs implement adjustable compression without requiring additional hardware.

We propose a mathematical model for DMA-based MIMO systems, accounting for their unique characteristics, such as moderate frequency selectivity and mutual coupling. We focus on the scenario where the frequency selectivity, induced by the physics of the metasurfaces, is identical among all the radiating elements. We characterize the maximal achievable average sum-rate among all UTs in the network, and compare them to the fundamental limits, achievable with unconstrained antenna arrays. We show that the effect of frequency selectivity can be accounted for in the configuration of the DMAs, and thus the fundamental limits can be approached using DMA based antenna arrays. Then, we derive an alternating optimization algorithm for configuring the DMA to approach the optimal performance, accounting for the specific characteristics of the metasurfaces. Our simulations demonstrate that the achievable performance of DMA-based massive MIMO systems is comparable to the theoretical fundamental limits of the channel.

The rest of this paper is organized as follows: Section II introduces the system model. Section III derives the performance limits, and proposes an algorithm for designing DMAs to approach these limits. Section IV provides simulation examples.

Throughout this paper, we use boldface lower-case letters for vectors, e.g., $\mathbf{x}$; the ith element of $\mathbf{x}$ is written as $(\mathbf{x})_i$. Matrices are denoted with boldface upper-case letters, e.g., $\boldsymbol{M}$, $(\boldsymbol{M})_{i,j}$ denotes its (i,j)th element, and $|\boldsymbol{M}|$ is its determinant. We use $\boldsymbol{I}_n$ to denote the $n \times n$ identity matrix. Stochastic expectation and Kronecker delta are denoted by $\mathbb{E}\{\cdot\}$ and $\delta_{\cdot,\cdot}$ respectively. Let $\|\cdot\|$ denote the Euclidean norm when applied to vectors and the Frobenius norm when applied to matrices, $\mathcal{C}$ and $\mathcal{N}$ are the sets of complex numbers and natural

This project has received funding from the Air Force Office of Scientific Research under grant No. FA9550-18-1-0187.

N. Shlezinger and Y. C. Eldar are with the faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel (e-mail: nirshlezinger1@gmail.com; yonina@weizmann.ac.il). O. Dicker is with the department of EE, Technion, Haifa, Israel (e-mail: dicker@technion.ac.il). M. F. Imani and D. R. Smith are with the department of ECE, Duke University, Durham, NC (e-mail: mohamad.imani@gmail.com; drsmith@duke.edu).

numbers, respectively.

II. System Model

A. Dynamic Metasurface Antennas

Metamaterials are a class of artificial materials whose physical properties, and particularly their permittivity and permeability, can be engineered to exhibit a broad set of desired characteristics [31]. Metasurface antennas implement antenna arrays using microstrips consisting of a multitude of sub-wavelength, frequency-selective resonant metamaterial radiating elements [22]. On the receive side, each microstrip feeds a single RF chain, whose digital output is obtained as a linear combination of the radiation observed by each element of the microstrip. To mathematically formulate the input-output relationship of such antennas, consider a metasurface antenna with n_d microstrips, each consisting of l_e elements. Let $\mathbf{y}[i] \in \mathcal{C}^{n_d \cdot l_e \times 1}$ be a vector such that $(\mathbf{y}[i])_{(p-1)\cdot l_e + l}$ is the radiation observed at the lth element of the pth microstrip at time i. The frequency response of the elements is modeled as a finite-memory causal linear filter $\{h[\tau]\}_{\tau=0}^{m_h}$. The output of the metasurface antenna at time i is the vector $\mathbf{z}[i] \in \mathcal{C}^{n_d}$ whose entries can be written as

$$(\mathbf{z}[i])_p = \sum_{l=1}^{l_e} q_{p,l} \sum_{\tau=0}^{m_h} h[\tau] \cdot (\mathbf{y}[i-\tau])_{(p-1)\cdot l_e + l}\,, \tag{1}$$

with $p \in \{1,2,\ldots,n_d\}$. Here, each coefficient $q_{p,l} \in \mathcal{Q}$ represents the physical properties of the lth element of the pth microstrip, and $\mathcal{Q}$ is the set of feasible values of $q_{p,l}$, determined by the physics of the metsurface. Due to the sub-wavelength proximity of the elements in microsrtip, the input vector $\mathbf{y}[i]$ is spatially correlated. The relationship between the multivariate processes $\mathbf{y}[i]$ and $\mathbf{z}[i]$ can be thus written as

$$\mathbf{z}[i] = \boldsymbol{Q} \sum_{\tau=0}^{m_h} h[\tau] \cdot \mathbf{y}[i-\tau], \tag{2}$$

where $\boldsymbol{Q}$ is an $n_d \times (n_d \cdot l_e)$ matrix representing the configurable weights of the DMAs. Using (1), we can write

$$(\boldsymbol{Q})_{p_1,(p_2-1)n_d + l} = q_{p_1,l} \cdot \delta_{p_1 - p_2}. \tag{3}$$

DMAs integrate a tuning mechanism into each independent antenna [28]. The dynamic tunning adds the flexibility to adjust the properties of the metamaterial elements, namely, to control the values of the coefficients $\{q_{p,l}\}$ in (1). The set of possible values of $\{q_{p,l}\}$, denoted $\mathcal{Q}$, represents the Lorentzian resonance response [23], and typically consists of a subset of the complex plain $\mathcal{C}$ of either of the following forms [23, Sec. III]:

- *Amplitude only*: $\mathcal{Q} = [a,b]$ for some $0 < a < b$.
- *Binary amplitude*: $\mathcal{Q} = c \cdot \{0,1\}$ for some fixed $c > 0$.
- *Lorentzian-constrained phase*: $\mathcal{Q} = \{q = \frac{j+e^{j\phi}}{2} : \phi \in [0, 2\pi]\}$.

In order to quantify the gains of utilizing DMAs, we recall that in standard antenna arrays, the digital output is the observed vector $\mathbf{y}[i]$. Clearly, any performance achievable with DMAs is also achievable with standard antenna arrays, as $\mathbf{z}[i]$ can be obtained from $\mathbf{y}[i]$, but not vice versa. However, standard antennas require each of the $n_d \cdot l_e$ to be connected to an RF chain, while DMAs require a single RF chain per microstrip. RF chain hardware tends to be costly [6], is typically a dominant source of power consumption [7], and require memory for storing their digital output [32]. Consequently, by utilizing DMAs, the resulting cost, memory usage, and power consumption, are reduced by a factor of l_e compared to standard antenna arrays. Additionally, unlike standard antenna arrays, DMAs can realize large-scale planar arrays of tunable antenna elements [9].

We note that reducing the number of RF chains can also be carried out with standard antenna arrays using dedicated analog combining hardware, see, e.g., [5], [6], [10], [32]. However, in the presence of standard antenna arrays, analog combining comes at the

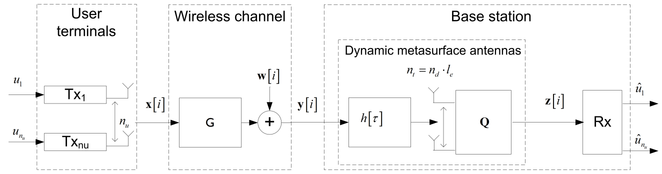

Fig. 1. System model illustration.

cost of additional hardware, increasing the overall size and cost. DMAs inherently implement adjustable analog combining in the physical structure of the metasurfaces, without additional hardware. Furthermore, standard analog combining implemented using dedicated hardware is also typically subject to constraints. In particular, while in DMAs the weights matrix $\boldsymbol{Q}$ must obey the structure in (3) and its entries must be in $\mathcal{Q}$, standard analog combiners must satisfy the architecture-based constraints detailed in [6, Sec. II]. Finally, we note that when $l_e = 1$, $\boldsymbol{Q} = \boldsymbol{I}_{n_d}$, and $h[\tau]$ is a Kronecker delta function, then $\mathbf{z}[i] \equiv \mathbf{y}[i]$, and the resulting DMA coincides with the standard array. However, this implementation requires the same amount of RF chains as standard arrays, and does not result in any gains in cost, power consumption, and memory requirement.

B. Problem Formulation

We consider a single-cell multi-user uplink MIMO system. The BS is equipped with a DMA, consisting of n_d microstrips, each with l_e elements, thus the overall number of elements is $n_t \triangleq n_d \cdot l_e$. The number of UTs served by the BS is $n_u \leq n_t$.

Let $\boldsymbol{G} \in \mathcal{C}^{n_t \times n_u}$ represent the channel matrix from the UTs to the BS, and $\mathbf{w}[i] \in \mathcal{C}^{n_t}$ be an i.i.d. zero-mean proper-complex Gaussian signal with covariance matrix $\boldsymbol{C}_W$, representing the additive channel noise at the BS. By letting $\mathbf{x}[i] \in \mathcal{C}^{n_u}$ be the transmitted signal of the UTs at time index i, the corresponding channel output at the BS is given by

$$\mathbf{y}[i] = \boldsymbol{G}\mathbf{x}[i] + \mathbf{w}[i]. \tag{4}$$

We assume that the UTs utilize Gaussian codebooks, i.e., $\mathbf{x}[i]$ is a zero-mean Gaussian vector with identity covariance matrix, and that the BS knows the realization of $\boldsymbol{G}$.

At the BS, the DMA converts the received $\mathbf{y}[i] \in \mathcal{C}^{n_t}$ into the vector $\mathbf{z}[i] \in \mathcal{C}^{n_d}$, used to decode the message. The relationship between $\mathbf{y}[i]$ and $\mathbf{z}[i]$ is given by (2). An illustration of this system is given in Fig. 1. Due to the moderate frequency variations exhibited by metasurfaces, its frequency selectivity, modeled via $h[\tau]$, cannot be ignored by assuming a narrowband model, as in the standard massive MIMO literature, e.g., [3].

In order to compare the performance achievable with DMA to optimal unconstrained antenna arrays with n_t RF chains, we also consider the case where the BS decodes the transmitted signals based on the channel output $\mathbf{y}[i]$, instead of $\mathbf{z}[i]$. This scenario is referred to henceforth as *optimal MIMO*. Recall that the maximal SE of optimal MIMO is not smaller than that achievable with DMAs, as the output of the DMA $\mathbf{z}[i]$ can be obtained from $\mathbf{y}[i]$. To guarantee fair comparison, the antenna spacing in the optimal MIMO setup is identical to that used with DMAs, thus the resulting wireless channel, i.e., the relationship between $\mathbf{x}[i]$ and $\mathbf{y}[i]$, is the same as in the DMA setup.

Our goal is to characterize the SE, given by the achievable average sum-rate, for the considered system with DMAs compared to the optimal MIMO case, and to provide guidelines for configuring DMAs such that performance is optimized.

III. Achievable Average Sum-Rates

In the following we study the achievable average sum-rate and the resulting DMA configuration for the setup presented in Section II. The maximal achievable average sum-rate for a fixed DMA weights matrix $\boldsymbol{Q}$ is stated in the following theorem:

Theorem 1. *The maximal achievable average sum-rate of the channel in (4) and (2) for a fixed weight matrix $\boldsymbol{Q}$ is given by*

$$R_{\rm s} = \frac{1}{n_u} \log \left| \boldsymbol{I}_{n_d} + \boldsymbol{Q}\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{Q}^H \left(\boldsymbol{Q}\boldsymbol{C}_W\boldsymbol{Q}^H \right)^{-1} \right|. \tag{5}$$

Proof: The theorem is obtained by incorporating (2) as part of the channel. The achievable sum-rate of the resulting finite-memory multiple-access channel is given in [34]. ■

Theorem 1 gives rise to the fundamental performance limits of the wireless channel, as stated in the following corollary:

Corollary 1. *Define $\tilde{\boldsymbol{G}} \triangleq \boldsymbol{C}_W^{-1/2}\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{C}_W^{-1/2}$, and let $\{\lambda_i\}_{i=1}^{n_u}$ be its eigenvalues arranged in descending order. The achievable average sum-rate of the optimal MIMO setup is given by*

$$R_{\rm s}^{\rm OM} = \frac{1}{n_u} \sum_{i=1}^{n_u} \log\left(1 + \lambda_i\right). \tag{6}$$

Proof: As noted in Section II, when $l_e = 1$, $\boldsymbol{Q} = \boldsymbol{I}_{n_d}$, and $\boldsymbol{\Gamma}(\omega) \equiv \boldsymbol{I}_{n_d}$, the resulting setup coincides with the optimal MIMO setup. Substituting this into (5) proves (6). ■

When DMAs are utilized, we note that due to the structure constraints on $\boldsymbol{Q}$, it is difficult to determine the DMA weights matrix $\boldsymbol{Q}$ such that (5) is maximized. Therefore, in order to design $\boldsymbol{Q}$ and obtain the resulting $R_{\rm s}$, we first derive the choice of $\boldsymbol{Q}$ which maximizes the achievable sum-rate, ignoring the structure constraints detailed in Section II. Then, we propose an iterative algorithm for configuring practical constrained DMAs.

A. Optimal Weights

To find $\boldsymbol{Q}$ which maximizes (5), we formulate the dependence of $R_{\rm s}$ on $\boldsymbol{Q}$ in the following lemma:

Lemma 1. *Define $\tilde{\boldsymbol{Q}} \triangleq \boldsymbol{Q}\boldsymbol{C}_W^{1/2}$ and let $\boldsymbol{V}$ be its right singular vectors matrix. By letting $\tilde{\boldsymbol{V}}$ be the $n_t \times n_d$ matrix consisting of the first n_d columns of the unitary matrix $\boldsymbol{V}$, the achievable sum-rate in (5) can be written as*

$$R_{\rm s} = \frac{1}{n_u} \log \left| \boldsymbol{I}_{n_d} + \tilde{\boldsymbol{V}}^H\tilde{\boldsymbol{G}}\tilde{\boldsymbol{V}} \right|. \tag{7}$$

Proof: By replacing $\boldsymbol{Q}$ in (5) with $\tilde{\boldsymbol{Q}} = \boldsymbol{Q}\boldsymbol{C}_W^{1/2}$ it follows from Sylvester's determinant theorem [37, Ch. 6.2] that

$$R_{\rm s} = \frac{1}{n_u} \log \left| \boldsymbol{I}_{n_t} + \tilde{\boldsymbol{G}} \left(\tilde{\boldsymbol{Q}}^H \left(\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^H \right)^{-1} \tilde{\boldsymbol{Q}} \right) \right|. \tag{8}$$

Next, we note that $\tilde{\boldsymbol{Q}}^H(\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^H)^{-1}\tilde{\boldsymbol{Q}}$ is a projection matrix, and can be written as $\tilde{\boldsymbol{Q}}^H(\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^H)^{-1}\tilde{\boldsymbol{Q}} = \tilde{\boldsymbol{V}}\tilde{\boldsymbol{V}}^H$ [37, Ch. 5.9]. Substituting this into (8) proves (7). ■

Lemma 1 implies that the SE depends on the weights matrix $\boldsymbol{Q}$ *only through the first n_d right eigenvectors* of $\tilde{\boldsymbol{Q}} = \boldsymbol{Q}\boldsymbol{C}_W^{1/2}$. If we ignore the structure constraints of $\boldsymbol{Q}$, then the maximal achievable sum-rate and the corresponding choice of $\tilde{\boldsymbol{V}}$ which maximizes (8) are given in the following corollary:

Corollary 2. *The maximal achievable average sum-rate when $\boldsymbol{Q}$ can be any complex matrix is given by*

$$R_{\rm s}^{\rm OD} = \frac{1}{n_u} \sum_{i=1}^{\min(n_d,n_u)} \log(1 + \lambda_i), \tag{9}$$

and is achieved by setting the columns of $\tilde{\boldsymbol{V}}$ to be the eigenvectors corresponding to $\{\lambda_i\}_{i=1}^{n_u}$.

Proof: The corollary follows directly from (8). ■

The number of non-zero eigenvalues of $\tilde{\boldsymbol{G}}$ is given by its rank, denoted n_g, which is at most n_u. It thus follows from (9) that increasing the number of microstrips n_d to be larger than n_g has no effect on the optimal sum-rate $R_{\rm s}^{\rm OD}$. In particular, comparing (9) to the fundamental limits in (6), we note that when $n_d \geq n_g$, then $R_{\rm s}^{\rm OD}$ achieves the fundamental limits $R_{\rm s}^{\rm OM}$. However, as each microstrip requires a single RF chain, increasing n_d implicitly increases the cost, power usage, and memory requirements of the resulting system. Furthermore, by letting $\tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}}\tilde{\boldsymbol{V}}^H$ be the compact singular valued decomposition (SVD) of the optimal $\tilde{\boldsymbol{Q}}$, it follows from Corollary 2 that the weights matrix which maximizes (5) can be written as

$$\boldsymbol{Q}^{\rm OD} = \tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}}\tilde{\boldsymbol{V}}^H\boldsymbol{C}_W^{-1/2}. \tag{10}$$

The matrix in (10) is optimal for any unitary $n_d \times n_d$ matrix $\tilde{\boldsymbol{U}}$ and diagonal $n_d \times n_d$ matrix $\tilde{\boldsymbol{D}}$ with positive diagonal entries.

B. Practical DMA Design

The derivation of Corollary 2 ignores the structure constraints on $\boldsymbol{Q}$, and assumes that the right eigenvectors matrix $\tilde{\boldsymbol{V}}$ can be any set of unitary vectors. Nonetheless, as detailed in the problem formulation, $\boldsymbol{Q}$ must be written as in (3), and its coefficients $\{q_{i,l}\}$ should belong to the feasible set $\mathcal{Q}$. Since finding the constrained matrix $\boldsymbol{Q}$ which maximizes (5) is a difficult task, we propose to set $\boldsymbol{Q}$ to be the closest feasible matrix to the unconstrained $\boldsymbol{Q}^{\rm OD}$ in the sense of minimal Frobenious norm. Here, as in [10], [33], we exploit the invariance of $R_{\rm s}^{\rm OD}$ to the selection of the left singular matrix $\tilde{\boldsymbol{U}}$ and the diagonal matrix $\tilde{\boldsymbol{D}}$, and set these matrices such that the Frobenious distance to the feasible approximation is minimized. To formulate the problem, we let $\mathcal{Q}^{n_d \times n_t}$ be the set of $n_d \times n_t$ which can be written as in (3) and whose non-zero entries belong to the feasible set $\mathcal{Q}$. Let $\mathcal{U}^{n_d}$ and $\mathcal{D}^{n_d}$ denote the sets of $n_d \times n_d$ unitary matrices and diagonal matrices with positive diagonal entries larger than some $\epsilon > 0$, respectively. We set the weights matrix $\boldsymbol{Q}$ to be the solution to:

$$\min_{\boldsymbol{Q}\in\mathcal{Q}^{n_d\times n_t}, \tilde{\boldsymbol{U}}\in\mathcal{U}^{n_d}, \tilde{\boldsymbol{D}}\in\mathcal{D}^{n_d}} \left\| \boldsymbol{Q} - \tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}}\tilde{\boldsymbol{V}}^H\boldsymbol{C}_W^{-1/2} \right\|^2. \tag{11}$$

Let $P_{\mathcal{Q}} : \mathcal{C}^{n_d\times n_t} \mapsto \mathcal{Q}^{n_d\times n_t}$ be the entry-wise projection into $\mathcal{Q}^{n_d\times n_t}$. By (3) for any $\boldsymbol{M} \in \mathcal{C}^{n_d\times n_t}$, $P_{\mathcal{Q}}(\boldsymbol{M})$ is given by

$$\left(P_{\mathcal{Q}}(\boldsymbol{M})\right)_{p_1,(p_2-1)n_d+l} = \mathop{\arg\min}_{q\in\mathcal{Q}} \left| q - (\boldsymbol{M})_{p_1,(p_2-1)n_d+l} \right|^2 \cdot \delta_{p_1-p_2}.$$

We propose to solve (11) using alternating minimization, based on the properties detailed in the following lemma:

Lemma 2. *For any $\boldsymbol{M} \in \mathcal{C}^{n_d\times n_t}$ we have that*

$$\boldsymbol{Q}^{\rm AM}(\boldsymbol{M}) \triangleq \mathop{\arg\min}_{\boldsymbol{Q}\in\mathcal{Q}^{n_d\times n_t}} \|\boldsymbol{Q} - \boldsymbol{M}\|^2 = P_{\mathcal{Q}}(\boldsymbol{M}). \tag{12a}$$

Additionally, for any $\boldsymbol{M}_1, \boldsymbol{M}_2 \in \mathcal{C}^{n_d\times n_t}$, let $\boldsymbol{U}_M$ and $\boldsymbol{V}_M$ be the left singular vectors matrix and the right singular vectors matrix of $\boldsymbol{M}_1\boldsymbol{M}_2^H$, then

$$\tilde{\boldsymbol{U}}^{\rm AM}(\boldsymbol{M}_1, \boldsymbol{M}_2) \triangleq \mathop{\arg\min}_{\tilde{\boldsymbol{U}}\in\mathcal{U}^{n_d}} \left\| \boldsymbol{M}_1 - \tilde{\boldsymbol{U}}\boldsymbol{M}_2 \right\|^2 = \boldsymbol{U}_M\boldsymbol{V}_M^H. \tag{12b}$$

Finally, by letting $\mathbf{m}_{1,i}$ and $\mathbf{m}_{2,i}$ be the ith column of $\boldsymbol{M}_1^H$ and $\boldsymbol{M}_2^H$, respectively, it holds that the diagonal entries of

$$\tilde{\boldsymbol{D}}^{\rm AM}(\boldsymbol{M}_1, \boldsymbol{M}_2) \triangleq \mathop{\arg\min}_{\tilde{\boldsymbol{D}}\in\mathcal{D}^{n_d}} \left\| \boldsymbol{M}_1 - \tilde{\boldsymbol{D}}\boldsymbol{M}_2 \right\|^2, \tag{12c}$$

are given by

$$\left(\tilde{\boldsymbol{D}}^{\rm AM}(\boldsymbol{M}_1, \boldsymbol{M}_2)\right)_{i,i} = \max\left(\frac{{\rm Re}\left(\mathbf{m}_{1,i}^H\mathbf{m}_{2,i}\right)}{\left\|\mathbf{m}_{2,i}\right\|^2}, \epsilon \right). \tag{12d}$$

Proof: The equality (12a) follows from the definition of the Frobenius norm. Since the feasible set $\mathcal{Q}^{n_d\times n_t}$ is defined entry-wise, the norm is minimized by entry-wise projection.

Similarly, the minimizing matrix in (12c)-(12d) holds as

$$\left\| \boldsymbol{M}_1 - \tilde{\boldsymbol{D}}\boldsymbol{M}_2 \right\|^2 = \sum_{i=1}^{n_d} \left\| \mathbf{m}_{1,i} - (\tilde{\boldsymbol{D}})_{i,i} \cdot \mathbf{m}_{2,i} \right\|^2$$

$$\stackrel{(a)}{=} \sum_{i=1}^{n_d} \left\|\mathbf{m}_{1,i}\right\|^2 - 2{\rm Re}\left(\mathbf{m}_{1,i}^H\mathbf{m}_{2,i}\right) \cdot (\tilde{\boldsymbol{D}})_{i,i} + \left\|\mathbf{m}_{2,i}\right\|^2 \cdot (\tilde{\boldsymbol{D}})_{i,i}^2,$$

where (a) follows since $(\tilde{\boldsymbol{D}})_{i,i}$ is real-valued. Consequently, the optimal setting of $(\tilde{\boldsymbol{D}})_{i,i} \geq \epsilon$ is given by (12d).

Finally, the unitary matrix in (12b) is obtained from the unitary Procrustes problem [36, Ch. 7.4]. ■

Based on Lemma 2, we propose solving the joint optimization problem (11) in an alternating fashion, i.e., optimize over $\boldsymbol{Q}$ for fixed $\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{D}}$, next optimize over $\tilde{\boldsymbol{U}}$ for fixed $\boldsymbol{Q}, \tilde{\boldsymbol{D}}$, then optimize over $\tilde{\boldsymbol{D}}$ for fixed $\boldsymbol{Q}, \tilde{\boldsymbol{U}}$, and continue until convergence. The resulting alternating minimization algorithm is summarized in Algorithm 1. As the objective in (11) is differentiable, convergence of alternating optimization is guaranteed [38, Thm. 2].

Algorithm 1 DMA weights configuration

1: Initialization: Set $k = 0$ and $\tilde{\boldsymbol{U}}_k = \boldsymbol{I}_{n_d}$, $\tilde{\boldsymbol{D}}_k = \boldsymbol{I}_{n_d}$.
2: Compute $\tilde{\boldsymbol{V}}$ using Corollary 2.
3: Set $\boldsymbol{Q}_{k+1} = \boldsymbol{Q}^{\rm AM}$ with $\boldsymbol{M} = \tilde{\boldsymbol{U}}_k \tilde{\boldsymbol{D}}_k \tilde{\boldsymbol{V}}^H \boldsymbol{C}_W^{-1/2}$ via (12a).
4: Set $\tilde{\boldsymbol{U}}_{k+1} = \tilde{\boldsymbol{U}}^{\rm AM}$ via (12b) with $\boldsymbol{M}_1 = \boldsymbol{Q}_{k+1}$ and $\boldsymbol{M}_2 = \tilde{\boldsymbol{D}}_k \tilde{\boldsymbol{V}}^H \boldsymbol{C}_W^{-1/2}$.
5: Set $\tilde{\boldsymbol{D}}_{k+1} = \tilde{\boldsymbol{D}}^{\rm AM}$ via (12c)-(12d) with $\boldsymbol{M}_1 = \tilde{\boldsymbol{U}}_{k+1}^H \boldsymbol{Q}_{k+1}$ and $\boldsymbol{M}_2 = \tilde{\boldsymbol{V}}^H \boldsymbol{C}_W^{-1/2}$.
6: If termination is inactive: Set $k := k + 1$ and go to Step 3.

In Algorithm 1 we exploit the fact that the optimal unconstrained $\boldsymbol{Q}^{\rm OD}$ achieves the same sum-rate for any setting of $\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{D}}$, and use these matrices as optimization variables. Consequently, we obtain feasible weight matrices which are within a small distance from the optimal unconstrained matrix. In Section IV we numerically demonstrate that BSs equipped with DMAs designed via Algorithm 1 are capable of achieving performance within a small gap of the fundamental limits of the channel, achievable using optimal impractical antenna arrays.

IV. Numerical Study

In this section we numerically evaluate the achievable performance using the proposed DMA configurations. We consider an uplink multi-user MIMO cell, based on the model used in [2]. Here, a BS equipped with a DMA serves $n_u = 10$ UTs, uniformly distributed in a hexagonal cell of radius 400 m. We use ρ_i to denote the distance of the ith UT from the BS. The channel matrix $\boldsymbol{G}$ is generated as $\boldsymbol{G} = \boldsymbol{H}\boldsymbol{D}$, where $\boldsymbol{H} \in \mathcal{C}^{n_t \times n_u}$ is a random proper-complex zero-mean Gaussian matrix with i.i.d. entires of unit variance, and $\boldsymbol{D} \in \mathcal{C}^{n_u \times n_u}$ is a diagonal matrix whose entires are $(\boldsymbol{D})_{i,i} = \frac{\zeta_i}{\rho_i^2}$, where $\{\zeta_i\}$ are the shadow fading coefficients, independently randomized from a log-normal distribution with standard deviation of 8 dB.

As the DMA elements are sub-wavelength separated, the noise $\mathbf{w}[i]$ is spatially correlated. We use an exponentially decaying spatial correlation profile, namely, $(\boldsymbol{C}_W)_{i,l} = \sigma_W^2 \cdot e^{-|i-l|}$, $i, l \in \{1, \ldots, n_t\}$. We consider the following setups:

- $R_{\rm s}^{\rm UC}$ - unconstrained weights, i.e., $\mathcal{Q} = \mathcal{C}$.
- $R_{\rm s}^{\rm AO}$ - amplitude only weights with $a = 0.001$, $b = 5$.
- $R_{\rm s}^{\rm BA}$ - binary amplitude weights with $c = 0.1$.
- $R_{\rm s}^{\rm LP}$ - Lorentzian-constrained phase.

In Fig. 2 we let the signal-to-noise ratio (SNR), defined as $1/\sigma_W^2$, vary in the range $[-5, 30]$ dB. For each SNR value we compare the performance achievable using DMAs with $n_d = 10$ microstrips, each with $l_e = 10$ elements, computed via Algorithm 1, to the optimal $R_{\rm s}^{\rm OD}$ computed via Corollary 2. Since $n_d \geq n_u$, then $R_{\rm s}^{\rm OD}$ equals the fundamental limit, $R_{\rm s}^{\rm OM}$. Observing Fig. 2, we note that for SNRs above 10 dB, $R_{\rm s}^{\rm LP}$ coincides with $R_{\rm s}^{\rm UC}$, indicating that the Lorentzian-constrained phase restriction induces negligible loss. The amplitude only restriction and the binary amplitude constraint achieve roughly the same performance, which is within a small gap of $R_{\rm s}^{\rm LP}$.

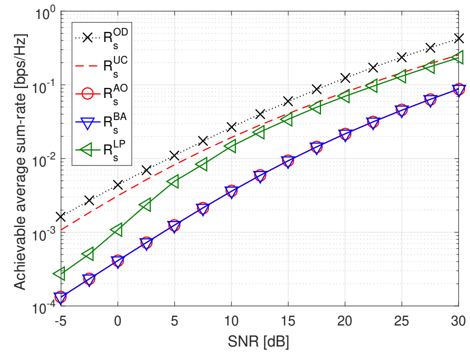

Fig. 2. Rate vs. SNR.

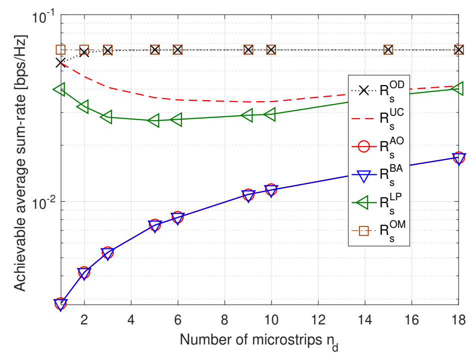

Fig. 3. Rate vs. number of microstrips.

Furthermore, the SNR loss induced by restricting the weights matrix to satisfy (3) is approximately 2.5 dB.

Next, in Fig. 3, we fix the SNR to 15 dB, the number of antennas to $n_t = 90$, and compute the SEs for $n_d \in [1, 18]$. Observing Fig. 3, we again note that $R_{\rm s}^{\rm LP}$ approaches the performance achieved with unconstrained weights, $R_{\rm s}^{\rm UC}$, for most considered values of n_d. As expected, for $n_d = 1$, $R_{\rm s}^{\rm UC}$, which is subject only to (3), coincides with the optimal $R_{\rm s}^{\rm OD}$, as (3) imposes no constraint on $\boldsymbol{Q}$ for $n_d = 1$. Next, we note that $R_{\rm s}^{\rm OD}$ is monotonically increasing for $n_d \leq 3$, and for $n_d > 3$ it remains constant and equals the fundamental limit of the channel, $R_{\rm s}^{\rm OM}$. This is in agreement with the discussion following Corollary 2, as for the considered scenario, most realizations of $\tilde{\boldsymbol{G}}$ have at most 3 dominant eigenvalues, thus $R_{\rm s}^{\rm OD}$ remains constant for $n_d > 3$. Since the constraint induced on $R_{\rm s}^{\rm UC}$ in (3) becomes less significant as n_d decreases, it is shown in Fig. 3 that, for a fixed number of elements n_t, both $R_{\rm s}^{\rm UC}$ and $R_{\rm s}^{\rm LP}$ do not monotonically increase with the number of microstrips n_d. Finally, we observe in Fig. 3 that, while $R_{\rm s}^{\rm OD}$ remains constant as the number of microstrips n_d increases above n_u, the performance achievable with DMAs is monotonically increasing. This follows since increasing the number of microstrips n_d allows designing the matrices $\tilde{\boldsymbol{U}}$ and $\tilde{\boldsymbol{D}}$ in (10), which have no effect on $R_{\rm s}^{\rm OD}$, such that the $\boldsymbol{Q}^{\rm OD}$ can be better approximated using a feasible weights matrix.

V. Conclusions

In this work we studied uplink massive MIMO systems where the large-scale antenna array at the BS is implemented using a DMA. We characterized the SE, and derived an alternating optimization algorithm for designing practical DMAs to approach the optimal performance. Our results illustrate that by properly adjusting the inherent combining and compression induced by the physics of DMAs, a practical massive MIMO system can be constructed which is capable of achieving performance comparable to the fundamental limits, obtained with optimal costly, power consuming, and large-sized antenna arrays.

References

[1] T. L. Marzetta. "Massive MIMO: An introduction". *Bell Labs Technical Journal*, vol. 20, Mar. 2015, pp. 11–22.

[2] T. L. Marzetta. "Noncooperative cellular wireless with unlimited numbers of base station antenna". *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, Nov. 2010, pp. 3950–3600.

[3] J. Hoydis, S. Ten Brink, and M. Debbah. "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?". *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, Feb. 2013, pp. 160–171.

[4] N. Shlezinger and Y. C. Eldar. "On the spectral efficiency of noncooperative uplink massive MIMO systems". *arXiv preprint*, arXiv:1712.03485, 2017.

[5] A. AlKhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath. "MIMO precoding and combining solutions for millimeter-wave systems". *IEEE Comm. Mag.*, vol. 52, no. 12, Dec. 2014, pp. 122-131.

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

[7] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath. "Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs". *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, Apr. 2017, pp. 2274-2287.

[8] A. Hoeher and N. Doose. "A massive MIMO terminal concept based on small-size multi-mode antennas". *Trans. Emerging Tel. Tech.*, vol. 28, no. 2, Feb. 2017.

[9] I. F. Akyildiz and J. M. Jornet. "Realizing Ultra-Massive MIMO (1024×1024) communication in the (0.06 - 10) Terahertz band". *Nano Commun. Networks*, vol. 8, 2016, pp. 46-54.

[10] S. Stein and Y. C. Eldar. "Hybrid analog-digital beamforming for massive MIMO systems". *arXiv preprint*, arXiv:1712.03485, 2017

[11] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu. "Channel estimation and performance analysis of one-bit massive MIMO systems". *IEEE Trans. Signal Process.*, vol. 65, no. 15, Aug. 2017, pp. 4075-4089.

[12] J. Choi, J. Mo, and R. W. Heath. "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs". *IEEE Trans. Commun.*, vol. 64, no. 5, May 2016, pp. 2005-2018.

[13] J. Choi, B. L. Evans, and A. Gatherer. "Resolution-adaptive hybrid MIMO architectures for millimeter wave communications". *IEEE Trans. Signal Process.*, vol. 65, no. 23, Dec. 2017, pp. 6201-6216.

[14] T. C. Zhang, C. K. Wen, S. Jin, and T. Jiang. "Mixed-ADC massive MIMO detectors: Performance analysis and design optimization". *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, Nov. 2016, pp. 7738-7752.

[15] J. Choi, J. Sung, B. L. Evans, and A. Gatherer. "Antenna selection for large-scale MIMO systems with low-resolution ADCs". *Proc. IEEE ICASSP*, Calgary, Canada, Apr. 2018.

[16] S. K. Mohammed and E. G. Larsson. "Per-antenna constant envelope precoding for large multi-user MIMO systems". *IEEE Trans. Commun.*, vol. 61, no. 3, Mar. 2013, pp. 1059-1071.

[17] C. Studer and E. G. Larsson. "PAR-aware large-scale multi-user MIMO-OFDM downlink". *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, Feb. 2013, pp. 303-313.

[18] S. Ghosh, T. N. Tran, and T. Le-Ngoc. "Dual-layer EBG-based miniaturized multi-element antenna for MIMO systems". *IEEE Trans. Antennas Propag.*, vol. 62, no. 8, Aug. 2014, pp. 3985-3997.

[19] C. L. Holloway, E. F. Kuester, J. A. Gordon, J. O'Hara, J. Booth, and D. R. Smith. "An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials". *IEEE Antennas Propag. Mag.*, vol. 54, no. 2, Apr. 2012, pp. 10 - 35.

[20] P. Mookiah and K. P. Dandekar. "Metamaterial-substrate antenna array for MIMO communication system". *IEEE Trans. Antennas Propag.*, vol. 57, no. 10, Oct. 2009, pp. 3283-3292.

[21] R. O. Ouedraogo, E. J. Rothwell, A. R. Diaz, K. Fuchi, and A. Temme. "Miniaturization of patch antennas using a metamaterial-inspired technique". *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, May 2012, pp. 2175-2182.

[22] J. Hunt, T. Driscoll, A. Mrozack, G. Lipworth, M. Reynolds, D. Brady, and D. R. Smith. "Metamaterial apertures for computational imaging". *Science*, vol. 339, no. 6117, Jan. 2013, pp. 310-313.

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

[24] A. V. Diebold, M. F. Imani, T. Sleasman, and D. R. Smith. "Phaseless computational ghost imaging at microwave frequencies using a dynamic metasurface aperture". *Applied Optics*, vol. 57, no. 9, 2018, pp. 2142 - 2149.

[25] T. Sleasman, M. Boyarsky, M. F. Imani, J. N. Gollub, and D. R. Smith. "Design considerations for a dynamic metamaterial aperture for computational imaging at microwave frequencies". *J. Opt. Soc. Am. B*, vol. 33, no. 6, 2016, pp. 1098 - 1111.

[26] T. Sleasman, M. Boyarsky, M. F. Imani, T. Fromenteze, J. N. Gollub, and D. R. Smith. "Single-frequency microwave imaging with dynamic metasurface apertures". *J. Opt. Soc. Am. B*, vol. 34, no. 8, 2017, pp. 1713 - 1736.

[27] C. M. Watts, A. Pedross-Engel, D. R. Smith, and M. S. Reynolds. "X-band SAR imaging with a liquid-crystal-based dynamic metasurface antenna". *J. Opt. Soc. Am. B*, vol. 34, no.28, 2017, pp. 300 - 306.

[28] T. Sleasman, M. Boyarsky, L. Pulido-Mancera, T. Fromenteze, M. F. Imani, M. S. Reynolds, and D. R. Smith. "Experimental synthetic aperture radar with dynamic metasurfaces". *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, Dec. 2017, pp. 6864 - 6877.

[29] M. Boyarsky, T. Sleasman, L. Pulido-Mancera, T. Fromenteze, A. Pedross-Engel, C. M. Watts, M. F. Imani, M. S. Reynolds, and D. R. Smith. "Synthetic aperture radar with dynamic metasurface antennas: a conceptual development". *Journal of the Optical Society of America A*, vol. 34, no. 5, 2017, pp. A22 - A36.

[30] M. C. Johnson, S. L. Brunton, N. B. Kundtz, and J. N. Kutz. "Sidelobe canceling for reconfigurable holographic metamaterial antenna". *IEEE Trans. Antennas Propag.*, vol. 63, no. 4, Apr. 2015, pp. 1881-1886.

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

[32] N. Shlezinger, Y. C. Eldar, and M. R. D. Rodrigues. "Hardware-limited task-based quantization". *arXiv preprint*, arXiv:1807.08305, 2018.

[33] N. Shlezinger, R. Dabora, and Y. C. Eldar. "Measurement matrix design for phase retrieval based on mutual information". *IEEE Trans. Signal Process.*, vol. 66, no. 2, Jan. 2018, pp. 324-339.

[34] S. Verdu. "Multiple-access channels with memory with and without frame synchronism," *IEEE Trans. on Inform. Theory*, vol. 35, no. 3, May 1989, pp. 605–619.

[35] D. N. C. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge, 2005.

[36] R. A. Horn and C. A. Johnson. *Matrix Analysis*. Cambridge, 1990.

[37] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, 2000.

[38] J. C. Bezdek and R. J. Hathaway. "Convergence of alternating optimization". *Neural, Parallel and Scientific Computations*, vol. 11, no. 4, Dec. 2003, pp. 351–368.