Beam Focusing for Near-Field Multiuser MIMO Communications

Haiyang Zhang, *Member, IEEE*, Nir Shlezinger, *Member, IEEE*, Francesco Guidi, *Member, IEEE*, Davide Dardari, *Senior Member, IEEE*, Mohammadreza F. Imani, *Member, IEEE*, and Yonina C. Eldar, *Fellow, IEEE*

***Abstract*— Large antenna arrays and high-frequency bands are two key features of future wireless communication systems. The combination of large-scale antennas with high transmission frequencies often results in the communicating devices operating in the near-field (Fresnel) region. In this paper, we study the potential of beam focusing, feasible in near-field operation, in facilitating high-rate multi-user downlink multiple-input multiple-output (MIMO) systems. As the ability to achieve beam focusing is dictated by the transmit antenna, we study near-field signalling considering different antenna structures, including fully-digital architectures, hybrid phase shifter-based precoders, and the emerging dynamic metasurface antenna (DMA) architecture for massive MIMO arrays. We first provide a mathematical model to characterize near-field wireless channels as well as the transmission pattern for the considered antenna architectures. Then, we formulate the beam focusing problem for the goal of maximizing the achievable sum-rate in multi-user networks. We propose efficient solutions based on the sum-rate maximization task for fully-digital, (phase shifters based-) hybrid and DMA architectures. Simulation results show the feasibility of the proposed beam focusing scheme for both single- and multi-user scenarios. In particular, the designed focused beams provide a new degree of freedom to mitigate interference in both angle and distance domains, which is not achievable using conventional far-field beam steering, allowing reliable communications for uses even residing at the same angular direction.**

***Index Terms*— Beam focusing, dynamic metasurface antennas, near-field multi-user communication.**

Manuscript received 27 May 2021; revised 1 January 2022 and 27 February 2022; accepted 7 March 2022. Date of publication 18 March 2022; date of current version 12 September 2022. This work was supported in part by the European Union's H2020 Research and Innovation Program under Grant 101000967; in part by the Air Force Office of Scientific Research under Grant FA9550-18-1-0208; in part by the Israel Science Foundation under Grant 0100101; in part by the Theory Lab, Central Research Institute, 2012 Labs, Huawei Technologies Co., Ltd.; and in part by the Project Dipartimenti di Eccellenza—DEI, University of Bologna. An earlier version of this paper was in part at the 2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) [DOI: 10.1109/ICASSP39728.2021.9413746]. The associate editor coordinating the review of this article and approving it for publication was W. Ni. *(Corresponding author: Haiyang Zhang.)*

Haiyang Zhang and Yonina C. Eldar are with the Faculty of Math and CS, Weizmann Institute of Science, Rehovot 7610001, Israel (e-mail: haiyang.zhang@weizmann.ac.il; yonina.eldar@weizmann.ac.il).

Nir Shlezinger is with the School of ECE, Ben-Gurion University of the Negev, Beer-Sheva 8499000, Israel (e-mail: nirshl@bgu.ac.il).

Francesco Guidi is with the National Research Council of Italy, Institute of Electronics, Computer and Telecommunication Engineering, 40136 Bologna, Italy (e-mail: francesco.guidi@ieiit.cnr.it).

Davide Dardari is with the Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi" (DEI-CNIT), University of Bologna, Cesena Campus, 47521 Cesena, Italy (e-mail: davide.dardari@unibo.it).

Mohammadreza F. Imani is with the School of ECEE, Arizona State University, Tempe, AZ 85287 USA (e-mail: mohammadreza.imani@asu.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2022.3158894.

Digital Object Identifier 10.1109/TWC.2022.3158894

I. INTRODUCTION

WIRELESS communication over high-frequency millimeter wave (mmWave) and terahertz (THz) spectrum is regarded as a key technology for beyond 5G communications, due to its capability of enhancing data-rates thanks to the large available bandwidth. In order to compensate for the dominant path loss characterizing transmissions in high frequencies, wireless base stations (BSs) operating in these bands will be equipped with electrically large antenna arrays [2]. A byproduct of utilizing large-scale antennas is that high-frequency communication may take place in the near-field (Fresnel) region, as opposed to conventional wireless systems, typically operating in the far-field regime. More specifically, the near-field distance can be several dozens of meters for relatively small antennas/surfaces at mmWave and THz [3]–[5]. This implies that the far-field model, assuming plane wavefronts of the electromagnetic (EM) field rather than spherical ones, no longer holds at practical distances [6]–[8]. Managing the spherical wavefront of the signals is translated into flexible transmit beamforming capabilities. In particular, it brings forth the possibility to generate radiation patterns which focus the beam (*beam focusing*) at a specific location, in contrast to only a specific direction as in far-field conditions via conventional *beam steering*. Beam focusing gives rise to the possibility to support multiple coexisting orthogonal links, even at similar angles [9].

Most existing works on near-field focusing appeared in the antenna theory literature (see, e.g., [10], [11] and the references therein), wherein the EM field in the Fresnel region was characterized and modeled for various antenna implementation technologies. For example, the authors in [10] studied the effect of the planar array's antenna size, inter-element distance, and focal distance on the near-field focusing performance. In [11], the authors developed a multi-focus antenna array to focus signals on multiple focal points in the near-field region. While antenna theory provides tools to achieve beam focusing, how to exploit this ability to facilitate near-field wireless communications[1] is still in its infancy, and only a small set of works have studied near-field focusing from a communication perspective. In [12], the authors considered

[1]Note that in this paper, with "near-field" we refer only to the Fresnel region (also known as the radiative near-field), thus neglecting the reactive near-field region, that entails distances in the order of the wavelength.

a point-to-point short-range MIMO communication system, which consists of two identical transceiver array antennas that face each other with a distance comparable to the size of the antenna aperture. More recently, near-field communications with antennas based on large intelligent surfaces (LISs), whose large aperture gives rise to near-field operation, was explored in [13]–[16]. In particular, [13] considered a single-user scenario, and characterized the path-loss and the available communication modes, while [14] studied a two-user uplink scenario in which the BS is equipped with an LIS. Both [13] and [14] studied ideal antenna architectures, where the transceiver has direct access to the signal observed at each element. In addition, [15], [16] considered near-field communication in reconfigurable intelligent surfaces (RISs)-assisted systems, where RISs act as a passive reflector. Nonetheless, the potential of near-field focusing in facilitating massive MIMO downlink communications with practical antenna technologies has not been thoroughly studied to date.

The ability to achieve focused beams in massive MIMO systems is highly dependent on the signal processing capabilities of the antenna array, which vary between different architectures. The most flexible solution for a given array of radiating elements is the fully-digital architecture, where each antenna element is connected to a dedicated radio frequency (RF) chain. In such architectures, the transceiver is capable of controlling beams at infinitely many directions at the same time, which greatly enhances the spatial flexibility [17]. However, towards the deployment of large-scale arrays in 5G and beyond communication systems, the implementation of a fully-digital architecture becomes extremely challenging due to its increased cost and power consumption. To alleviate this, hybrid analog/digital architectures are commonly considered for massive MIMO communications [18]–[20]. Such hybrid architectures operate with fewer RF chains than antenna elements by combining low-dimensional digital processing and high-dimensional analog precoding, typically implemented using an interconnection of phase-shifters. An alternative emerging technology for efficiently realizing large-scale arrays is based on DMAs. DMAs are a practical implementation of LISs, i.e., they enable programmable control of the transmit/receive beam patterns, which also provide advanced analog signal processing capabilities [21]–[24], and naturally implement RF chain reduction without dedicated analog circuitry. Furthermore, DMAs facilitate densification of the antenna elements which can be exploited to improve the focusing performance. It is emphasized that these previous works on DMAs, e.g., [21], [22], studied their application to facilitate conventional far-field communications, and did not consider near-field communications. The fact that the signal processing capabilities of the aforementioned antenna architectures affect their ability to generate focused beams motivates the study of near-field multi-user communications with these different antennas.

In this paper we study multi-user downlink MIMO systems operating in the near-field region. We focus on the exploration of utilizing various antenna architectures, including fully-digital arrays, phase-shifters based hybrid architectures, and DMAs, to facilitate multi-user communications via near-field signalling. In particular, we aim to quantify the capabilities of massive MIMO architectures in forming focused beams, as well as the effect of such an operation on downlink multi-user systems. To the best of our knowledge, this work represents the first study on the design of focused beams (e.g., beam focusing) as means of optimizing multi-user communication objectives, and utilizing this capability to facilitate simultaneous communication with multiple users.

We begin by formulating a mathematical model for downlink near-field multi-user MIMO systems. Our model incorporates both the digital signal and analog signal processing carried out by the BS, as well as the propagation of the transmitted EM waves in near-field wireless communications. Then, we study beam focusing design in order to maximize the sum-rate under each of the considered antenna architectures; We first consider fully-digital antenna systems, which is the most flexible architecture as it allows independent control of the signal fed to each transmitting element. We then use the obtained fully-digital beam focusing configuration as a baseline for deriving the corresponding setting for phase shifter based hybrid architectures. In particular, we show that the sum-rate maximization problem for such hybrid antennas can be tackled using manifold optimization techniques for sum-rate optimization in far-field communications aided by RISs [25]–[28]. For DMAs, where the analog signal processing capabilities follow the Lorentzian-form response of metamaterial elements [29], we cannot adapt design methods previously proposed for far-field systems as we do for phase shifter based hybrid antennas, and thus we derive a dedicated configuration algorithm. To that aim, we first focus on a single-user case, for which we are able to optimize the DMA configuration. Then, we consider the case of multiple users, where the resulting optimization problem is non-convex, and propose an alternating design algorithm to jointly optimize the DMA configuration and digital precoding.

While parts of our technical derivations build upon design algorithms proposed for far-field communications, we demonstrate that the incorporation of the near-field characteristics results in fundamentally different beam patterns compared to the far-field. In particular, simulation results show that our proposed designs for different types of antenna architectures are all capable of concentrating the transmissions to the desired focal points, illustrating the beam focusing ability of our proposed designs. Furthermore, it is demonstrated that by exploiting the beam focusing capabilities of near field transmissions via the proposed configuration methods, one can reliably simultaneously communicate with multiple users located in the same angular direction with different ranges, which is not achievable using conventional beam steering techniques. Finally, we show that by accounting for the near-field capabilities in transmission, one can achieve notable gains in achievable sum-rate compared to designs assuming conventional far-field operation.

The rest of this paper is organized as follows: Section II presents the near-field channel model, reviews the considered antenna architectures, and formulates the near-field-aware precoding problem. Section III presents efficient algorithms for tuning the beam focus solutions for all considered antenna architectures, while Section IV numerically demonstrates the beam-focusing ability and evaluates its effect on the achievable rates. Finally, Section V concludes the paper.

Throughout the paper, we use boldface lower-case and upper-case letters for vectors and matrices, respectively. Calligraphic letters are used for sets. The ℓ_2 norm, vectorization operator, transpose, conjugation, Hermitian transpose, trace, Kronecker product, Hadamard product, and stochastic expectation are written as $\|\cdot\|$, $\mathrm{Vec}(\cdot)$, $(\cdot)^T$, $(\cdot)^\dagger$, $(\cdot)^H$, $\mathrm{Tr}\,(\cdot)$, $\otimes$, $\circ$, and $\mathbb{E}\{\cdot\}$, respectively. Finally, for any vector $\mathbf{x}$, $(\mathbf{x})_i$ denotes the ith entry of $\mathbf{x}$.

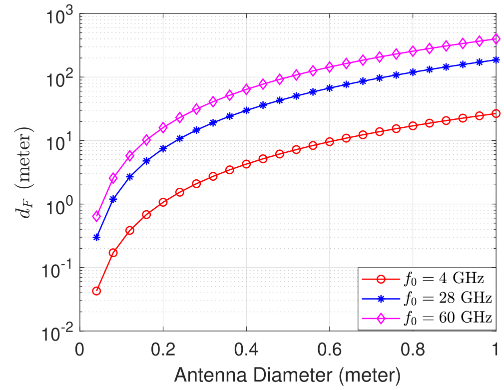
II. System Model

In this section, we describe the considered near-field multi-user MIMO communication system. We first introduce the concept of near-field transmission in Section II-A. Then, we formulate the mathematical model of near-field wireless channels in Section II-B. After that, we present three types of antenna architectures and their corresponding signal models in Section II-C. Finally, in Section II-D, we formulate the optimization problem of designing the transmission beam pattern to maximize the sum-rate for near-field communications.

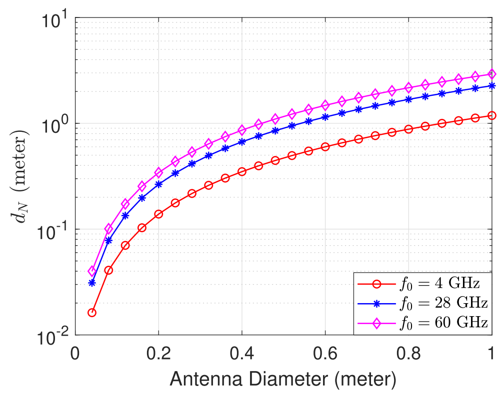
A. Near-Field Region

According to conventional notation, transmission is considered to take place in the far-field if the distance between the transmitter and the receiver is larger than the Fraunhofer distance, denoted by $d_{\rm F} = \frac{2\,D^2}{\lambda}$, where D is the antenna diameter and λ the wavelength. For distances larger than d_F, the signal wavefront can be faithfully approximated as being planar. When the distance is shorter than $d_{\rm F}$ but larger than the Fresnel distance, typically denoted by $d_{\rm N} = \sqrt[3]{\frac{D^4}{8\,\lambda}}$, the receiver is considered to lie in the radiative near-field Fresnel region, referred to henceforth as the near-field region. The near-field accounts for distances in between $d_{\rm F}$ and $d_{\rm N}$. The boundary $d_{\rm N}$ constitutes the minimal distance from which reactive field components from the antenna itself can be neglected.

Conventional wireless communications lie in the far-field due to the entailed distances, antenna sizes and frequencies. For instance, for an antenna of diameter $D = 0.1$ meters at carrier frequency of 5 GHz, any receiver located at a distance of more than $d_F = 0.33$ meters is considered to lie in the far-field. However, for mmWave frequency bands, particularly when combined with antenna arrays of relatively large physical size, this approximation no longer holds, and one must account for the spherical wavefront shape. This is demonstrated in Fig. 1, which illustrates the values of $d_{\rm F}$ (upper limit) and $d_{\rm N}$ (lower limit) for different antenna diameters D and carrier frequencies f_0, that together delimit the expected near-field operating region. From Fig. 1, we can clearly see that when the system operates at mmWave frequency bands, the near-field distance can be up to dozen of meters for relative small antennas/surfaces. For instance, for a BS equipped with an antenna of diameter $D = 0.5$ meters at carrier frequency 28 GHz, any user closer than 47 meters from the antenna resides in its near-field. Therefore, it is of interest to investigate how to exploit the non-negligible spherical wavefronts of the near-field to increase communication rates. In particular, this gives rise to the possibility of generating focused beams and to enhance the communication performance of wireless networks by alleviating multi-user interference.



(a) $d_{\rm F}$ vs. antenna diameter.



(b) $d_{\rm N}$ vs. antenna diameter.

Fig. 1. Near-field region boundaries for different frequencies and antenna diameter values.

B. Near-Field Channel Model

To evaluate the ability to exploit near-field operation in MIMO communications, we focus on downlink multi-user systems. In particular, we consider a downlink multi-user MIMO system where the BS employs a uniform planar array (UPA), i.e., a two-dimensional antenna surface, with N_e uniformly spaced radiating elements in the horizontal direction and N_d elements in the vertical direction. The total number of antenna elements is thus $N = N_d \times N_e$. We denote the Cartesian coordinate of the lth element of the ith row as $\mathbf{p}_{i,l} = (x_l, y_i, 0)$, $l = 1, 2, \ldots N_e$, $i = 1, 2, \ldots N_d$. The BS communicates with M single-antenna receivers, as illustrated in Fig. 2. We consider that the receivers' positioning information is known at the BS via high-accuracy wireless positioning techniques [3]. We focus on communications in the near-field, i.e., where the distance between the BS and the users is not larger than the Fraunhofer distance $d_{\rm F}$ and not smaller than the Fresnel limit $d_{\rm N}$. The properties of near-field spherical waves allow for the generation of focused beams to facilitate communications.

To start, we model the near-field wireless channels following existing modelling techniques for EM propagation in the

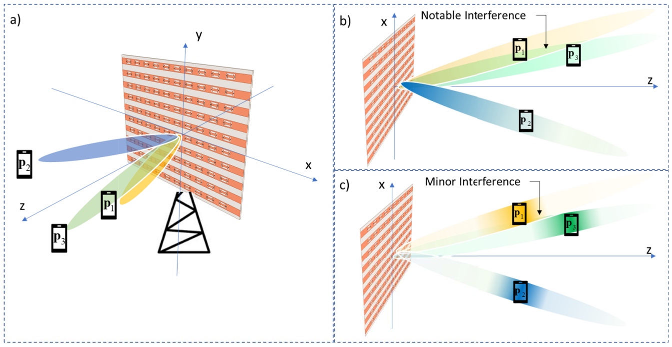


Fig. 2. Near-field communications with $M = 3$ receivers, with dedicated beams directed towards each user: (a) illustration in three-dimensional space; (b) beam steering based on far-field design, resulting in notable interference among users sharing the same angular direction; (c) beam focusing, resulting in minor interference.

radiating near-field, e.g., [30]. The signal received in free-space conditions by the mth user, $m \in \mathcal{M} \triangleq \{1, 2, \ldots, M\}$, located at $\mathbf{p}_m = (x_m, y_m, z_m)$ is given by

$$r(\mathbf{p}_m) = \sum_{i=1}^{N_d} \sum_{l=1}^{N_e} A_{i,l}(\mathbf{p}_m)\, e^{-\jmath k |\mathbf{p}_m - \mathbf{p}_{i,l}|}\, s_{i,l} + n_m, \quad (1)$$

where $s_{i,l}$ denotes the signal emitted by the antenna at position $\mathbf{p}_{i,l}$; the term $e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{i,l}|}$ contains the phase due to the distance travelled by the wave from $\mathbf{p}_{i,l}$ to $\mathbf{p}_m$; $k = 2\pi/\lambda$ is the wave number; $A_{i,l}(\mathbf{p}_m)$ denotes the channel gain coefficient; and $n_m \sim \mathcal{CN}\left(0, \sigma^2\right)$ is the additive white Gaussian noise (AWGN) at user m. Following [30], we write

$$A_{i,l}(\mathbf{p}_m) = \sqrt{F(\Theta_{i,l,m})} \frac{\lambda}{4\,\pi |\mathbf{p}_m - \mathbf{p}_{i,l}|}, \quad (2)$$

where $\Theta_{i,l,m} = (\theta_{i,l,m}, \phi_{i,l,m})$ is the elevation-azimuth pair from the lth element of the ith row to the mth user, while $F(\Theta_{i,l,m})$ is the radiation profile of each element, modeled as

$$F(\Theta_{i,l,m}) = \begin{cases} 2(b+1)\cos^b(\theta_{i,l,m}) & \theta_{i,l,m} \in [0, \pi/2], \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

In (3), the parameter b determines the Boresight gain, whose value depends on the specific technology adopted [30]. As an example, for the dipole case we have $b = 2$, which yields $F(\Theta_{i,l,m}) = 6\, \cos^2 \theta_{i,l,m}$. Here, the model accounts for the fact that the transmitted power is doubled by the reflective ground behind the antenna.

To obtain a more compact formulation of the received signal in (1), we define the vector

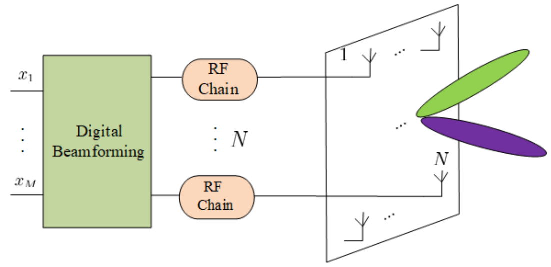
$$\mathbf{a}_m = \left[A_{1,1}(\mathbf{p}_m)\, e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{1,1}|}, A_{1,2}(\mathbf{p}_m)\, e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{1,2}|}, \right.$$
$$\left. \cdots, A_{N_d,N_e}(\mathbf{p}_m) e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{N_d,N_e}|}\right]^H. \quad (4)$$

For convenience, we omit the location index $\mathbf{p}_m$ in $\mathbf{a}_m(\mathbf{p}_m)$ for the rest of this paper. Using (4), we can then write the received signal at the mth user as

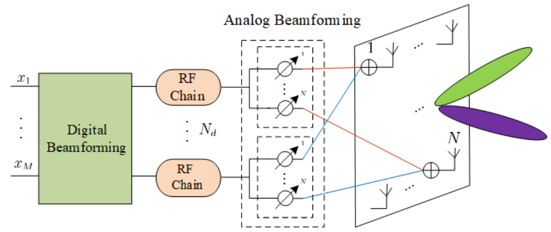
$$r(\mathbf{p}_m) = \mathbf{a}_m^H \mathbf{s} + n_m, \quad m \in \mathcal{M}, \quad (5)$$

where $\mathbf{s} = [s_{1,1}, s_{1,2} \cdots, s_{N_d,N_e}]$ collects the transmitted signals of all antennas.

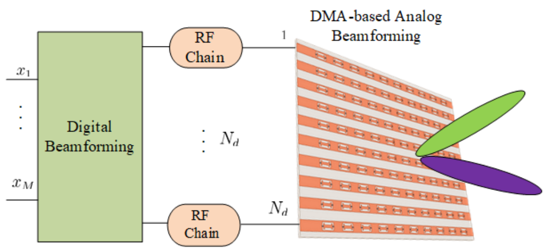
Near-field operation is encapsulated in the vector $\mathbf{a}_m$ defined in (4). When the far-field approximation holds, the outputs of all the elements experience the same path loss (e.g., $A(\mathbf{p}_m) = A_{i,l}(\mathbf{p}_m)\ \forall\, i, l$) and a phase shift with constant gradient along the array aperture, with $\Theta_{i,l,m} = \Theta_m\, \forall\, i, l$. More specifically, $\mathbf{a}_m$ becomes the traditional beamsteering vector given by $\mathbf{a}_m = A(\mathbf{p}_m)\, [e^{-\jmath k \Psi_{1,1}(\Theta_m)}, \ldots, e^{-\jmath k \Psi_{i,l}(\Theta_m)}, \ldots, e^{-\jmath k \Psi_{N_d,N_e}(\Theta_m)}]$, where $\Psi_{i,1}(\Theta_m)$ depends only on the direction of the mth user and on the spacing among the radiating elements. The diversity among the elements of $\mathbf{a}_m$ in the near-field gives rise to the possibility to focus the beam towards an intended position in space, rather than just steer it at a given angle, as enabled in the far-field.



(a) The fully-digital architecture



(b) The phase-shifters based hybrid architecture



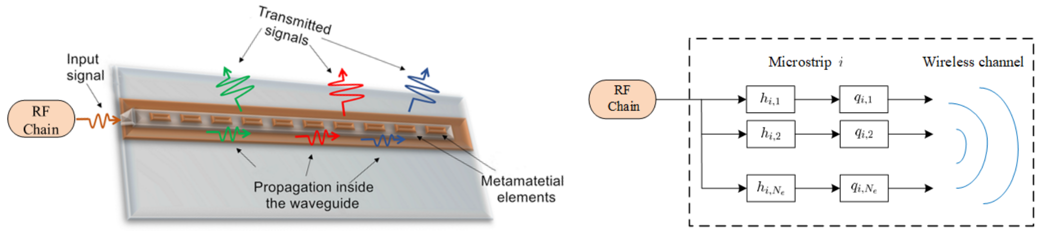
(c) DMA-based architecture.

Fig. 3. Three typical antenna architectures.

C. Antenna Architectures

The beam pattern in (5) depends on the transmitted signal $\mathbf{s}$, which in turn depends on the signal processing capabilities supported by the antenna architecture. We consider three types of antenna schemes as shown in Fig. 3: a fully-digital architecture, phase-shifters based analog precoder, and a DMA. In the following subsections, we will introduce each of the antenna architectures and provide insights on the corresponding signal model of $\mathbf{s}$.

1) Fully-Digital Antenna: In fully-digital antennas each element is connected to a dedicated RF chain, as illustrated in Fig. 3(a). Such architectures provide the most flexible signal processing capabilities, as the input to each element can be separately processed. Nonetheless, fully-digital antennas are typically costly, particularly in massive MIMO systems, since



(a) Physical model (b) Mathematical model

Fig. 4. Illustration of signal transmission using a microstrip. (a) Physical model; (b) Mathematical model.

the number of RF chains is equal to the number of antenna elements [19], [20]. Thus, we consider the fully-digital architecture as a baseline system, representing the beam focusing capabilities achievable in near-field multi-user communications with unconstrained linear precoding. In this case, the signal transmitted by the BS can be written as

$$\mathbf{s} = \sum_{m=1}^{M} \tilde{\mathbf{w}}_m x_m, \tag{6}$$

where x_m is the normalized data symbol intended for the mth user, i.e., $\mathbb{E}[x_m^2] = 1$, and $\tilde{\mathbf{w}}_m \in \mathbb{C}^N$ is the precoding vector for x_m.

By expressing the channel input $\mathbf{s}$ via (6), the received signal of the mth user is given by

$$r(\mathbf{p}_m) = \mathbf{a}_m^H \sum_{j=1}^{M} \tilde{\mathbf{w}}_j x_j + n_m, \quad m \in \mathcal{M}. \tag{7}$$

Based on (7), the achievable rate of the mth user for the fully-digital antenna case is given by

$$R_m\left(\{\tilde{\mathbf{w}}_m\}\right) = \log_2\left(1 + \frac{\left|\mathbf{a}_m^H \tilde{\mathbf{w}}_m\right|^2}{\sum_{j \neq m} \left|\mathbf{a}_m^H \tilde{\mathbf{w}}_j\right|^2 + \sigma^2}\right), \quad m \in \mathcal{M}. \tag{8}$$

Expression (8) characterizes the achievable rate at each user for a given precoding configuration, and is computed assuming that the users treat the interference as noise. While in some scenarios one can achieve higher rates by decoding the interference [31], the common practice in downlink massive MIMO systems is to treat it as noise [32], resulting in the rate in (8).

2) Phase Shifter Based Hybrid Antenna: Hybrid antennas combine digital signal processing with some constrained level of analog signal processing. Here, the number of RF chains, denoted by $N_{\rm RF}$, is smaller than the number of antenna elements N. In fully-connected phase shifting analog precoders, each RF chain output is connected to all the transmit antennas through a phase-shifters based analog beamforming network, as shown in Fig. 3(b). In this case, the signal transmitted by the BS is given by

$$\mathbf{s} = \sum_{m=1}^{M} \mathbf{Q}\, \mathbf{w}_m x_m. \tag{9}$$

Here, the digital precoding vector $\mathbf{w}_m$ is an $N_{\rm RF} \times 1$ vector, while $\mathbf{Q} \in \mathbb{C}^{N \times N_{\rm RF}}$ represents analog precoding, which maps the $N_{\rm RF} \times 1$ digital vector into the N antenna elements. For phase-shifters based analog precoding, the elements of $\mathbf{Q}$, denoted by $\{q_{i,l}\}$, satisfy

$$q_{i,l} \in \mathcal{F} \triangleq \left\{e^{j\phi} | \phi \in [0, 2\pi]\right\}, \quad \forall i, l. \tag{10}$$

Comparing (9) with (6), it holds that the received signal model for the hybrid precoder is a special case of the fully-digital one, obtained by setting $\tilde{\mathbf{w}}_m = \mathbf{Q}\mathbf{w}_m$ for each $m \in \mathcal{M}$. As a result, the achievable rate for a given configuration of the analog precoder $\mathbf{Q}$ and digital vectors $\{\mathbf{w}_j\}$ is given by $R_m(\{\mathbf{Q}\mathbf{w}_j\})$, computed via (8).

3) DMA: DMAs utilize radiating metamaterial elements embedded onto the surface of a waveguide to realize reconfigurable antennas of low cost and power consumption [24]. The typical DMA architecture is comprised of multiple waveguides, e.g. microstrips, and each microstrip contains multiple metamaterial elements. The elements are typically sub-wavelength spaced, implying that one can pack a larger number of elements in a given aperture compared to conventional architectures based on, e.g., patch arrays [24]. The frequency response of each individual element can be externally adjusted by varying its local electrical properties [33].

For DMA-based transmitting architectures, each microstrip is fed by one RF-chain, and the input signal is radiated by all the elements located on the microstrip, as shown in Fig. 3(c). Fig. 4 shows an example of transmitting a signal using a single microstrip with multiple elements. To formulate its input-output relationship, consider a DMA consisting of $N = N_d \cdot N_e$ metamaterial elements, where here N_d and N_e are the numbers of microstrips and elements in each microstrip, respectively. The equivalent baseband signal radiated from the lth element of the ith microstrip is $s_{i,l} = h_{i,l}\, q_{i,l}\, z_i$, where z_i is the baseband signal fed to the ith microstrip, $q_{i,l}$ denotes the tunable response of the lth element of the ith microstrip, and $h_{i,l}$ encapsulates the effect of signal propagation inside the microstrip. We consider the case where the response of the elements is frequency flat as in [21], and focus on the Lorentzian-constrained phase model of the metamaterial elements frequency response [34], i.e.,

$$q_{i,l} \in \mathcal{Q} \triangleq \left\{\frac{j + e^{j\phi}}{2} | \phi \in [0, 2\pi]\right\}, \quad \forall i, l. \tag{11}$$

The signal propagation inside the microstrip is formulated as

$$h_{i,l} = e^{-\rho_{i,l}(\alpha_i + j\beta_i)}, \quad \forall i, l \tag{12}$$

where α_i is the waveguide attenuation coefficient, β_i is the wavenumber, and $\rho_{i,l}$ denotes the location of the lth element in the ith microstrip.

Letting $\mathbf{z} = [z_1, \ldots, z_{N_d}]^T$ be the microstrips input, the baseband representation of the signal transmitted by the DMA output is given by $\mathbf{s} = \mathbf{H}\,\mathbf{Q}\,\mathbf{z}$, where $\mathbf{H}$ is a $N \times N$ diagonal matrix with elements $\mathbf{H}_{((i-1)N_e+l,(i-1)N_e+l)} = h_{i,l}$, and $\mathbf{Q} \in \mathbb{C}^{N \times N_d}$ denotes the configurable weights of the DMAs, with each element given by

$$\mathbf{Q}_{(i-1)N_e+l,n} = \begin{cases} q_{i,l} & i = n, \\ 0 & i \neq n. \end{cases} \tag{13}$$

The DMA input signal is given by $\mathbf{z} = \sum_{m=1}^{M} \mathbf{w}_m x_m$, where $\mathbf{w}_m \in \mathbb{C}^{N_d}$ is the digital precoding vector for x_m. The baseband channel input transmitted by the DMA is thus given by $\mathbf{s} = \sum_{m=1}^{M} \mathbf{H}\mathbf{Q}\mathbf{w}_m x_m$. We again note that the transmitted signal is formally equal to that of a fully-digital architecture with precoding vectors $\tilde{\mathbf{w}}_m = \mathbf{H}\mathbf{Q}\mathbf{w}_m$ for each $m \in \mathcal{M}$. Consequently, the resulting achievable rate of the mth user for a given DMA configuration matrix $\mathbf{Q}$ and digital precoding vectors $\{\mathbf{w}_j\}$ is computed as $R_m(\{\mathbf{H}\mathbf{Q}\mathbf{w}_j\})$ using (8).

D. Problem Formulation

Based on the above model, we investigate multi-user communications in the near-field, considering the possibility of achieving reliable communications when different users share similar directions but are located at different distances from the BS. The aim here to is to design the transmission beam pattern to maximize the achievable sum-rate, reflecting the overall number of bits which can be reliably conveyed per channel use. Based on the different antenna architectures, for a given transmit power constraint $P_{\max} > 0$, the task of interest can be written as:

$$\begin{aligned} \max_{\{\tilde{\mathbf{w}}_m\}} \quad & \sum_{m=1}^{M} R_m\left(\{\tilde{\mathbf{w}}_j\}_{j\in\mathcal{M}}\right) \\ s.t. \quad & \sum_{m=1}^{M} \|\tilde{\mathbf{w}}_m\|^2 \leq P_{\max}, \quad \{\tilde{\mathbf{w}}_m\} \in \mathcal{W}, \end{aligned} \tag{14}$$

where $R_m\left(\{\tilde{\mathbf{w}}_j\}_{j\in\mathcal{M}}\right) = \log_2\left(1 + \frac{|\mathbf{a}_m^H \tilde{\mathbf{w}}_m|^2}{\sum_{j\neq m} |\mathbf{a}_m^H \tilde{\mathbf{w}}_j|^2 + \sigma^2}\right)$.

The problem formulated in (14) is similar to the corresponding problem encountered in far-field communications. The fact that communications is carried out in the near-field is encapsulated in the vectors $\{\mathbf{a}_m\}$. As a result, some of the tools used for tackling this problem in the sequel are adopted from studies considering far-field communications.

The set of feasible precoding vectors $\mathcal{W}$ in (14) captures the unique constraints imposed by the antenna architecture: for fully-digital UPAs, $\mathcal{W}_{\text{FD}}$ is the set of all M-tuples of vectors in $\mathbb{C}^N$; For hybrid beamformers, the feasible set $\mathcal{W}_{\text{HB}}$ is expressed as

$$\mathcal{W}_{\text{HB}} = \{\{\tilde{\mathbf{w}}_m\}_{m\in\mathcal{M}} | \tilde{\mathbf{w}}_m = \mathbf{Q}\mathbf{w}_m; \mathbf{Q} \in \mathcal{F}^{N \times N_d}\}; \tag{15}$$

For DMAs, the set of feasible precoders can be written as

$$\mathcal{W}_{\text{DMA}} = \{\{\tilde{\mathbf{w}}_m\}_{m\in\mathcal{M}} | \tilde{\mathbf{w}}_m = \mathbf{H}\mathbf{Q}\mathbf{w}_m\}, \tag{16}$$

where $\mathbf{H}$ is the fixed diagonal matrix representing the propagation inside the microstrips, while $\mathbf{Q}$ can be written as in (13) in which the non-zero elements take values in $\mathcal{Q}$ defined in (11).

For both the hybrid antenna and the DMA architectures, we optimize the corresponding $\{\mathbf{w}_m\}$ and $\mathbf{Q}$ to obtain the feasible precoders $\{\tilde{\mathbf{w}}_m\}$. For the hybrid antenna, each element of $\mathbf{Q}$ need to satisfy the unit modulus constraint in (10), whereas for the DMA, the non-zero elements of $\mathbf{Q}$ should satisfy the Lorentzian constraint in (11).

While the formulation of (14) does not explicitly account for the fact that transmission takes place in the near-field, this property is embedded in the equivalent channel vectors $\{\mathbf{a}_m\}$. As we show in the sequel, while the objective in (14) appears invariant of the shape of the resulting beams, maximizing the achievable sum-rate in the near field yields focused beams allowing to mitigate interference between users lying in the same angular direction.

III. Beam Focusing Aware Precoding Design

In this section, we study beam focusing-aware precoding design to maximize the sum-rate. We begin with unconstrained fully-digital antennas in Section III-A. Then, we derive the hybrid phase shifter setting in Section III-B, and consider DMA-based antenna architectures in Section III-C. We conclude with a discussion in Section III-D.

A. Fully-Digital Beam Focusing

For the fully-digital beam focusing design, the feasible precoding set $\mathcal{W}_{\text{FD}}$ is unconstrained, and includes all combinations of M vectors in $\mathbb{C}^N$. For the single-user case, i.e., $M = 1$, the rate in (14) is maximized by setting $\tilde{\mathbf{w}}_1 = \sqrt{P_{\max}} \frac{\mathbf{a}_1}{|\mathbf{a}_1|}$. However, for the general case of $M > 1$, problem (14) is non-convex, and thus it is difficult to find the optimal solution. Nevertheless, due to the similarity between (14) and the corresponding sum-rate maximization for interference broadcast channels operating in the far-field, one can utilize tools derived for far-field systems. A candidate strategy is to use the weighted sum mean-squared error (MSE) minimization approach [35] to deal with problem (14), which guarantees convergence to a stationary point.

By exploiting the relationship between sum-rate maximization and MSE minimization [35, Thm. 1], we have the following lemma.

Lemma 1: Problem (14) *with* $\mathcal{W} = \mathcal{W}_{\text{FD}}$ *is equivalent (in the sense of having the same global optimum) to the following problem*

$$\begin{aligned} \max_{\{\tilde{\mathbf{w}}_m, u_m, v_m\}} \quad & \sum_{m=1}^{M} \log_2(v_m) - v_m e_m\left(u_m, \{\tilde{\mathbf{w}}_m\}\right) \\ s.t. \quad & \sum_{m=1}^{M} \|\tilde{\mathbf{w}}_m\|^2 \leq P_{\max}, \quad v_m \geq 0, \ m \in \mathcal{M}, \end{aligned} \tag{17}$$

where u_m and v_m are auxiliary variables, and $e_m\big(u_m, \{\tilde{\mathbf{w}}_m\}\big)$ is given by $e_m\big(u_m, \{\tilde{\mathbf{w}}_m\}\big) = \big|1 - u_m\, \mathbf{a}_m^H\, \tilde{\mathbf{w}}_m\big|^2 + \sum_{j\neq m} \big|u_m\, \mathbf{a}_m^H\, \tilde{\mathbf{w}}_j\big|^2 + \sigma^2 \left|u_m\right|^2$.

Although problem (17) involves more optimization variables than (14), it is concave for each set of the optimization variables when the remaining two sets are fixed. Hence, the block coordinate descent method can be applied to solve (17), resulting in the procedure summarized as Algorithm 1, which is based on the method proposed in [35, Sec. III].

Algorithm 1 Alternating Optimization of Fully-Digital Precoders

Initialize: $\left\{\tilde{\mathbf{w}}_m^0\right\}_{m=1}^M$;

for $t = 1, \ldots, t_{\max}$ **do**

Update $u_m^t = \frac{\mathbf{a}_m^H\, \tilde{\mathbf{w}}_m^{t-1}}{\sum_{j=1}^M \left|\mathbf{a}_m^H\, \tilde{\mathbf{w}}_j^{t-1}\right|^2 + \sigma^2}$, $\forall m$;

Update $v_m^t = \left(e_m\left(u_m^t, \left\{\tilde{\mathbf{w}}_m^{t-1}\right\}\right)\right)^{-1}$, $\forall m$;

Update $\tilde{\mathbf{w}}_m^t = u_m^t\, v_m^t \left(\sum_{j=1}^M v_j^t \left|u_j^t\right|^2 \mathbf{a}_j \mathbf{a}_j^H + \lambda_p \mathbf{I}\right)^{-1} \mathbf{a}_m$;

end for

Output: $\{\mathbf{w}_m^t\}$.

In Algorithm 1, the parameter λ_p in step 4 is the Lagrangian multiplier associated with the transmit power constraint of the BS. The selection of λ_p can be set by hyperparameter optimization schemes, e.g., using the bisection method [27], [35]. Algorithm 1 is ignorant of the fact that communications takes place in the near-field, as this property is only encapsulated in the equivalent channel vectors $\{\mathbf{a}_m\}$. Nonetheless, as we show numerically in Section IV, this optimization method, that targets the sum-rate and does not explicitly account for the resulting beam pattern, yields focused beams, which allows multiple users to co-exist with minimal cross interference while residing in the same angular direction.

B. Beam Focusing via Phase-Shifters Based Hybrid Precoding

We next focus on the design of phase-shifters based hybrid antenna architecture. In this case, by defining the $N_d \times M$ matrix $\mathbf{W} \triangleq [\mathbf{w}_1, \ldots, \mathbf{w}_M]$, (14) is re-expressed as

$$\max_{\{\mathbf{w}_m\},\mathbf{Q}} \sum_{m=1}^M \log_2\left(1 + \frac{\left|\mathbf{a}_m^H\, \mathbf{Q}\mathbf{w}_m\right|^2}{\sum_{j\neq m} \left|\mathbf{a}_m^H\, \mathbf{Q}\mathbf{w}_j\right|^2 + \sigma^2}\right),$$
$$s.t.\ \ \|\mathbf{Q}\mathbf{W}\|_F^2 \leq P_{\max},\ \ \mathbf{Q} \in \mathcal{F}^{N\times N_{\rm RF}}. \tag{18}$$

Problem (18) is non-convex due to both the coupled optimization variables and the unit modulus constraints. A reasonable way out is to let the hybrid solution to (18) be sufficiently "close" to the fully-digital solution of problem (14). In particular, our design seeks to identify the hybrid precoding mapping which is the closest to the resulting fully-digital precoding in the Frobenious norm sense. This strategy is quite often used for optimizing hybrid analog/digital systems, see, e.g., [19], [21], [36], [37]. According to [18], the hybrid solution can be rigorously shown to be a tight approximation of the fully-digital solution when the number of RF chains is at least twice the number of users (data streams). Specifically, let $\tilde{\mathbf{W}}_{\rm opt} = [\tilde{\mathbf{w}}_1, \cdots, \tilde{\mathbf{w}}_M]$ be the unconstrained precoding matrix obtained via Algorithm 1. The resulting surrogate optimization problem is given by

$$\min_{\mathbf{Q},\mathbf{W}} \ \left\|\tilde{\mathbf{W}}_{\rm opt} - \mathbf{Q}\mathbf{W}\right\|^2$$
$$s.t.\ \mathbf{Q} \in \mathcal{F}^{N\times N_{\rm RF}}. \tag{19}$$

Note that we have temporarily neglected the power constraint in (19). Nonetheless, once $\mathbf{Q}$ and $\mathbf{W}$ are tuned to optimize (19), we update the digital precoder by multiplying a factor, i.e., $\mathbf{W} = \frac{\sqrt{P_{\max}}\mathbf{W}}{\|\mathbf{Q}\mathbf{W}\|^2}$. Thus, we tackle (19) using alternating optimization. In particular, for a given $\mathbf{Q}$, the digital precoding matrix $\mathbf{W}$ which minimizes (19) is stated in the following lemma:

Lemma 2: For a given $\mathbf{Q}$, the matrix $\mathbf{W}$ which minimizes (19) *is*

$$\mathbf{W} = \left(\mathbf{Q}^H\mathbf{Q}\right)^{-1} \mathbf{Q}^H \tilde{\mathbf{W}}_{\rm opt}. \tag{20}$$

Proof: The lemma is obtained as the least-squares solution to (19) with fixed $\mathbf{Q}$. ∎

To optimize $\mathbf{Q}$ for a given $\mathbf{W}$, we exploit the fact that the unit- modulus constraint on the norm of $\mathbf{Q}$ bears similarity to constraints encountered in optimizing RISs. In particular, by defining $\mathbf{q}_{\rm S} = {\rm Vec}\left(\mathbf{Q}\right)$, we equivalently reformulate (19) for a given $\mathbf{W}$ as

$$\min_{\mathbf{q}_{\rm S}\in\mathcal{M}} \ f\left(\mathbf{q}_{\rm S}\right) \triangleq \left\|{\rm Vec}\left(\tilde{\mathbf{W}}_{\rm opt}\right) - \left(\mathbf{W}^T \otimes \mathbf{I}\right)\mathbf{q}_{\rm S}\right\|^2 \tag{21}$$

where $\mathcal{S} = \left\{\mathbf{q}_{\rm S} \in \mathbb{C}^L : |\mathbf{q}_{{\rm S},1}| = \cdots = |\mathbf{q}_{{\rm S},L}| = 1\right\}$ denotes the search space, and $L \triangleq N \times N_{\rm RF}$.

Note that the search space $\mathcal{S}$ in (21) is the product of L complex circles, which is a Riemannian submanifold of $\mathbb{C}^L$. Therefore, following [25]–[27], we tackle (21) using the Riemannian conjugate gradient algorithm. The solution to (21) is thus updated based on the following formula

$$\mathbf{q}_{\rm S}^{t+1} = \mathcal{R}_t\left(\mathbf{q}_{\rm S}^t + \varsigma^t \boldsymbol{\eta}^t\right), \tag{22}$$

where $\mathbf{q}_{\rm S}^t \in \mathcal{S}$ and $\mathbf{q}_{\rm S}^{t+1} \in \mathcal{S}$ denote the current point and the next point, respectively; $\varsigma^t > 0$ and $\boldsymbol{\eta}^t$ are the Armijo step size [38] and the search direction at the point $\mathbf{q}_{\rm S}^t$, respectively; $\mathcal{R}_t\left(\cdot\right)$ denotes the retraction operator, which projects the vector $\mathbf{q}_{\rm S}^t + \varsigma^t \boldsymbol{\eta}^t$ to the search space $\mathcal{S}$ via element-wise retraction, i.e., $\left(\mathbf{q}_{\rm S}^{t+1}\right)_l = \frac{\left(\mathbf{q}_{\rm S}^t + \varsigma^t \boldsymbol{\eta}^t\right)_l}{\left|\left(\mathbf{q}_{\rm S}^t + \varsigma^t \boldsymbol{\eta}^t\right)_l\right|}$ for $l = 1, \cdots, L$. Let $T_{\mathbf{q}_{\rm S}^t}\mathcal{S}$ denote the tangent space of the complex circle manifold $\mathcal{S}$ at the point $\mathbf{q}_{\rm S}^t$, which is composed of all the tangent vectors that tangentially pass through $\mathbf{q}_{\rm S}^t$. The search direction $\boldsymbol{\eta}^t$ lies in $T_{\mathbf{q}_{\rm S}^t}\mathcal{S}$ [39], given by

$$\boldsymbol{\eta}^t = -\,{\rm grad}\ f\left(\mathbf{q}_{\rm S}^t\right) + \alpha^t \mathcal{T}_{\mathbf{q}_{\rm S}^{t-1}\rightarrow\mathbf{q}_{\rm S}^t}\left(\boldsymbol{\eta}^{t-1}\right), \tag{23}$$

where α^t is chosen as the Polak-Ribiere parameter [38]; ${\rm grad}\ f\left(\mathbf{q}_{\rm S}^t\right)$ represents the Riemannian gradient of $f\left(\mathbf{q}_{\rm S}\right)$ at the point $\mathbf{q}_{\rm S}^t$, which is obtained by orthogonally projecting its Euclidean gradient, denoted by $\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)$, onto the tangent space $T_{\mathbf{q}_{\mathrm{S}}^{t}}\mathcal{S}$, i.e.,

$$\begin{aligned}&\mathrm{grad}\ f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)\\&=\begin{bmatrix}\left(\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)\right)_1-\mathrm{Re}\left\{\left(\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)\right)_1\times\left(\mathbf{q}_{\mathrm{S}}^{t}\right)_1^{\dagger}\right\}\left(\mathbf{q}_{\mathrm{S}}^{t}\right)_1\\ \vdots\\ \left(\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)\right)_L-\mathrm{Re}\left\{\left(\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)\right)_L\times\left(\mathbf{q}_{\mathrm{S}}^{t}\right)_L^{\dagger}\right\}\left(\mathbf{q}_{\mathrm{S}}^{t}\right)_L\end{bmatrix}\\&=\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)-\mathrm{Re}\left\{\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)\circ\left(\mathbf{q}_{\mathrm{S}}^{t}\right)^{\dagger}\right\}\circ\mathbf{q}_{\mathrm{S}}^{t},\end{aligned}\tag{24}$$

where $(\mathbf{x})_i$ denotes the ith entry of a vector $\mathbf{x}$, $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number, and the notation $\circ$ represents the Hadamard production (element-wise multiplication) operation. The Euclidean gradient is defined as $\nabla f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)=2\left(\mathbf{W}^{\dagger}\otimes\mathbf{I}\right)\left(\left(\mathbf{W}^{T}\otimes\mathbf{I}\right)\mathbf{q}_{\mathrm{S}}^{t}-\mathrm{Vec}\left(\tilde{\mathbf{W}}_{\mathrm{opt}}\right)\right)$.

In (23), $\mathcal{T}_{\mathbf{q}_{\mathrm{S}}^{t-1}\to\mathbf{q}_{\mathrm{S}}^{t}}\left(\boldsymbol{\eta}^{t-1}\right)$ denotes the vector transport, which maps the previous search direction $\boldsymbol{\eta}^{t-1}$ (lying in the tangent space $T_{\mathbf{q}_{\mathrm{S}}^{t-1}}\mathcal{S}$) to the tangent space $T_{\mathbf{q}_{\mathrm{S}}^{t}}\mathcal{S}$. As a result, $\mathrm{grad}\ f\left(\mathbf{q}_{\mathrm{S}}^{t}\right)$ and $\mathcal{T}_{\mathbf{q}_{\mathrm{S}}^{t-1}\to\mathbf{q}_{\mathrm{S}}^{t}}\left(\boldsymbol{\eta}^{t-1}\right)$ lie in the same tangent space $T_{\mathbf{q}_{\mathrm{S}}^{t}}\mathcal{S}$, and thus the sum operation in (23) makes sense. The vector transport is given by $\mathcal{T}_{\mathbf{q}_{\mathrm{S}}^{t-1}\to\mathbf{q}_{\mathrm{S}}^{t}}\left(\boldsymbol{\eta}^{t-1}\right)=\boldsymbol{\eta}^{t-1}-\mathrm{Re}\left\{\boldsymbol{\eta}^{t-1}\circ\left(\mathbf{q}_{\mathrm{S}}^{t}\right)^{\dagger}\right\}\circ\mathbf{q}_{\mathrm{S}}^{t}$.

Finally, the overall resulting configuration algorithm for hybrid analog precoders is summarized as Algorithm 2.

Algorithm 2 Alternating Optimization of Hybrid Precoders

Initialize: $\mathbf{W}^0$, $\mathbf{Q}^0$;
Obtain $\tilde{\mathbf{W}}_{\mathrm{opt}}$ via Algorithm 1;
for $j=1,\dots,j_{\max}$ **do**
  $\mathbf{q}_{\mathrm{S}}^{0}=\mathrm{Vec}\left(\mathbf{Q}^{0}\right)$, $\boldsymbol{\eta}^{0}=-\,\mathrm{grad}\ f\left(\mathbf{q}_{\mathrm{S}}^{0}\right)$;
  for $t=1,\dots,t_{\max}$ **do**
    Find the next point $\mathbf{q}_{\mathrm{S}}^{t+1}$ *according to* (22);
    Update the search direction $\boldsymbol{\eta}^{t+1}$ *according to* (23);
  end for
  Set $\mathbf{Q}^{j}=\mathrm{Vec}^{-1}\left(\mathbf{q}_{\mathrm{S}}^{t_{\max}}\right)$;
  Set $\mathbf{W}^{j}$ via Lemma 2 with $\mathbf{Q}=\mathbf{Q}^{j}$.
end for
Output: $\mathbf{Q}^{j}$, $\mathbf{W}^{j}$.

C. DMA-Based Beam Focusing

Here, we consider the configuration of DMAs for maximizing the sum-rate in near-field downlink communications. We note that the architectures considered in the previous subsections are relatively well-studied in the wireless communications literature, and thus we were able to utilize methods previously derived for similar setups to optimize the precoder. However, as the application of DMAs for wireless communications is a relatively new area of research, in the following we derive a dedicated algorithm for configuring their weights based on (14). In particular, we first reformulate (14) for the case of $\mathcal{W}=\mathcal{W}_{\mathrm{DMA}}$ as

$$\begin{aligned}&\max_{\{\mathbf{w}_m\},\mathbf{Q}}\ \sum_{m=1}^{M}\log_2\left(1+\frac{\left|\mathbf{a}_m^{H}\mathbf{H}\mathbf{Q}\,\mathbf{w}_m\right|^2}{\sum_{j\neq m}\left|\mathbf{a}_m^{H}\mathbf{H}\mathbf{Q}\,\mathbf{w}_j\right|^2+\sigma^2}\right)\\&\quad s.t.\ (13),\quad q_{i,l}\in\mathcal{Q},\forall i,l,\ \ \sum_{m=1}^{M}\left\|\mathbf{w}_m\right\|^2\le P_{\max}.\end{aligned}\tag{25}$$

We note that (25) is slightly different from (14), as the power constraint here is imposed on the digital output, and not the transmitted signal. Nonetheless, as discussed in the previous subsection, one can derive the overall system based on (25), and scale the digital precoder such that the power constraint in (14) is satisfied. As the joint design of the configuration of the DMA weights along with the digital precoding vector based on the non-convex problem (25) is challenging, we begin by considering a single-user setup to gain more design insights. After that, we extend our study to the multi-user case of $M>1$, and propose an alternating algorithm to deal with the resulting non-convex optimization problem.

1) Single-User Case: For the single-user case, there is no inter-user interference. Consequently, the achievable rate is given by $R=\log_2\left(1+\frac{1}{\sigma^2}\left|\mathbf{a}^{H}\mathbf{H}\mathbf{Q}\,\mathbf{w}\right|^2\right)$, where we have dropped the user index subscript m. Due to the monotonicity of the logarithm function, the optimization problem (25) is equivalently rewritten as

$$\begin{aligned}&\max_{\mathbf{w},\mathbf{Q}}\ \left|\mathbf{a}^{H}\mathbf{H}\mathbf{Q}\,\mathbf{w}\right|^2\\&s.t.\ (13),\quad q_{i,l}\in\mathcal{Q},\forall i,l,\ \ \left\|\mathbf{w}\right\|^2\le P_{\max}.\end{aligned}\tag{26}$$

Although (26) is notably simpler to tackle compared to (25), it still involves coupled optimization variables in the objective function, as well as the non-trivial element-wise constraint $\mathcal{Q}$. Specifically, each element response $q_{i,l}$ should take the Lorentzian-constrained form in (11), which is represented by a circle in the top half of the complex plane (inner circle in Fig. 5). Thus, the phase and amplitude of $q_{i,l}$ are coupled, which makes it challenging to solve (26). To tackle this problem, we employ the constant amplitude approach proposed in [34] to relax the Lorentzian constraint to the phase-only weights constraint with constant amplitude and arbitrary phase, given by $q_{i,l}\in\mathcal{F}$ as defined in (10), for each i,l. The feasible set $\mathcal{F}$ is a circle with unit radius and centered at the origin (outer circle in Fig. 5).

We next focus on problem (26) with $\mathcal{Q}$ replaced by $\mathcal{F}$, and then use the solution to tune the Lorentzian-constrained weights via projection as illustrated in Fig. 5. Due to the coupled optimization variables, the problem is still non-convex. Nonetheless, we are able to solve it in closed-form, as stated in the following theorem.

Theorem 1: Let $(\mathbf{Q}^*,\mathbf{w}^*)$ *be the solution to* (26) *with* $\mathcal{Q}=\mathcal{F}$. *According to the structure constraint* (13)*, each non-zero element of* $\mathbf{Q}^*$ *is* $q_{i,l}^*=e^{\jmath\psi_{i,l}^*}$*, with* $\psi_{i,l}^*=k|\mathbf{p}_m-\mathbf{p}_{i,l}|+\beta_i\rho_{i,l}$*, and* $\mathbf{w}^*=\frac{\sqrt{P_{\max}}\left(\mathbf{a}^{H}\mathbf{H}\mathbf{Q}^*\right)^{H}}{\left\|\mathbf{a}^{H}\mathbf{H}\mathbf{Q}^*\right\|}$.

Proof: The proof is given in Appendix A. ■

From Theorem 1, we can see that the optimized phase $\psi_{i,l}^*$ of each metamaterial element includes two parts: The first is the term $k|\mathbf{p}-\mathbf{p}_{i,l}|$ for achieving near-field focusing, i.e., coherent sum of signal components in location $\mathbf{p}$; the other part compensates for the transmission delay in the microstrip, given by $\beta_i\rho_{i,l}$ (12), such that the signals are transmitted synchronously. It is noted that a similar phase profile would be found using an holographic design process [40].

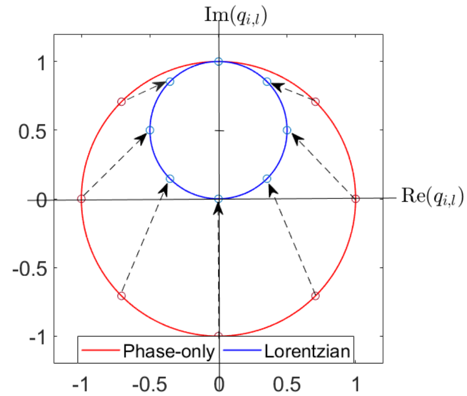


Fig. 5. Illustration of the phase-only weights (outer circle) and Lorentzian weights (inner circle) in the complex plane. Arrows indicate the mapping between the phase-only weights and Lorentzian weights points.

As the weighting $q^*_{i,l} = e^{\jmath\psi^*_{i,l}}$ does not satisfy the Lorentzian form defined in (11), we project it onto $\mathcal{Q}$ as illustrated in Fig. 5. In particular, the resulting non-zero weights are given by $\hat{q}_{i,l} = \frac{j+e^{j\psi^*_{i,l}}}{2}$ [34], and thus $\hat{q}_{i,l} \in \mathcal{Q}$. Hence, the resulting $\hat{\mathbf{Q}}$ can be used as an approximate solution to the original optimization problem (26). The numerical results provided in Section IV verify that deriving the elements response assuming weights of the form $\mathcal{F}$ followed by their projection onto $\mathcal{Q}$ yields accurate focused beams. The proposed relaxation and projection approach facilitates the design of effective DMA weights for the single-user case.

2) Multi-User Case: Next, we study the formulation of focused beams using DMAs for multi-user setups by tackling (25) with $M > 1$. Following the strategy used in Section III-B, we optimize $\mathbf{Q}$ and $\{\mathbf{w}_m\}$ in an alternating manner, while relaxing the constraints on the feasible elements response as done in the single-user setup. This procedure is iterated until convergence. In the following, we show how to solve (25) for fixed $\mathbf{Q}$ and for fixed $\{\mathbf{w}_m\}$, respectively.

a) Solving (25) *w.r.t.* $\{\mathbf{w}_m\}$*:* When $\mathbf{Q}$ is fixed, the form of problem (25) is similar to the sum-rate maximization problem studied in the previous Section III-A. Hence, following Lemma 1, problem (25) with fixed $\mathbf{Q}$ is equivalent to

$$\max_{\{\mathbf{w}_m, u_m, v_m\}} \sum_{m=1}^{M} \log_2(v_m) - v_m e_m^{\mathrm{DMA}}(u_m, \{\mathbf{w}_m\})$$
$$s.t. \sum_{m=1}^{M} \|\mathbf{w}_m\|^2 \le P_{\max}, \quad v_m \ge 0, \ m \in \mathcal{M}. \quad (27)$$

where u_m and v_m are auxiliary variables, and $e_m^{\mathrm{DMA}}(u_m, \{\mathbf{w}_m\}) = \left|1 - u_m \mathbf{a}_m^H \mathbf{H}\mathbf{Q}\,\mathbf{w}_m\right|^2 + \sum_{j\neq m}\left|u_m\,\mathbf{a}_m^H\,\mathbf{H}\mathbf{Q}\,\mathbf{w}_j\right|^2 + \sigma^2\left|u_m\right|^2$. One can verify that problem (27) has the same structure as problem (17). Hence, Algorithm 1 is applied to solve (27).

b) Solving (25) *w.r.t.* $\mathbf{Q}$*:* To proceed, we first define the $N_d^2 \cdot N_e \times 1$ vectors $\mathbf{q} = \mathrm{Vec}(\mathbf{Q})$, and $\mathbf{z}_{j,m} = \left(\mathbf{w}_j^T \otimes (\mathbf{a}_m^H \mathbf{H})\right)^H$. Using these definitions, we then identify an equivalent optimization problem, as stated in following theorem.

Theorem 2: Problem (25) *with fixed* $\{\mathbf{w}_m\}$ *is equivalent to the following problem*:

$$\max_{\bar{\mathbf{q}}} \ f(\bar{\mathbf{q}}) \triangleq \sum_{m=1}^{M} \log_2\left(1 + \frac{\left|\bar{\mathbf{q}}^H \bar{\mathbf{z}}_{m,m}\right|^2}{\sum_{j\neq m}\left|\bar{\mathbf{q}}^H \bar{\mathbf{z}}_{j,m}\right|^2 + \sigma^2}\right)$$
$$s.t. \ \bar{q}_l \in \mathcal{Q}, \quad l \in \mathcal{A}_q, \quad (28)$$

where $\mathcal{A}_q$ *is the set of all non-zero elements of* $\mathbf{q}$, $\bar{\mathbf{q}}$ *is the modified version of* $\mathbf{q}$ *obtained by removing all the zero elements of* $\mathbf{q}$; $\bar{\mathbf{z}}_{j,m}$ *is the modified version of* $\mathbf{z}_{j,m}$, *which is obtained by removing the elements having the same index as the zero elements of* $\mathbf{q}$.

Proof: The proof is given in Appendix B. ∎

The equivalence between the relaxed (25) and (28) holds in the sense that they have the same optimal value, and the solution to the relaxation of (25) can be recovered from the solution to (28) according to (13). Although (28) is still non-convex, we can find its approximate solution using alternating optimization, by iteratively optimizing each element of $\bar{\mathbf{q}}$ separately while keeping the remaining elements fixed. The resulting optimization problem for the lth element is

$$\max_{0\le\phi_l\le 2\pi} \sum_{m=1}^{M} \log_2\left(1 + \frac{\left|\bar{\mathbf{q}}(\phi_l)^H \bar{\mathbf{z}}_{m,m}\right|^2}{\sum_{j\neq m}\left|\bar{\mathbf{q}}(\phi_l)^H \bar{\mathbf{z}}_{j,m}\right|^2 + \sigma^2}\right), \quad (29)$$

where $\bar{\mathbf{q}}(\phi_l) \triangleq \left[\bar{q}_1, \cdots, \bar{q}_{l-1}, \frac{j+e^{j\phi_l}}{2}, \bar{q}_{l+1}, \cdots, \bar{q}_{N_d N_e}\right]$, $l \in \mathcal{A}_q$.

Problem (29) is a single variable optimization problem with respect to ϕ_l. Although it is challenging to find a closed-form solution for (29), we can easily solve it by using numerical techniques such as a one-dimensional search. The proposed algorithm for solving (25) for the case of $M > 1$ is summarized as Algorithm 3. In Algorithm 3, steps 3 to 7 solve (27) for a fixed $\mathbf{Q}$ using the similar procedure as in Algorithm 1, and steps 11 to 14 solve (28) for a fixed $\{\mathbf{w}_m\}$ using the alternating optimization approach.

D. Discussion

In line with traditional multi-user communication schemes, our derivation in the previous sections considers the achievable sum-rate as the objective function to be optimized. Consequently, while our work focuses on beam focusing, we do not explicitly design the transmission beam patterns to generate focused beams, but rather to establish reliable high-rate communications. The fact that we are operating in the near-field, from which the beam focusing ability arises, is implicitly encapsulated in the objective via the vectors $\{\mathbf{a}_m\}$, as discussed in Section II-B. This property allowed us to combine in our derivations methods developed for far-field systems. Specifically, the optimization of the precoding steps using fully-digital and hybrid antennas are based on the identification of existing methods in the far-field literature. Nonetheless, by applying these derivations in the radiating near-field regime, we reveal the ability to achieve beam focusing in multi-user wireless communications and its effect on mutual interference, and do so in a manner which is balanced among candidate

Algorithm 3 Alternating Optimization of DMA Precoders for $M > 1$

Initialize: $\{\mathbf{w}_m^0\}_{m=1}^M$, $\mathbf{Q}^0$;
for $t = 1, \ldots, t_{\max}$ **do**
  Update $\mathbf{g}_m \triangleq (\mathbf{Q}^{t-1})^H \mathbf{H}^H \mathbf{a}_m, \forall m$
  for $t_1 = 1, \ldots, t_{\max}$ **do**
    Update $u_m^{t_1} = \frac{\mathbf{g}_m^H \mathbf{w}_m^{t_1-1}}{\sum_{j=1}^M \left|\mathbf{g}_m^H \mathbf{w}_j^{t_1-1}\right|^2 + \sigma^2}$, $\forall m$;
    Update $v_m^{t_1} = (e_m^{\text{DMA}}(u_m^{t_1}, \{\mathbf{w}_m^{t_1-1}\}))^{-1}$, $\forall m$;
    Update $\mathbf{w}_m^{t_1} = u_m^{t_1} v_m^{t_1} \left(\sum_{j=1}^M v_j^{t_1} \left|u_j^{t_1}\right|^2 \mathbf{g}_j \mathbf{g}_j^H + \lambda_p \mathbf{I}\right)^{-1} \mathbf{g}_m$;
  end for
  Update $\mathbf{w}_m = \mathbf{w}_m^{t_1}, \forall m$;
  Update $\bar{\mathbf{z}}_{j,m}, \forall m, j$;
  Update $\bar{\mathbf{q}} = \text{Vec}\left(\mathbf{Q}^{t-1}\right)$;
  for $l = 1, \ldots, N_d N_e$ **do**
    Update ϕ_l by solving problem (29);
    Update the lth element of $\bar{\mathbf{q}}$: $\bar{q}_l = \frac{j + e^{j\phi_l}}{2}$;
  end for
  Update $\mathbf{Q}^t$ with non-zero entries taken from $\bar{\mathbf{q}}$;
end for
Output: $\{\mathbf{w}_m^{t_1}\}$ and $\mathbf{Q}^t$.

future antenna architectures. Despite the fact that we do not directly enforce the generation of focused beams, our numerical study in Section IV demonstrates that such beams are indeed generated when seeking to optimize the sum-rate, enabling users with identical angles to simultaneously achieve high rates with little interference.

Among the considered architectures, the fully-digital antenna supports the most flexible design, and it is expected to achieve the largest sum-rates among all considered architectures for a given UPA with fixed element placing. However, such architectures assign a costly RF chain to each element, and thus may be prohibitively costly when the number of antennas elements N becomes very large. To achieve cost-effective power transmission, hybrid antenna architectures with limited RF chains are often adopted. Nonetheless, as the total number of required phase shifters is very large, the power consumption of active phase-shifters based analog precoding may also become significant. The third considered architecture, i.e., DMAs, is the most scalable in terms of cost and power efficiency. Nonetheless, it is also the most challenging to design due to the Lorentzian form of its elements, whose gain and phase are coupled. For this reason, our design method resorted to optimizing each element separately in an alternating manner, as detailed in Algorithm 3. Furthermore, DMAs are typically utilized with sub-wavelength element spacing, allowing to pack a larger number of elements in a given physical area compared to conventional antennas based on, e.g., patch arrays [41]. Therefore, for a given antenna aperture, DMAs are in fact capable of achieving the most focused beams among all considered architecture for the single-user case, as will be shown by the numerical results in Section IV.

The model used for the frequency response of the DMA elements is the Lorentzian-phase constrained form (11), which does not vary in frequency. Nonetheless, DMA elements can also be configured to exhibit controllable frequency selectivity, i.e., a different response can be configured in each frequency bin in a coupled manner [22]. This advanced frequency selective analog signal processing capabilities, which are not available in conventional hybrid antenna architectures based on phase shifters [19], give rise to the possibility of generating frequency-variant beam focusing patterns, facilitating high-rate wideband communications with a large number of users. Furthermore, we focus on the typical DMA architecture comprised of a set of one-dimensional microstrips, which result in the equivalent partially-connected model in (13) and hence brings in some performance loss compared to fully-connected analog combiners for the multi-user scenario. However, one can also use two-dimensional waveguides [42] resulting in an equivalent fully connected model, i.e., $\mathbf{Q}$ is not restricted to take the form (13), although such an architecture results in a more complex model for the propagation inside the waveguide. Finally, in the current work, we study the design of focused beams under the assumption of free-space condition and the perfect knowledge of users' location, e.g., the vectors $\mathbf{a}_m$ in (5) is perfectly known. In practice, one usually has to estimate it, giving rise to estimation errors which are likely to affect the ability to generate focused beams. The extension of the current study to near-field communications with estimated channels is left for future work.

IV. Numerical Evaluations

Here, we provide numerical results to verify the near-field beam focusing capability of our proposed designs under three different antenna architectures. We first consider the single-user scenario in Section IV-A, which demonstrates the gains of near-field beam focusing over beam steering in enhancing the signal strength of the target point. Then, in Section IV-B, we show the advantage of beam focusing to distinguish different users for multi-user communications.

Throughout the experimental study, we consider a planar array positioned in the xy-plane, with carrier frequency set to $f_c = 28$ GHz ($\lambda = 1.07$ cm). To make a fair comparison, We consider that the three different types of antennas have the same physical aperture and are all equal to $D = \sqrt{2}L$, with L denoting antenna length. We consider $\lambda/2$ antenna separation for fully-digital and hybrid antennas, and $\lambda/5$ spacing between DMA elements within the same row unless other wise stated. The separation between rows could still be $\lambda/2$. Hence, the number of rows and elements in each row for the fully-digital and hybrid antennas are $N_d = N_e = \lfloor 2L/\lambda \rfloor$, where $\lfloor \cdot \rfloor$ is the integer floor function. For the phase-shifter based analog precoder we set the number of RF chains to $N_{\text{RF}} = N_d$, while for the DMA, the number of microstrips and that of metamaterial elements are $N_d = \lfloor 2L/\lambda \rfloor$ and $N_e = \lfloor 5L/\lambda \rfloor$ respectively. We use $\alpha = 0.6$ $[\text{m}^{-1}]$ and $\beta = 827.67$ $[\text{m}^{-1}]$ to represent the propagation inside the DMA waveguides, assuming a microstrip implemented in Duroid 5880 with 30 mill thickness. In addition, we set the maximum

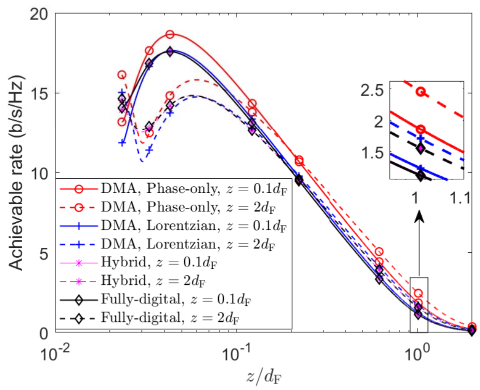


Fig. 6. Achievable rates of beam focusing at F_{near} and beam steering at F_{far}.

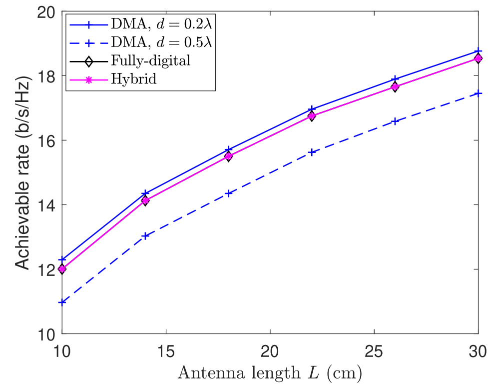


Fig. 7. Achievable rate versus the antenna length.

transmit power to $P_{\max} = -13\,\text{dBm}$, and the noise power to $\sigma^2 = -114\,\text{dBm}$.[2]

A. Single-User Scenario

We first show the beam focusing design in the single-user scenario, where the antenna length is set to be $L = 30$ cm. Fig. 6 illustrates the achievable rates along the z-axis direction, namely, when the single user is located in each specific point in the z-axis, achieved by the beam focusing and beam steering solutions under the three different considered antenna architectures. For DMAs, we depict here the rate achieved with phase-only elements in addition to that achieved using their true Lorentizan, as phase-only DMA weights are derived as an intermediate step in the optimization algorithm detailed in Section III-C. Beam focusing is obtained by setting the focusing points at the near-field $\text{F}_{\text{near}}((x, y, z) = (0, 0, 0.1d_{\text{F}}))$, whereas the beam steering is at the far-field $\text{F}_{\text{far}}((x, y, z) = (0, 0, 2d_{\text{F}}))$. In Fig. 6, we clearly observe that for all antenna architectures, near-field focusing can increase the signal strength, which in turn improves the achievable rate, when the user is located in the proximity of the focusing point of $z = 0.1d_{\text{F}}$. Secondly, when the user is located in the direction of the focusing point but at a different distance from the antenna plane, the observed radiation is notably reduced compared with the corresponding far-field beam steering design, resulting in lower rates.

The achievable rate of beam focusing in Fig. 6 is much larger than that of the beam steering solution at the near-field focusing point F_{near}. In particular, the DMA architecture achieves higher rates compared to fully-digital and hybrid antennas. This is because DMA stacks more antenna elements within the same antenna area due to the smaller element spacing than the other two types of antennas. Moreover, we note that the hybrid antenna achieves the same performance as that of the fully-digital antenna. This observation is consistent with existing results in far-field communications, where it is established that any fully-digital beamforming solution can be realized by a hybrid beamforing solution when the number of RF chains is at least twice the number of users (data streams) [18]. Finally, we note that the achievable rate of the Lorentzian weights constraint is comparable to that of the phase-only weights constraints, which verifies the effectiveness of our DMA configuration approach. In the remainder of the experimental study we thus consider only DMAs with lorentzian-constrained weights.

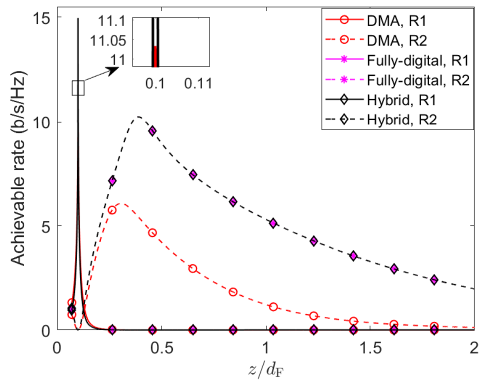
Next, Fig. 7 illustrates the effect of antenna size (or antenna aperture) on the achievable rate of each of the three antenna architectures. We set the focusing point at $\text{F}_{\text{near}}((x, y, z) = (0, 0, 150\,\lambda))$. From Fig. 7, we observe that as the antenna length L increases, the achievable rate substantially improves for each of the three antenna architectures. This reveals an important insight that large antenna arrays cannot only increase the near-field region ($d_{\text{F}} = \frac{2D^2}{\lambda}$), but also enhance the signal strength at the near-field focusing point. Moreover, it is observed that the achievable rate of DMAs with $0.2\,\lambda$ element spacing is higher than that of both fully-digital and hybrid antennas, while the performance of DMAs with $0.5\,\lambda$ element spacing, which deviates from conventional DMA deployment as their ability to stack more elements in a given aperture is not exploited [24], is worse than the other two antenna architectures. This is expected since the smaller the element spacing, the more antenna elements can be patched within the same antenna area.

B. Multi-User Scenario

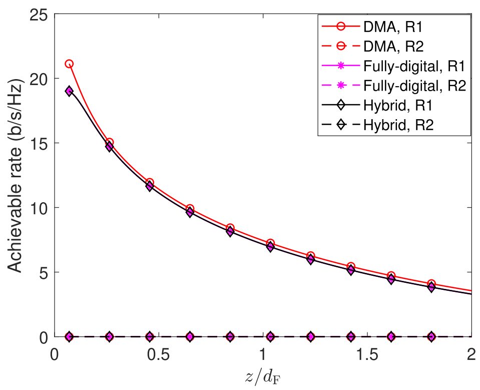
We proceed to show the advantage of beam focusing to distinguish different users for downlink communications. In particular, we would show that beam focusing provides a new degree of freedom to mitigate multiuser interference, i.e., it cannot only control the multiuser interference in the angle domain, as traditional beam steering but also control the interference in the distance domain. Having demonstrated the effect of antenna length on the achievable rate in the previous section, we fix the antenna length here to $L = 10$ cm. We first study the two-user case in Figs. 8 and 9. Then we study the general multiple users ($M > 2$) case where the users are randomly distributed on the xz-plane within the near-field region in Fig. 10.

Fig. 8 illustrates the achievable rates of each of the two users, i.e. $\{R_1, R_2\}$, measured along the z-axis. The two focal points of the users are located at: (a) near-field region: $\text{F}_1((x, y, z_1) = (0, 0, 0.1d_{\text{F}}))$ and $\text{F}_2((x, y, z_2) = (0, 0, 0.4d_{\text{F}}))$; and (b) far-field region: $\text{F}_1((x, y, z_1) = (0, 0, 1.5d_{\text{F}}))$ and $\text{F}_2((x, y, z_2) = (0, 0, 1.8d_{\text{F}}))$. From Fig. 8(a),

[2]This setting corresponds to the noise power spectrum density at users is -174 dBm/Hz and signal bandwidth is $120\,\text{KHz}$, assuming the noise figure of each user to be 9 dB.

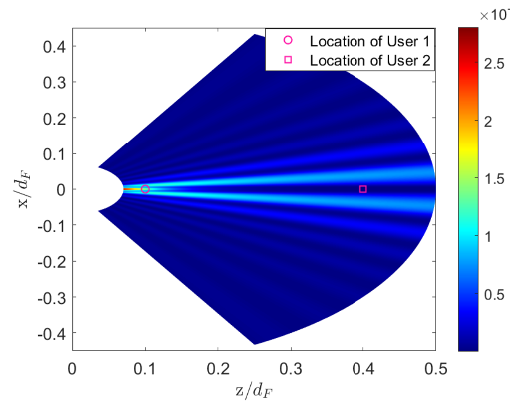


(a) Near-field: $z_1 = 0.1d_\mathrm{F}$ and $z_2 = 0.4d_\mathrm{F}$.

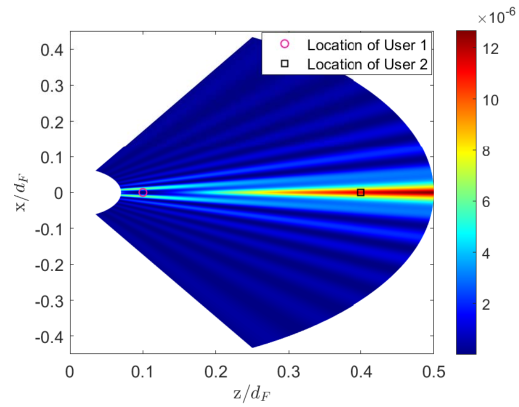


(b) Far-field: $z_1 = 1.5d_\mathrm{F}$ and $z_2 = 1.8d_\mathrm{F}$.

Fig. 8. Achievable rates per user versus location along the z-axis.



(a) The normalized signal power of user 1



(b) The normalized signal power of user 2

Fig. 9. The normalized signal power measurement of two users.

it is observed that for all the studied antenna architectures, the peak achievable rates of each of the two users occur when they are located around their corresponding focal points, implying that the designed focused beams for each type of the antennas are all capable of yielding reliable communications with minimal degradation due to interference. For example, the achievable rate of user 1 is maximized when it is located at $(0, 0, z_1)$. Moreover, it is observed that the fully-digital and hybrid antennas achieve higher rates than the DMA architecture for the near-filed two-user scenario. This is due to the fact that we consider the DMA with one-dimensional microstrips, which results in a partially-connected analog combiner and hence brings in some performance loss compared to fully-connected analog combiners under the multiuser scenario. When utilizing conventional beam steering based on far-field assumptions, it is observed in Fig. 8(b) that user 2 achieves negligible rates, i.e., $R_2 \approx 0$, regardless of its distance from the transmit antenna. This is because, for far-field communications, conventional beam steering is unable to distinguish the two users with the same angular direction. Hence, in order to maximize the sum-rate, conventional beam steering allocates essentially all the transmit power to one user that has the better channel (i.e., smaller path loss).

In order to more explicitly illustrate the distinguishing ability of near-field beam focusing, in Fig. 9 we further show the normalized signal power of the signal transmitted to each user along the whole xz-plane, using a fully-digital antenna. The received signal power of each user at each point is normalized by the corresponding channel gain to remove the effect of the distance between the user and the BS. Fig. 9(a) shows the normalized signal power of user 1 at each point of near-field xz-plane, from which we can clearly see that the maximum normalized signal power is achieved at around the focusing point of user 1, while the minimum normalized signal power is achieved at around the focusing point of user 2. This result verifies our conclusion that near-field focusing cannot only enhance the signal strength at the focusing point, but also eliminate the co-channel interference to other users, even if the two users lies in the same angular direction. The same is observed in Fig. 9(b), which illustrates the normalized signal power of user 2 along the near-field xz-plane. From Figs. 8 and 9, we conclude that our proposed near-field beam focusing design allows to simultaneously communicate with multiple users located at the same angular direction, while such separation ability is not achievable for conventional beam steering.

Finally, we study a more general scenario with different number of users randomly located at the near-field xz-plane. Fig. 10 shows the achievable sum-rate versus the number of users M. Particularly, we successively add users to the near-field xz-plane. From Fig. 10, it is observed that the achievable sum-rate of the three different antenna architectures all monotonically increase with the number of users M when M is small. However, as the number of users M increases,

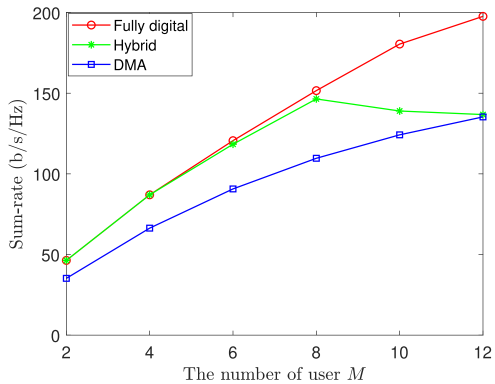


Fig. 10. Achievable sum-rate versus the number of users.

the growth trend of the sum-rate achieved by both fully-digital and DMA architectures becomes slower. This is because as the number of users M increases, the co-channel interference among users limits the increase of achievable sum-rate. On the other hand, it is observed that the achievable sum-rate of hybrid antenna starts to decrease when the number of users $M \geq 8$. This is because as the number of RF chains is constant, the hybrid beamforming solution cannot perfectly realize the optimal digital beamforming solution when the number of users is greater than half the number of RF chains.

V. Conclusion

In this work, we studied the potential of beam focusing for a near-field multi-user MIMO communication scenario. We first provided the mathematical model for the near-field wireless channels and the transmission pattern for the three different types of antenna architectures, including fully-digital, (phase shifters based-) hybrid and DMA architectures. We then formulated a near-field beam focusing problem for maximizing the achievable sum-rate. After that, we proposed efficient solutions based on the sum-rate maximization task for each of the antenna architectures. Numerical results demonstrated that beam focusing leads to an improved achievable rate in the near-field, and also has the potential of decreasing co-channel interference in multi-user communication scenarios. It is shown that the achievable sum-rate of hybrid architectures and DMAs is comparable to that of fully-digital architectures. In particular, DMAs are shown to achieve the most focused beams among all considered architectures with a given aperture for the single-user case, whereas for the near-field multi-user case, hybrid phase-shifter based antennas where each element can be connected to each RF chain are shown to achieve improved rates compared to DMAs.

Appendix

A. Proof of Theorem 1

For a fixed weighting matrix $\mathbf{Q}$, the digital precoding vector $\mathbf{w}$ that maximizes (26) is

$$\mathbf{w}^* = \sqrt{P_{\max}} \frac{\left(\mathbf{a}^H \mathbf{HQ}\right)^H}{\|\mathbf{a}^H \mathbf{HQ}\|}. \tag{A.1}$$

This implies that the maximal ratio transmission with maximum available power is the optimal digital precoding vector for any fixed $\mathbf{Q}$.

By substituting (A.1) into (26) with $\mathcal{Q}$ replaced by $\mathcal{F}$, we obtain the optimization problem

$$\max_{\mathbf{Q}} \ \left\|\mathbf{a}^H \mathbf{HQ}\right\|^2, \quad s.t. \ (13), \quad q_{i,l} \in \mathcal{F}, \ \forall i, l. \tag{A.2}$$

To drop the non-convex structure constraint on $\mathbf{Q}$, we rewrite the objective function of (A.2) as $\left\|\mathbf{a}^H \mathbf{HQ}\right\|^2 = \sum_{i=1}^{N_d} \left|\sum_{l=1}^{N_e} A_{i,l}(\mathbf{p}_m) \, e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{i,l}|} h_{i,l} q_{i,l}\right|^2$. By substituting this into (A.2), we obtain

$$\max_{\{q_{i,l}\}} \ \sum_{i=1}^{N_d} \left|\sum_{l=1}^{N_e} A_{i,l}(\mathbf{p}_m) \, e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{i,l}|} h_{i,l} q_{i,l}\right|^2$$
$$s.t. \ q_{i,l} \in \mathcal{F}, \quad \forall i, l. \tag{A.3}$$

The maximization problem in (A.3) can be decomposed into N_d subproblems, each with an identical structure. In particular, each subproblem individually designs the weighting coefficients of a single microstrip, with the ith subproblem given by

$$\max_{\{q_{i,l}\}} \ \left|\sum_{l=1}^{N_e} A_{i,l}(\mathbf{p}_m) \, e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{i,l}|} h_{i,l} q_{i,l}\right|^2$$
$$s.t. \ q_{i,l} \in \mathcal{F}, \quad \forall l. \tag{A.4}$$

Substituting the expression of $h_{i,l}$ in (12) into (A.4), we obtain

$$\max_{\{\psi_{i,l}\}} \ \left|\sum_{l=1}^{N_e} A_{i,l}(\mathbf{p}_m) \, e^{-\alpha_i \rho_{i,l}} \, e^{-\jmath k|\mathbf{p}_m - \mathbf{p}_{i,l}|} e^{-\jmath \beta_i \rho_{i,l}} e^{\jmath \psi_{i,l}}\right|^2. \tag{A.5}$$

Hence, according to the triangle inequality, the solution to problem (A.5) is $\psi_{i,l}^* = k|\mathbf{p}_m - \mathbf{p}_{i,l}| + \beta_i \rho_{i,l}$, for each i, l, thus proving the theorem.

B. Proof of Theorem 2

By using the fact that $\mathbf{x}^T \mathbf{Q} \mathbf{y} = (\mathbf{y}^T \otimes \mathbf{x}^T)\mathrm{Vec}(\mathbf{Q})$ is valid for arbitrary vectors $\mathbf{x}$, $\mathbf{y}$, and matrix $\mathbf{Q}$, we have

$$\left|\mathbf{a}_m^H \mathbf{HQ} \, \mathbf{w}_m\right|^2 = \left|(\mathbf{w}_m^T \otimes (\mathbf{a}_m^H \mathbf{H}))\mathrm{Vec}(\mathbf{Q})\right|^2, \tag{B.1}$$
$$\left|\mathbf{a}_m^H \mathbf{HQ} \, \mathbf{w}_j\right|^2 = \left|(\mathbf{w}_j^T \otimes (\mathbf{a}_m^H \mathbf{H}))\mathrm{Vec}(\mathbf{Q})\right|^2. \tag{B.2}$$

For brevity, we define $L = N_d^2 \times N_e$. Letting $\mathbf{z}_{m,m} = \left(\mathbf{w}_m^T \otimes (\mathbf{a}_m^H \mathbf{H})\right)^H \in \mathbb{C}^{L \times 1}$, $\mathbf{z}_{j,m} = \left(\mathbf{w}_j^T \otimes (\mathbf{a}_m^H \mathbf{H})\right)^H \in \mathbb{C}^{L \times 1}$, and $\mathbf{q} = \mathrm{Vec}\left(\mathbf{Q}\right) \in \mathbb{C}^{L \times 1}$, we can reformulate the problem (25) with fixed $\{\mathbf{w}_m\}$ as

$$\max_{\mathbf{q}} \ \sum_{m=1}^{M} \log_2 \left(1 + \frac{\left|\mathbf{q}^H \mathbf{z}_{m,m}\right|^2}{\sum_{j \neq m} \left|\mathbf{q}^H \mathbf{z}_{j,m}\right|^2 + \sigma^2}\right)$$
$$s.t. \ \begin{cases} |q_i| = 1 & \text{if } c_d^i N_e + 1 \leq c_m^i \leq (c_d^i + 1) N_e, \\ 0 & \text{otherwise} \end{cases}$$
$$\forall i = 1, \cdots, L, \tag{B.3}$$

where $c_d^i = \mathrm{div}(i, N_d N_e)$ and $c_m^i = \mathrm{mod}(i, N_d N_e)$ are the quotient and remainder of the division of i by $N_d N_e$, respectively.

It is easy to verify that the zero elements of $\mathbf{q}$ have no effect on the objective function of (B.3). Hence, (B.3) can be simplified as

$$\max_{\bar{\mathbf{q}}} \; f(\bar{\mathbf{q}}) = \sum_{m=1}^{M} \log_2 \left(1 + \frac{\left|\bar{\mathbf{q}}^H \bar{\mathbf{z}}_{m,m}\right|^2}{\sum_{j \neq m} \left|\bar{\mathbf{q}}^H \bar{\mathbf{z}}_{j,m}\right|^2 + \sigma^2} \right)$$
$$s.t. \; \bar{q}_l \in \mathcal{Q}, \quad \forall l \in \mathcal{A}_q, \tag{B.4}$$

where $\mathcal{A}_q$ denotes the set of all non-zero elements of $\mathbf{q}$, $\bar{\mathbf{q}}$ is the modified version of $\mathbf{q}$ obtained by removing all the zero elements of $\mathbf{q}$; and $\bar{\mathbf{z}}_{m,m}$ and $\bar{\mathbf{z}}_{j,m}$ are respectively the modified versions of $\mathbf{z}_{m,m}$ and $\mathbf{z}_{j,m}$, which are obtained by removing the elements having the same index as the zero elements of $\mathbf{q}$.

References

[1] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for multi-user MIMO communications with dynamic metasurface antennas," in *Proc. IEEE ICASSP*, Jun. 2021, pp. 4780–4784.

[2] J. Zhang *et al.*, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.

[3] F. Guidi and D. Dardari, "Radio positioning with EM processing of the spherical wavefront," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3571–3586, Jun. 2021.

[4] E. Bjornson and L. Sanguinetti, "Power scaling laws and near-field behaviors of massive MIMO and intelligent reflecting surfaces," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1306–1324, 2020.

[5] D. Dardari and N. Decarli, "Holographic communication using intelligent surfaces," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 35–41, Jun. 2021.

[6] E. Björnson, Ö. T. Demir, and L. Sanguinetti, "A primer on near-field beamforming for arrays and reconfigurable intelligent surfaces," 2021, *arXiv:2110.06661*.

[7] A. Guerra, F. Guidi, D. Dardari, and P. Djuric, "Near-field tracking with large antenna arrays: Fundamental limits and practical algorithms," *IEEE Trans. Signal Process.*, vol. 69, pp. 5723–5738, 2021.

[8] A. Elzanaty, A. Guerra, F. Guidi, and M. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Trans. Signal Process.*, vol. 69, pp. 5386–5402, 2021.

[9] P. Nepa and A. Buffi, "Near-field-focused microwave antennas: Near-field shaping and implementation," *IEEE Antennas Propag. Mag.*, vol. 59, no. 3, pp. 42–53, Jun. 2017.

[10] A. Buffi, P. Nepa, and G. Manara, "Design criteria for near-field-focused planar arrays," *IEEE Antennas Propag. Mag.*, vol. 54, no. 1, pp. 40–50, Feb. 2012.

[11] R. Liu and K. Wu, "Antenna array for amplitude and phase specified near-field multifocus," *IEEE Trans. Antennas Propag.*, vol. 67, no. 5, pp. 3140–3150, May 2019.

[12] K. Nishimori, N. Honma, T. Seki, and K. Hiraga, "On the transmission method for short-range MIMO communication," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1247–1251, Mar. 2011.

[13] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2526–2537, Nov. 2020.

[14] A. de Jesus Torres, L. Sanguinetti, and E. Björnson, "Near- and far-field communications with large intelligent surfaces," 2020, *arXiv:2011.13835*.

[15] W. Tang *et al.*, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, Jan. 2021.

[16] J. C. B. Garcia, A. Sibille, and M. Kamoun, "Reconfigurable intelligent surfaces: Bridging the gap between scattering and reflection," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2538–2547, Nov. 2020.

[17] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.

[18] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[19] S. S. Ioushua and Y. C. Eldar, "A family of hybrid analog–digital beamforming methods for massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, no. 12, pp. 3243–3257, Jun. 2019.

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

[21] 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, Oct. 2019.

[22] H. Wang *et al.*, "Dynamic metasurface antennas for MIMO-OFDM receivers with bit-limited ADCs," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2643–2659, Apr. 2021.

[23] I. Yoo, M. F. Imani, T. Sleasman, H. D. Pfister, and D. R. Smith, "Enhancing capacity of spatial multiplexing systems using reconfigurable cavity-backed metasurface antennas in clustered MIMO channels," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1070–1084, Feb. 2019.

[24] 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, Apr. 2021.

[25] X. Yu, D. Xu, and R. Schober, "MISO wireless communication systems via intelligent reflecting surfaces," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2019, pp. 735–740.

[26] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, May 2020.

[27] C. Pan *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[28] A. Abrardo, D. Dardari, M. Di Renzo, and X. Qian, "MIMO interference channels assisted by reconfigurable intelligent surfaces: Mutual coupling aware sum-rate optimization based on a mutual impedance channel model," *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2624–2628, Dec. 2021.

[29] 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 Antenn. Propag. Mag.*, vol. 54, no. 4, pp. 10–35, Jul. 2012.

[30] S. W. Ellingson, "Path loss in reconfigurable intelligent surface-enabled channels," in *Proc. IEEE 32nd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2021, pp. 829–835.

[31] N. Shlezinger and Y. C. Eldar, "On the spectral efficiency of noncooperative uplink massive MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1956–1971, Mar. 2019.

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

[33] T. Sleasman *et al.*, "Waveguide-fed tunable metamaterial element for dynamic apertures," *IEEE Antennas Wireless Propag. Lett.*, vol. 15, pp. 606–609, 2016.

[34] D. R. Smith, O. Yurduseven, L. P. Mancera, P. Bowen, and N. B. Kundtz, "Analysis of a waveguide-fed metasurface antenna," *Phys. Rev. A, Gen. Phys.*, vol. 8, no. 5, Nov. 2017, Art. no. 054048.

[35] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[36] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Feb. 2016.

[37] T. Gong, N. Shlezinger, S. S. Ioushua, M. Namer, Z. Yang, and Y. C. Eldar, "RF chain reduction for MIMO systems: A hardware prototype," *IEEE Syst. J.*, vol. 14, no. 4, pp. 5296–5307, Dec. 2020.

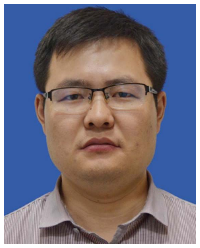
[38] J. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-CS-94-125, Aug. 1994.

[39] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton Press, 2009.

[40] V. R. Gowda, M. F. Imani, T. Sleasman, O. Yurduseven, and D. R. Smith, "Focusing microwaves in the Fresnel zone with a cavity-backed holographic metasurface," *IEEE Access*, vol. 6, pp. 12815–12824, 2018.

[41] I. F. Akyildiz and J. M. Jornet, "Realizing ultra-massive MIMO (1024×1024) communication in the (0.06–10) terahertz band," *Nano Commun. Netw.*, vol. 8, pp. 46–54, Jun. 2016.

[42] M. F. Imani, T. Sleasman, and D. R. Smith, "Two-dimensional dynamic metasurface apertures for computational microwave imaging," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 12, pp. 2299–2303, Dec. 2018.



Haiyang Zhang (Member, IEEE) received the B.S. degree in communication engineering from Lanzhou Jiaotong University, Lanzhou, China, in 2009, the M.S. degree in information and communication engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2012, and the Ph.D. degree in information and communication engineering from Southeast University, Nanjing, in 2017. From 2017 to 2020, he was a Post-Doctoral Research Fellow with the Singapore University of Technology and Design, Singapore. He is currently a Post-Doctoral Research Fellow with the Weizmann Institute of Science, Israel. His research interests include 6G near-field communications, learning and sampling theory, and physical-layer security.

Nir Shlezinger (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical and computer engineering from the Ben-Gurion University of the Negev, Israel, in 2011, 2013, and 2017, respectively. From 2017 to 2019, he was a Post-Doctoral Researcher at the Technion. From 2019 to 2020, he was a Post-Doctoral Researcher with the Weizmann Institute of Science, where he was awarded the FGS Prize for outstanding research achievements. He is currently an Assistant Professor with the School of Electrical and Computer Engineering, Ben-Gurion University of the Negev. His research interests include communications, information theory, signal processing, and machine learning.

Francesco Guidi (Member, IEEE) received the B.S. and M.S. degrees (both summa cum laude) in biomedical and in electronics and telecommunications engineering from the University of Bologna, Italy, in 2006 and 2009, respectively, and the joint Ph.D. degree in electronics, telecommunications and information technologies from Ecole Polytechnique ParisTech, France, and the University of Bologna in 2013. From Mid 2013 to Mid 2015, he was a Post-Doctoral Researcher with the University of Bologna. He is currently a Researcher with IEIIT-CNR, Italy. His research interests include RFID and radar technologies, joint antenna and channel characterization, signal processing, UWB, and mm-waves technologies. From 2015 to 2017, he was a recipient of an Individual European Marie Skłodowska-Curie Fellowship at French Atomic Energy Commission (CEA-LETI), Grenoble, France. He was a recipient of the Best Student Paper Award at the 2014 IEEE International Conference on Ultra-Wideband and the Best Paper Award at the 2021 IEEE International Conference on Autonomous Systems, Montreal, Canada. He has organized a number of special sessions or workshop at international conferences.

Davide Dardari (Senior Member, IEEE) is currently a Full Professor at the University of Bologna, Italy. He has been a Research Affiliate at the Massachusetts Institute of Technology, USA. He has published more than 250 technical papers and played several important roles in various national and European projects. His interests are in wireless communications, localization techniques, and distributed signal processing. He has received the IEEE Aerospace and Electronic Systems Society's M. Barry Carlton Award (2011) and the IEEE Communications Society's Fred W. Ellersick Prize (2012). He was the Chair of the Radio Communications Committee and Distinguished Lecturer (2018–2019) of the IEEE Communication Society. He was the Co-General Chair of the 2011 IEEE International Conference on Ultra-Wideband and a Co-Organizer of the IEEE International Workshop on Advances in Network Localization and Navigation (ANLN)—ICC 2013–2016 editions. He was also the TPC Chair of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2018), the TPC Co-Chair of the Wireless Communications Symposium of the 2007/2017 IEEE International Conference on Communications, and the TPC Co-Chair of the 2006 IEEE International Conference on Ultra-Wideband. He has served as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2006 to 2012 and a guest editor for several journals.

Mohammadreza F. Imani (Member, IEEE) received the B.S.E. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2007, and the M.S.E. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2010 and 2013, respectively. From 2014 to 2018, he has worked as a Post-Doctoral Associate with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA. From 2018 to August 2020, he was a Research Scientist at the Department of Electrical and Computer Engineering. In August 2020, he joined the Arizona State University School of Electrical, Computer, and Energy Engineering as an Assistant Professor. His research includes analytical and applied electromagnetics, metamaterials and metasurfaces, computational imaging and sensing, wireless power transfer, antenna analysis and synthesis, and MIMO communication systems.

Yonina C. Eldar (Fellow, IEEE) received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering from Tel-Aviv University, Tel-Aviv, Israel, 1995 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2002.

She is currently a Professor with the Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel. She was previously a Professor with the Department of Electrical Engineering, Technion, where she held the Edwards Chair in engineering. She is also a Visiting Professor with MIT, a Visiting Scientist with the Broad Institute, an Adjunct Professor with Duke University, and was a Visiting Professor with Stanford. She is the author of the book *Sampling Theory: Beyond Bandlimited Systems* and the coauthor of four other books published by Cambridge University Press. Her research interests include statistical signal processing, sampling theory and compressed sensing, learning and optimization methods, and their applications to biology, medical imaging, and optics.

Dr. Eldar was a member of the Young Israel Academy of Science and Humanities and the Israel Committee for Higher Education. She is a member of the IEEE Sensor Array and Multichannel Technical Committee and serves for several other IEEE committees. She is a member of the Israel Academy of Sciences and Humanities (elected 2017) and a EURASIP Fellow. She was a Horev Fellow of the Leaders in Science and Technology Program with the Technion and an Alon Fellow. She has received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award (2013), the IEEE/AESS Fred Nathanson Memorial Radar Award (2014), and the IEEE Kiyo Tomiyasu Award (2016). She received the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, the Wolf Foundation Krill Prize for Excellence in Scientific Research, the Henry Taub Prize for Excellence in Research (twice), the Hershel Rich Innovation Award (three times), the Award for Women with Distinguished Contributions, the Andre and Bella Meyer Lectureship, the Career Development Chair with the Technion, the Murieland David Jacknow Award for Excellence in Teaching, and the Technions Award for Excellence in Teaching (two times). She received several best paper awards and best demo awards together with her research students and colleagues, including the SIAM Outstanding Paper Prize, the UFFC Outstanding Paper Award, the Signal Processing Society Best Paper Award, and the *IET Circuits, Devices and Systems* Premium Award. She was selected as one of the 50 most influential women in Israel and Asia. She is also a Highly Cited Researcher. She was the co-chair and the technical co-chair of several international conferences and workshops. In the past, she was a Signal Processing Society Distinguished Lecturer, a member of the IEEE Signal Processing Theory and Methods and Bio Imaging Signal Processing Technical Committees, and served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, *EURASIP Journal on Advances in Signal Processing*, *SIAM Journal on Matrix Analysis and Applications*, and *SIAM Journal on Imaging Sciences*. She is the Editor-in-Chief for *Foundations and Trends in Signal Processing*.