Dynamic Metasurface Antennas for MIMO-OFDM Receivers With Bit-Limited ADCs

Hanqing Wang, *Member, IEEE*, Nir Shlezinger, *Member, IEEE*, Yonina C. Eldar, *Fellow, IEEE*, Shi Jin, *Senior Member, IEEE*, Mohammadreza F. Imani, *Member, IEEE*, Insang Yoo, *Graduate Student Member, IEEE*, and David R. Smith, *Senior Member, IEEE*

***Abstract*— The combination of orthogonal frequency modulation (OFDM) and multiple-input multiple-output (MIMO) techniques plays an important role in modern communication systems. In order to meet the growing throughput demands, future MIMO-OFDM receivers are expected to utilize a massive number of antennas, operate in dynamic environments, and explore high frequency bands, while satisfying strict constraints in terms of cost, power, and size. An emerging technology to realize massive MIMO receivers of reduced cost and power consumption is based on dynamic metasurface antennas (DMAs), which inherently implement controllable compression in acquisition. In this work we study the application of DMAs for MIMO-OFDM receivers operating with bit-constrained analog-to-digital converters (ADCs). We present a model for DMAs which accounts for the configurable frequency selective profile of its metamaterial elements, resulting in a spectrally flexible hybrid structure. We then exploit previous results in task-based quantization to show characterized the achievable OFDM recovery accuracy for a given DMA configuration in the presence of bit-constrained ADCs, and propose methods for adjusting the DMA parameters based on channel state information. Our numerical results demonstrate that by properly exploiting the spectral diversity of DMAs, notable performance gains are obtained over existing designs of conventional hybrid architectures, demonstrating the potential of DMAs for realizing high performance massive antenna arrays of reduced cost and power consumption.**

***Index Terms*— Metasurface antennas, bit-constrained analog-to-digital converter (ADC), multiple-input multiple-output-orthogonal frequency division multiplexing (MIMO-OFDM).**

Manuscript received April 22, 2020; revised September 9, 2020; accepted November 16, 2020. Date of publication November 26, 2020; date of current version April 16, 2021. The work of H. Wang and S. Jin was supported in part by the National Science Foundation of China (NSFC) for Distinguished Young Scholars of China with Grant 61625106 and the NSFC under Grant 61531011, and the Scientific Research Foundation of Graduate School of Southeast University under Grant YBJJ1759. The work of N. Shlezinger and Y. C. Eldar was supported by the Benoziyo Endowment Fund for the Advancement of Science, the Estate of Olga Klein – Astrachan, the European Union's Horizon 2020 research and innovation program under grant No. 646804-ERC-COG-BNYQ, and the Air Force Office of Scientific Research under grants No. FA9550-18-1-0187 and FA9550-18-1-0208. This article was presented in part at the 2020 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2020. The associate editor coordinating the review of this article and approving it for publication was S. Affes. *(Corresponding author: Shi Jin.)*

Hanqing Wang and Shi Jin are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: hqwanglyt@seu.edu.cn; jinshi@seu.edu.cn).

Nir Shlezinger is with the School of Electrical Engineering, Ben-Gurion University of the Negev, Be'er-Sheva 8410501, Israel (e-mail: nirshl@bgu.ac.il).

Yonina C. Eldar is with the Faculty of Math and CS, Weizmann Institute, Rehovot 7610001, Israel (e-mail: yonina@weizmann.ac.il).

Mohammadreza F. Imani is with the Department ECEE, Arizona State University, Tempe, AZ 85281 USA (e-mail: mohammadreza.imani@asu.edu).

Insang Yoo and David R. Smith are with the Department of ECE, Duke University, Durham, NC 27708 USA (e-mail: insang.yoo@duke.edu; drsmith@duke.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCOMM.2020.3040761.

Digital Object Identifier 10.1109/TCOMM.2020.3040761

I. Introduction

WIRELESS networks are subject to constantly growing throughput demands. To satisfy these requirements, cellular base stations (BSs) are equipped with a large number of antennas while serving multiple remote users [1], utilizing wideband orthogonal frequency division multiplexing (OFDM) transmissions. Such multi-user (MU) multiple-input multiple-output (MIMO) OFDM architectures are capable of reliably providing increased data rates to the service users [2].

In addition to their performance requirements, BSs are expected to be cost efficient, operate under strict power constraints. A major challenge associated with realizing such MU-MIMO-OFDM systems stems from the increased cost of analog-to-digital convertors (ADCs) [3], which allow the analog signals observed by each antenna to be processed in digital. The power usage of an ADC is related to the signal bandwidth and the number of bits used for digital representation [4]. Thus, when the number of antennas and ADCs operating at wide bands is large, limiting the number of quantization bits, is crucial to keep feasible cost and power usage [5].

Focusing on uplink communications, quantization constraints imply that the BS cannot process the channel output directly but rather a discretized distorted representation of it. The distortion induced by the continuous-to-discrete quantization mapping degrades the ability to extract the desired information, such as recovering the transmitted signal, from the observed channel output. An attractive strategy to mitigate the effect of quantization error is to incorporate pre-quantization processing in analog resulting in a hybrid architecture. Jointly designing the analog processing along with the quantization rule and the digital mapping, referred to as *task-based quantization*, was shown to facilitate recovery of the underlying information in the digital domain [6]–[9]. An advantage of such hybrid MIMO receivers, originally proposed as a method to decrease the number of radio frequency (RF) chains [10]–[12], is that they can

be used to reduce the number of quantized samples, and accordingly, the number of ADCs, compared to assigning an ADC to each antenna [13]. Nonetheless, such designs require an additional dedicated hardware for the pre-quantization mapping [14], and the ability to adapt its parameters based on the channel conditions is typically limited and dictated by the analog components [10], [11].

An emerging receiver architecture which implements adjustable analog combining in the hardware level is based on dynamic metasurface antennas (DMAs) [15], [16], which consist of a set of microstrips, each embedded with configurable radiating metamaterial elements [17], [18]. Recent years have witnessed a growing interest in the application of metasurfaces as reflecting surfaces for wireless communications [19]–[27]. In such applications, a metasurface is placed in a physical location where it can aid the BS by reflecting and steering the transmitted waveforms, such that the reconfigurability of the metasurface is exploited to improve coverage and propagation conditions. An additional emerging application of metasurfaces in wireless communications, which is the focus in the current work, is to exploit their controllable radiation and reception patterns as antennas, i.e., as active transmitting and receiving devices rather than passive configurable reflectors, as recently studied in [28]–[31]. Such antenna structures typically consume much less power and cost less than architectures based on standard arrays [32], facilitating the implementation of a large number of tunable elements in a given physical area [33]. In the context of wireless communications, it is shown in [28], [29] that the achievable rate when utilizing DMAs without quantization constraints is comparable to using ideal antenna arrays. The potential of DMAs for realizing massive antenna arrays combined with the need to operate with low resolution quantization motivates the study of bit-constrained DMA-based BSs, which is the focus of the current work.

Here, we study uplink MU-MIMO-OFDM communications in which a bit-constrained BS is equipped with a DMA. We first extend the model formulated in [28], which was built upon approximations of the DMA properties proposed in [15] that hold for narrowband signals, to faithfully capture the reconfigurable frequency selective nature of DMAs in wideband setups, such as OFDM systems. Then, we show how the extended DMA model can be incorporated into the MU-MIMO-OFDM model, resulting in a form of hybrid receiver. However, while conventional hybrid architectures require a dedicated analog hardware whose mapping is typically frequency flat [10], [11], [14], DMAs implement a controllable frequency selective profile as an inherent byproduct of their structure. We use this model to formulate the following problem: How can the dynamic properties of DMAs, i.e., their configurable reception parameters, be exploited to facilitate recovering transmitted OFDM signals using low resolution ADCs?

Based on this formulation, we cast the problem as a task-based quantization setup, in which the task is to accurately recover the transmitted OFDM symbols. Using this framework, we derive a scheme for jointly optimizing the DMA weights along with the dynamic range of the ADCs and the digital processing, under a given bit constraint. Our proposed method consists of two algorithms. The first of these two methods is based on the approximations of DMA characteristics proposed in [15], which ignore the spectral flexibility of the elements, tuning a frequency flat hybrid receiver most suitable for narrowband signals; The latter makes usage of the full frequency selective profile of the metamaterial elements, and is thus preferable for wideband transmissions. Both proposed algorithms exploit the unique structure of the hybrid system which arises from the task-based quantization framework, while building upon the dynamic nature of DMAs, which allows to set their parameters in light of the channel conditions. We also discuss how these task-based quantization schemes can facilitate the design of common conventional hybrid receivers, such as partially connected phase shifter networks operating with low resolution ADCs [34].

The performance of the resulting receivers applying these proposed algorithms, in terms of OFDM signal recovery accuracy and uncoded bit error rate (BER), are evaluated in a simulation study. Our numerical results demonstrate the ability of bit-constrained DMAs to achieve notable performance gains over conventional hybrid architectures designed using previously proposed methods. These performance gains, which arise from the combination of task-based quantization tools and the spectral flexibility of DMAs, add to their practical benefits over conventional hybrid structures, which follow from the fact that DMAs do not require additional dedicated hardware for implementing their analog processing. Our numerical results also demonstrate that the proposed algorithms can be utilized to design conventional phase shifter based hybrid architectures of improved performance compared to previous methods.

The rest of this paper is organized as follows: Section II presents the mathematical model of DMA-based receivers in bit-constrained MIMO-OFDM systems, introducing the formulation of the configurable frequency selectivity of the metamaterial elements and how it is incorporated in the considered communication setup. Section III details the proposed DMA configuration methods, while Section IV discusses the properties, advantages, and challenges which arise from the proposed receiver architecture. Numerical examples are presented in Section V. Finally, Section VI concludes the paper. Proofs of the results stated in the paper are detailed in the appendix.

Throughout the paper, we use boldface lower-case letters for vectors, e.g., $\mathbf{x}$; boldface upper-case letters for matrices, e.g., $\boldsymbol{M}$; and calligraphic letters are used for sets, e.g., $\mathcal{C}$ and $\mathcal{Z}$ are the set of complex numbers and integers, respectively. For a vector, e.g., $\mathbf{x}$, $(\mathbf{x})_i$ denotes the i-th entry of it. For a matrix, e.g., $\mathbf{M}$, $(\mathbf{M})_{i,:}$ and $(\mathbf{M})_{:,j}$ denote the i-th row and j-th column of it, respectively. The ℓ_2 norm, Frobenius norm, Kronecker product, Hadamard product, transpose, Hermitian transpose, conjugate, and trace are denoted by $\|\cdot\|$, $\|\cdot\|_F$, $\otimes$, $\odot$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, and $\mathrm{tr}[\cdot]$, respectively. We use $\mathbf{I}_n$ and $\mathbf{O}_n$ for the $n \times n$ identity matrix and all-zero matrix, respectively. Finally, $\mathrm{blkdiag}(\mathbf{A}_1, \ldots, \mathbf{A}_n)$ is a block diagonal matrix comprised of $\mathbf{A}_1, \ldots, \mathbf{A}_n$.

II. System Model

In this section, we present the mathematical model for bit-constrained DMA-based uplink MU-MIMO-OFDM systems. We begin with modelling the DMA operation in Subsection II-A, which extends the one proposed in [28] to fully capture the frequency selective characteristics of metasurface antennas. Then, in Subsection II-B, we formulate the input-output relationship of bit-constrained DMA-based uplink MU-MIMO-OFDM systems.

A. Dynamic Metasurface Antennas

Metamaterials are a class of artificial materials whose physical properties, such as their permittivity and permeability, can be externally configured to achieve some desired electromagnetic properties [35]. Metamaterial elements stacked in surface configurations are referred to as metasurfaces. These two-dimensional structures can be tuned element-wisely, allowing the metasurface to carry out desired operations, such as radiation, reflection, beamforming, and reception of propagating waves [36]. In particular, metasurfaces can be utilized as antennas by incorporating such surfaces on top of a guiding structure. A simple and common metasurface antenna architecture is comprised of a set of microstrips, each consisting of a multitude of sub-wavelength, frequency-dependent resonant metamaterial radiating elements [37], whose radiation properties are dynamically adjustable. A larger antenna array can be thus formed by increasing the number of microstrips, or alternatively, by tiling several such metasurface antennas together.

When used as a receive antenna array, the signals observed by the elements are combined at a single output port for each microstrip, feeding an RF chain and an ADC with Nyquist rate sampling. An illustration of signals captured using a single microstrip is depicted in Fig. 1. The relationship between the observed signals and the microstrip output is dictated by the following properties:

P1 Each element acts as resonant circuit, whose frequency response is described by the Lorentzian form [15], [16]:

$$q(f) = \frac{F \cdot f^2}{(f^R)^2 - f^2 - j\kappa f}, \tag{1}$$

where $F > 0$ is the oscillator strength, $f^R > 0$ is the resonance frequency, and $\kappa \triangleq \frac{f^R}{2\chi} > 0$ is the damping factor with χ being the quality factor. In DMAs, these parameters can be tuned by external control for each element individually [17].

P2 Since the output port is located on the edge of the microstrip while the elements are uniformly placed along it, each signal which propagates from an element to that port undergoes a different path, and thus accumulates a different delay, depending on the element location. Let β be the wavenumber along the microstrip and ρ_l characterize the relative location of the lth element, which is proportional to the distance between the output port and the lth element. The propagation of the received signal on the l-th element inside the microstrip can be modeled as a filter whose frequency response consists of an attenuation term dictated by the waveguide characteristic impedance, and a phase shift term typically given by $e^{-j\beta\rho_l}$.

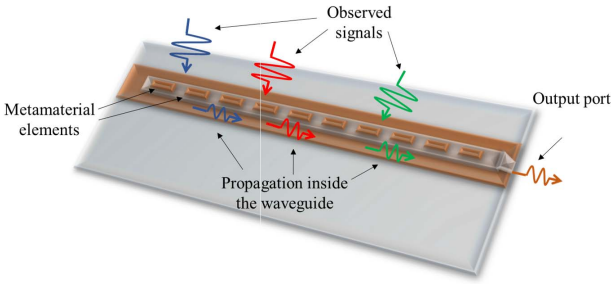


Fig. 1. Illustration of signal reception using a microstrip.

The frequency response of the metamaterial elements in *P1* is often approximated in the literature to be *frequency flat*, i.e., $q(f) \equiv q$ for each considered f. This approximation, which facilitates the configuration of the DMA elements, holds for narrowband signalling. Under this approximation, the ability to externally control the response of each element is typically modeled as setting $q \in [a_{\min}, a_{\max}]$ for some $0 < a_{\min} < a_{\max}$, referred to as *amplitude-only weights* [15, Sec. III-A], or alternatively, $q \in \{\frac{j+e^{j\phi}}{2} | \phi \in [0, 2\pi]\}$, referred to as *Lorentzian-constrained phase weights* [15, Sec. III-D]. Our previous work [28] used this frequency invariance assumption to analyze the achievable rate of uplink DMA-based MIMO systems. However, as the current work considers wideband OFDM signals, we further consider DMA configuration based on the general frequency selective model in (1), rather than its simplified narrowband approximation. In this model, one can control the resonance frequency f^R, the damping factor κ, and the oscillator strength F individually for each element, allowing each of the DMA elements to exhibit a variety of different frequency selective profiles, as well as frequency flat ones.

As an example, we depict in Figs. 2-3 the frequency response of a single element with magnitude normalized to unity at resonance, i.e., $\frac{q(f)}{q(f^R)}$, in which the quality factor $\chi = f^R/2\kappa$ is fixed to 50. The frequency response is evaluated for several different resonance frequencies f^R, focusing on the frequency band around 1.9 GHz. We note in Figs. 2-3 that each element approximates a bandpass filter, by setting its resonance frequency to be within the observed bandwidth, or alternatively, a filter with a frequency monotonic and even a frequency flat profile, when f^R is outside of the band of interest. In the sequel we show that this frequency selectivity can be exploited to facilitate OFDM signal recovery in the presence of quantized measurements.

Processing of the DMA output is carried out in digital baseband. Therefore, we next formulate the resulting model in discrete-time. Consider a DMA with $N \triangleq N_d \cdot N_e$ tunable metamaterial elements, where N_d and N_e are the number of microstrips and elements in each microstrip, respectively. Let $y_{i,l}[t]$ denote the equivalent baseband signal received from the wireless channel on metamaterial element l of microstrip i at time slot $t \in \{0, 1, \ldots, M-1\} \triangleq \mathcal{M}$, where M is the transmission block size, and let $y_{i,l}(\omega)$ be its discrete-time Fourier

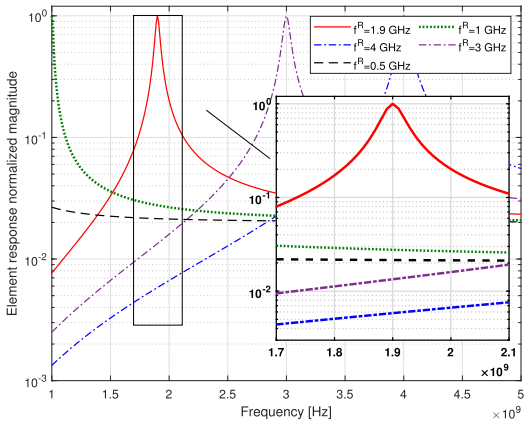


Fig. 2. Element response magnitude vs. frequency.

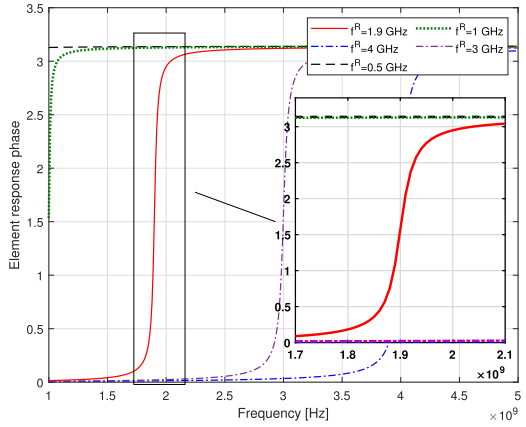


Fig. 3. Element response phase vs. frequency.

transform (DTFT). The frequency domain representation of the ith microstrip output $z_i(\omega)$ can be written as

$$z_i(\omega) = \sum_{l=1}^{N_e} h_{i,l}(\omega) q_{i,l}(\omega) y_{i,l}(\omega), \tag{2}$$

where $h_{i,l}(\omega)$ characterizes the effect of the signal propagation inside the microstrip *P2, while* $q_{i,l}(\omega)$ *denotes the tunable weight of the* l*th element of the* i*th microstrip. Based on property* P1, this frequency dependent profile, which represents the DTFT equivalent of (1), is given by

$$q_{i,l}(\omega) = \frac{F_{i,l}\Omega^2(\omega)}{(\Omega_{i,l}^R)^2 - \Omega^2(\omega) - j\Omega(\omega)\kappa_{i,l}}. \tag{3}$$

In the above equation, $F_{i,l}$, $\Omega_{i,l}^R$, and $\kappa_{i,l}$ are the configurable oscillator strength, angular resonance frequency, and damping factor of the lth element of microstrip i, respectively, where $\kappa_{i,l}$ is tuned equivalently via the configuration of the quality factor $\chi_{i,l}$. Let $\Omega(\omega)$ denote the analogous angular frequency corresponding to the continuous-time frequency ω, i.e., how the bandwidth of interest in continuous-time is mapped into angular frequencies of the DTFT of the sampled signal. This mapping is dictated by the carrier frequency f_c and the sampling rate f_s, and can be written as $\Omega(\omega) = 2\pi f_c + \omega f_s$ for $|\omega| < \pi$. An illustration of the system is given in Fig. 4.

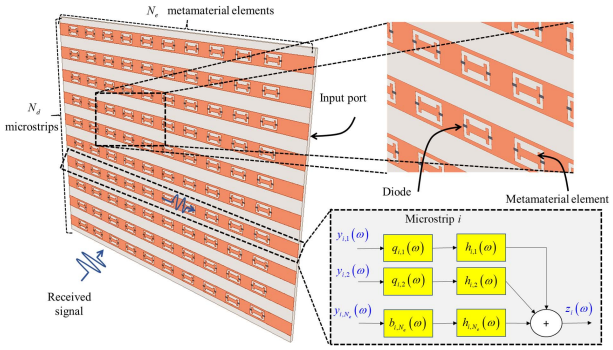


Fig. 4. Illustration of signal reception and propagation in DMA.

Next, we express the DMA model in (2) compactly in vector form. To that aim, we define a set of $N \times N$ diagonal matrices $\mathbf{H}(\omega)$ whose $[(i-1)N_e + l]$th diagonal element is $h_{i,l}(\omega)$. We also define $\mathbf{y}(\omega) \in \mathcal{C}^N$ as the vector comprised of the frequency domain received signal of the complete array such that its $[(i-1)N_e + l]$th element is $y_{i,l}(\omega)$. Using these definitions and letting $\mathbf{z}(\omega) \in \mathcal{C}^{N_d}$ be the DTFT of the DMA output, where $(\mathbf{z}(\omega))_i = z_i(\omega)$, the frequency domain formulation of the DMA operation can be formulated as

$$\mathbf{z}(\omega) = \mathbf{Q}(\omega)\mathbf{H}(\omega)\mathbf{y}(\omega), \tag{4}$$

where the matrix $\mathbf{Q}(\omega) \in \mathcal{C}^{N_d \times N}$ collects tunable weights of all DMA elements. For each $l \in \{1, 2, \cdots N_e\} \triangleq \mathcal{N}_e$, $i, k \in \{1, 2, \cdots N_d\} \triangleq \mathcal{N}_d$, the entries of $\mathbf{Q}(\omega)$ are

$$(\mathbf{Q}(\omega))_{k,(i-1)N_e+l} = \begin{cases} q_{i,l}(\omega) \in \mathcal{W} & i = k, \\ 0 & i \neq k. \end{cases} \tag{5}$$

where $\mathcal{W}$ is the set of all possible values for the coefficients $q_{i,l}(\omega)$. Throughout this paper, we denote $\mathcal{W}_{\text{DMA}}$ as the set of matrices satisfying the above constraints. By its definition, $\mathbf{Q}^H(\omega)$ is block diagonal with diagonal blocks being vectors denoted by $\mathbf{q}_i(\omega) = [q_{i,1}(\omega), q_{i,2}(\omega), \cdots, q_{i,N_e}(\omega)]^H \in \mathcal{C}^{N_e \times 1}$ for $i \in \mathcal{N}_d$. The expression (4) is then utilized to formulate bit-constrained DMA-based MU-MIMO-OFDM systems, as detailed in the next subsection.

B. Received Signal Model and Problem Formulation

In this paper, we consider the uplink scenario of a single-cell MU-MIMO-OFDM system. Here, the BS is equipped with a DMA consisting of N_d microstrips with a total of $N = N_e \cdot N_d$ metamaterial elements, and serves K single-antenna users. The users simultaneously transmit OFDM symbols with M subcarriers each. Due to power or memory constraints, the BS processes coarsely quantized samples of the DMA output obtained using ADCs, modeled as identical uniform scalar quantizers with resolution b. In particular, we focus on the recovery of the transmitted OFDM symbols in a hybrid manner by jointly configuring the DMA weights along with a digital linear filter applied to the ADCs outputs.

We consider a frequency-selective wireless channel, which follows a tapped delay line model with L_G taps, represented by a set of $N \times K$ matrices $\{\mathbf{G}[\tau]\}_{\tau=0}^{L_G-1}$. Let $\mathbf{s}[t] \in \mathcal{C}^{K \times 1}$ be the OFDM symbols transmitted by each user at the t-th

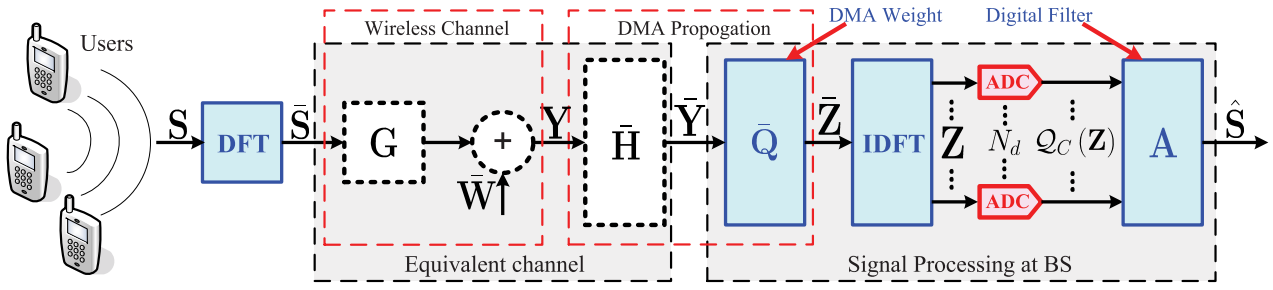


Fig. 5. Illustration of signal processing procedure. While the propagation inside the waveguide ($\bar{\mathbf{H}}$) takes place after the signals are captured by the element ($\bar{\mathbf{Q}}$), it can be treated as part of the equivalent channel. The configurable receiver parameters are the digital filter $\mathbf{A}$, the DMA weights $\bar{\mathbf{Q}}$, and the ADC support γ.

time slot, assumed to be i.i.d. and with covariance matrix $\mathbf{I}_K$. We denote the DTFT of $\{\mathbf{G}[\tau]\}_{\tau=0}^{L_G-1}$ and $\{\mathbf{s}[t]\}_{t=0}^{M-1}$ as $\mathbf{G}(\omega)$ and $\mathbf{s}(\omega)$, respectively. Then, the DTFT representation of the channel output vector $\mathbf{y}(\omega)$ is expressed as

$$\mathbf{y}(\omega) = \mathbf{G}(\omega)\mathbf{s}(\omega) + \mathbf{w}(\omega), \tag{6}$$

where $\mathbf{w}(\omega) \in \mathcal{C}^{N\times 1}$ is the DTFT of the additive noise vectors $\{\mathbf{w}[t]\}_{t=0}^{M-1}$, which are independent of $\mathbf{s}[t]$ and follow a zero-mean proper-complex Gaussian distribution with covariance $\mathbf{C}_W$. Since the elements in every microstrip are commonly sub-wavelength spaced, they are typically spatially correlated, and thus $\mathbf{C}_W$ is not restricted to be diagonal. The combined effect of the wireless channel and the propagation inside the microstrips in the frequency domain can be represented by the equivalent channel $\hat{\mathbf{G}}(\omega) = \mathbf{H}(\omega)\mathbf{G}(\omega) \in \mathcal{C}^{N\times K}$.

We assume the cyclic prefix (CP) length is larger than the memory length of the equivalent channel $\hat{\mathbf{G}}(\omega)$. Thus, the frequency response of all considered signals is fully captured by its M-point discrete Fourier transform (DFT), i.e., the DTFT in angular frequencies $\{\omega_m \triangleq \frac{2\pi m}{M}\}_{m\in\mathcal{M}}$. For brevity, we henceforth define $\mathbf{q}_{m,i} \triangleq \mathbf{q}_i(\omega_m)$, $\hat{\mathbf{G}}_m \triangleq \hat{\mathbf{G}}(\omega_m)$, and similarly define $\mathbf{Q}_m \in \mathcal{C}^{N_d\times N}$, $\mathbf{H}_m \in \mathcal{C}^{N\times N}$, $\mathbf{z}_m \in \mathcal{C}^{N_d\times 1}$, $\mathbf{s}_m \in \mathcal{C}^{K\times 1}$ and $\mathbf{w}_m \in \mathcal{C}^{N\times 1}$. Using these notations and substituting (4) into (6) for all M subchannels, the input-output relationship of an OFDM block after cyclic prefix (CP) removal in the frequency domain can be written in matrix form via

$$\bar{\mathbf{Z}} = \bar{\mathbf{Q}}\bar{\mathbf{G}}\bar{\mathbf{S}} + \bar{\mathbf{Q}}\bar{\mathbf{H}}\bar{\mathbf{W}}, \tag{7}$$

where $\bar{\mathbf{Z}} \triangleq [\mathbf{z}_0^T, \mathbf{z}_1^T, \cdots, \mathbf{z}_{M-1}^T]^T$, $\bar{\mathbf{S}} \triangleq [\mathbf{s}_0^T, \mathbf{s}_1^T, \cdots, \mathbf{s}_{M-1}^T]^T$, and $\bar{\mathbf{W}} \triangleq [\mathbf{w}_0^T, \mathbf{w}_1^T, \cdots, \mathbf{w}_{M-1}^T]^T$ denote the vertically concatenated vectors of the DMA output, transmitted signal and additive noise observed on each subchannel, respectively; $\bar{\mathbf{Q}} \triangleq \mathrm{blkdiag}\left(\mathbf{Q}_0, \mathbf{Q}_1, \cdots, \mathbf{Q}_{M-1}\right)$, $\bar{\mathbf{G}} \triangleq \mathrm{blkdiag}\big(\hat{\mathbf{G}}_0, \hat{\mathbf{G}}_1, \cdots, \hat{\mathbf{G}}_{M-1}\big)$ and $\bar{\mathbf{H}} \triangleq \mathrm{blkdiag}\left(\mathbf{H}_0, \mathbf{H}_1, \cdots, \mathbf{H}_{M-1}\right)$ denote the block diagonal formulation of DMA weights, equivalent channel and DMA propagation characterization on each subchannel, respectively. The expression (7) models the DMA output, which is fed to the ADCs, in the frequency domain. Since the ADC operation is formulated in the time domain, we transform $\bar{\mathbf{Z}}$ into the time domain by multiplying it with $\mathbf{V}_1 \triangleq \left(\mathbf{F}_M^H \otimes \mathbf{I}_{N_d}\right)$, where $\mathbf{F}_M$ is the $M \times M$ normalized DFT matrix. By defining $\mathbf{V}_2 \triangleq (\mathbf{F}_M \otimes \mathbf{I}_K)$, $\mathbf{S} = [\mathbf{s}[0]^T, \mathbf{s}[1]^T, \cdots, \mathbf{s}[M-1]^T]^T$, and $\mathbf{Z} = [\mathbf{z}[0]^T, \mathbf{z}[1]^T, \cdots, \mathbf{z}[M-1]^T]^T$, the resulting input of ADC is

$$\mathbf{Z} = \left(\mathbf{F}_M^H \otimes \mathbf{I}_{N_d}\right)\bar{\mathbf{Z}} = \mathbf{V}_1\bar{\mathbf{Q}}\bar{\mathbf{G}}\mathbf{V}_2\mathbf{S} + \mathbf{V}_1\bar{\mathbf{Q}}\bar{\mathbf{H}}\bar{\mathbf{W}}. \tag{8}$$

The ADC inputs $\{\mathbf{z}[t]\}_{t=0}^{M-1}$ are quantized using N_d identical pairs of ADCs which discretize the real and imaginary parts of each analog input sample independently. We denote complex-valued quantizer by $\mathcal{Q}_C(x+jy) = \mathcal{Q}(x) + j\mathcal{Q}(y)$, where $\mathcal{Q}(\cdot)$ is the uniform real-valued quantization operator:

$$\mathcal{Q}(x) = \begin{cases} -\gamma + \frac{2\gamma}{b}\left(l + \frac{1}{2}\right) & x - l\cdot\frac{2\gamma}{b} + \gamma \in \left[0, \frac{2\gamma}{b}\right], \\ & l \in \{0, 1, \ldots, b-1\}, \\ \mathrm{sign}\left(x\right)\left(\gamma - \frac{\gamma}{b}\right) & |x| > \gamma, \end{cases} \tag{9}$$

where γ is the dynamic range of the quantizer and b denotes its resolution, namely, the total number of output levels. The recovered symbols are obtained by linearly processing the quantized DMA outputs. Denoting $\mathbf{A} \in \mathcal{C}^{MK\times MN_d}$ as the digital linear processing matrix, then the resulting estimate of $\mathbf{S}$ is expressed as

$$\hat{\mathbf{S}} = \mathbf{A}\mathcal{Q}_C\left(\mathbf{Z}\right). \tag{10}$$

An illustration of the overall system is depicted in Fig. 5. It is noted that resulting formulation models the processing of the received signal as a bit-constrained hybrid system with frequency selective analog combining, represented by the weights $\bar{\mathbf{Q}}$. As such, the system in Fig. 5 specializes a broad range of conventional hybrid receiver architectures, in which the analog processing is frequency flat, e.g., [10], [11]. For example, this model specializes bit-constrained hybrid receivers with partially-connected phase shifter networks, as considered in [34], by fixing $\mathbf{H}(\omega)$ to be the identity matrix, i.e., canceling the effect of the propagation inside the microstrips *P2*, while restricting the elements $q_{i,l}(\omega)$ of $\bar{\mathbf{Q}}$ to be independent of ω and have a unit magnitude, namely, $q_{i,l}(\omega) \equiv e^{j\phi_{i,l}}$ for some $\phi_{i,l} \in [0, 2\pi)$ for all $(i,l) \in \mathcal{N}_d \times \mathcal{N}_e$. Nonetheless, unlike conventional architectures, the controllable frequency selective profile of the metamaterial elements also allows to realize different analog combining in each frequency bin in a coupled manner dictated by the Lorentzian response illustrated in Figs. 2-3.

Problem Formulation - Our aim is to jointly design the DMA weights $\bar{\mathbf{Q}}$, the digital linear filter $\mathbf{A}$, and the ADC dynamic range γ to minimize the mean-squared error (MSE)

$\mathbb{E}\{\|\mathbf{S}-\hat{\mathbf{S}}\|^2\}$, namely to produce an accurate estimate of $\mathbf{S}$. The estimator $\hat{\mathbf{S}}$ can be then used for the hard or soft decision of each transmitted symbol, taking values in some discrete constellation. The joint design problem at hand is challenging due to its two unique characteristics. These are the non-linear nature of the mapping induced by low-resolution ADCs, and the non-conventional feasible set of analog processing using DMAs. As discussed above, the analog processing of DMAs is notably different from analog precoders commonly studied in the literature, based on phase shifters and adders [10]–[12]. Furthermore, the resulting problem is fundamentally different from traditional digital-only designs, which operate only in the digital domain without including pre-quantization analog mapping [38]–[40]. In the following we tackle these challenges by utilizing the task-based quantization framework [6], which studies the design of hybrid analog/digital architectures for recovering information embedded in multivariate analog measurements. This strategy is known to allow achieving improved performance over digital-only designs by exploiting the joint analog/digital architecture to facilitate accurate recovery of the desired information under bit constraints [9]. However, existing task-based quantization results assume arbitrary linear analog processing. Our study thus involves introducing dedicated optimization techniques for handling the constraints representing the physical properties of DMAs, as detailed in the following.

III. DMA-Based Receiver Design

In this section, we propose a receiver design in light of the signal recovery MSE minimization. The primary problem of joint configuration of the DMA weights $\bar{\mathbf{Q}}$ and the digital linear filter $\mathbf{A}$ aiming to minimize MSE of the estimate (10) is basically intractable. The main challenge is that the non-linearity of low-resolution ADCs and the constraints on the structure and value of the DMA weight matrices prevents explicitly characterizing the minimal achievable MSE. To tackle this difficulty, we first characterize the non-linearity induced by the ADCs based on the principle of task-based quatization [6], [9] in Subsection III-A, and accordingly characterize the achievable OFDM recovery accuracy for given DMA configuration. Using the resulting formulation, we formulate the joint design of the DMA weights and the digital filter into a tractable optimization problem. Then, in Subsection III-B we translate the problem into a quadratic form using the *matrix quadratic transform*, eliminating the constraints imposed by the DMA structure for tractability. Finally, we propose an alternating algorithm for configuring DMAs under different types of feasible constraints on their weight matrix in Subsection III-C.

A. Task-Based Quantization Formulation

The goal of the bit-constrained receiver is to linearly process and discretize the wireless channel output for the task of recovering the transmitted OFDM symbols. The formulations of the DMA output in (8) and the estimated symbols in (10) indicate that the DMA induces a configurable pre-quantization mapping, represented by the matrix $\mathbf{V}_1\bar{\mathbf{Q}}$, while the filter $\mathbf{A}$ corresponds to the digital post-quantization processing. Therefore, the joint design of the DMA weights and digital filter can be treated as a task-based quantization setup, in which a hybrid receiver comprised of analog and digital processing are jointly optimized to facilitate the recovery of a desired vector. This formulation enables characterizing the achievable OFDM recovery accuracy for a given DMA configuration, and further facilitates the joint optimization of the quantization rule along with its pre and post quantization mappings [6]–[9], [13].

Following this framework, we assume the scalar quantizers (9) satisfy the statistical properties of dithered quantization [41], for which the quantizer output can be represented as the quantizer input corrupted by an additive zero-mean white noise uncorrelated with the quantizer input. This favorable property is approximately satisfied in uniform quantization without dithering for a broad range of input distributions [42]. In addition, the uniform quantizer is designed to guarantee that the ADCs are not overloaded, namely, the magnitudes of the real and imaginary parts of $\mathbf{Z}$ are not larger than the dynamic range γ, with sufficiently large probability. Accordingly, we set γ to be some multiple $\eta > 0$ of the maximal standard deviation of the ADC inputs, which is mathematically expressed as $\gamma^2 = \eta^2 \max_{t\in\mathcal{M}} \max_{i\in\mathcal{N}_d} \mathbb{E}\{|(\mathbf{z}[t])_i|^2\}$, such that the overload probability is bounded via Chebyshev's inequality [43, Pg. 64]. For the considered received signal model detailed in Section II, the resulting expression for γ is given in the following lemma:

Lemma 1: Let $\mathbf{E}_i \triangleq (\mathbf{I}_N)_{:,(i-1)\cdot N_e+1:i\cdot N_e}$ *be a matrix representing the placement of the nonzero elements into* $\left(\mathbf{Q}_m^H\right)_{:,i}$, *and define* $\boldsymbol{\Upsilon}_m \triangleq \hat{\mathbf{G}}_m\hat{\mathbf{G}}_m^H + \mathbf{H}_m\mathbf{C}_W\mathbf{H}_m^H$ *as the covariance of the equivalent channel output at frequency bin* $m \in \mathcal{M}$. *Then, for the received OFDM signal model in* (8)*, the dynamic range* γ *is set as*

$$\gamma^2 = \eta^2 \max_{i\in\mathcal{N}_d} \frac{1}{M} \sum_{m\in\mathcal{M}} \mathbf{q}_{m,i}^H \mathbf{E}_i^H \boldsymbol{\Upsilon}_m \mathbf{E}_i \mathbf{q}_{m,i}, \tag{11}$$

Proof: See Appendix A. ∎

When the quantizers are not overloaded, i.e., for each $l \in \{1, 2, \cdots, MN_d\}$ it holds that $\Pr\left(|(\mathbf{z}[t])_l| > \gamma\right) \approx 0$, the additive quantization noise has zero-mean entries with variance $\sigma_q^2 \triangleq \frac{4\gamma^2}{3b^2}$ [41, Theorem 2]. Furthermore, for any estimate $\hat{\mathbf{S}}$, it follows from the orthogonality principle that the MSE $\mathbb{E}\{\|\mathbf{S}-\hat{\mathbf{S}}\|^2\}$ equals the error of the linear minimal MSE (LMMSE) estimate of $\mathbf{S}$ from $\bar{\mathbf{Y}} \triangleq \bar{\mathbf{G}}\mathbf{V}_2\mathbf{S} + \bar{\mathbf{H}}\bar{\mathbf{W}}$ (denoted by $\tilde{\mathbf{S}}$), namely $e^{\mathrm{o}} \triangleq \mathbb{E}\{\|\mathbf{S}-\tilde{\mathbf{S}}\|^2\}$, plus the distortion of $\hat{\mathbf{S}}$ with respect to $\tilde{\mathbf{S}}$, namely $\mathbb{E}\{\|\hat{\mathbf{S}}-\tilde{\mathbf{S}}\|^2\}$. Consequently, for a given $\bar{\mathbf{Q}}$, the optimal digital filter and the resulting MSE derived for an asymptotically large number of subcarriers M is stated in the following lemma:

Lemma 2: The MSE minimizing digital filter $\mathbf{A}$ *under a fixed DMA configuration* $\bar{\mathbf{Q}}$ *and ADC support* γ *is given by*

$$\mathbf{A}^*(\bar{\mathbf{Q}}) = \mathbf{V}_2^H\bar{\mathbf{G}}^H\bar{\mathbf{Q}}^H\left(\sigma_q^2\mathbf{I}_{MN_d} + \bar{\mathbf{Q}}\bar{\boldsymbol{\Upsilon}}\bar{\mathbf{Q}}^H\right)^{-1}\mathbf{V}_1^H, \tag{12}$$

where $\bar{\boldsymbol{\Upsilon}} \triangleq \mathrm{blkdiag}\left(\boldsymbol{\Upsilon}_0, \boldsymbol{\Upsilon}_1, \ldots, \boldsymbol{\Upsilon}_M\right)$. *The corresponding MSE is given by* $\mathrm{EMSE}(\bar{\mathbf{Q}}) + e^{\mathrm{o}}$, *where the excess MSE, which*

corresponds to $\mathbb{E}\{\|\hat{\mathbf{S}}-\tilde{\mathbf{S}}\|^2\}$*, is given by*

$$\mathrm{EMSE}(\bar{\mathbf{Q}}) = \sum_{m\in\mathcal{M}} \mathrm{tr}\Big\{\hat{\mathbf{G}}_m^H \boldsymbol{\Upsilon}_m^{-1}\hat{\mathbf{G}}_m - \hat{\mathbf{G}}_m^H\mathbf{Q}_m^H\left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\boldsymbol{\Upsilon}_m\mathbf{Q}_m^H\right)^{-1}\mathbf{Q}_m\hat{\mathbf{G}}_m\Big\}. \tag{13}$$

Proof: See Appendix B. ∎

The requirement of M to be asymptotically large, which implies the entries of $\mathbf{S}$ approach being Gaussian in light of the central limit theorem, is necessary to guarantee that the linear minimal MSE (LMMSE) estimate of $\mathbf{S}$ from the channel output $\bar{\mathbf{Y}}$ is MSE optimal. For a finite M, the lemma also rigorously holds when the entries of $\mathbf{S}$ obey a jointly Gaussian distribution, i.e., the transmitters utilize Gaussian symbols, as commonly assumed in massive MIMO studies [1], [44], [45]. The minimization excess MSE in (13) is used in the sequel as the objective for setting the DMA weights.

Lemma 2 reveals the formation of the digital filter and characterizes the achievable MSE of the OFDM symbol recovery for a given DMA weight configuration $\bar{\mathbf{Q}}$. Once the DMA weight matrix $\bar{\mathbf{Q}}$ is determined, the digital filter and ADC dynamic range can be configured by (12) and (11), respectively. Specifically, the formation of digital filter $\mathbf{A}^*(\bar{\mathbf{Q}})$ in (12) implies that the processing in the digital domain is performed by first transforming the output of the low resolution ADCs, $\mathcal{Q}_C(\mathbf{Z})$, into the frequency domain, and subsequently implementing the linear filter $\hat{\mathbf{G}}_m^H\mathbf{Q}_m^H\left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\boldsymbol{\Upsilon}_m\mathbf{Q}_m^H\right)^{-1}$ at each frequency bin $m\in\mathcal{M}$. The key step in configuring the DMA based on Lemma 2 is to minimize $\mathrm{EMSE}(\bar{\mathbf{Q}})$ over the feasible set of $\bar{\mathbf{Q}}$. However, the constraints on the structure of $\bar{\mathbf{Q}}$ and on the feasible values of its nonzero entries, as well as the complicated nonconvex expression of the objective function (13), render the optimization problem a challenging task. This challenge motives further translation of the objective function and manipulation to the structure constraint on $\bar{\mathbf{Q}}$ using the matrix quadratic transformation [46].

B. Matrix Quadratic Transformation

Here, we translate the objective function (13) into to an equivalent solvable form using the matrix quadratic transformation. We start by examining the characteristics of the expression (13), which takes the form of the sum-of-functions-of-matrix-ratio problem. Specifically, the matrix $\hat{\mathbf{G}}_m^H\mathbf{Q}_m^H\left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\boldsymbol{\Upsilon}_m\mathbf{Q}_m^H\right)^{-1}\mathbf{Q}_m\hat{\mathbf{G}}_m$ can be treated as the ratio between the numerator matrix $\mathbf{Q}_m\hat{\mathbf{G}}_m\hat{\mathbf{G}}_m^H\mathbf{Q}_m^H$ and the denominator matrix $\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\boldsymbol{\Upsilon}_m\mathbf{Q}_m^H$, and $\mathrm{tr}(\cdot)$ is a non-decreasing matrix function in the sense that $\mathrm{tr}(\mathbf{R}_1)\geq\mathrm{tr}(\mathbf{R}_2)$ if $\mathbf{R}_1-\mathbf{R}_2$ is positive semidefinite. The *matrix quadratic transform* [46, Theorem 1] allows casting the matrix ratio form (13) into a matrix quadratic form, as stated in the following proposition:

Proposition 1: The minimization of (13) *is equivalent to*

$$\max_{\substack{\{\mathbf{Q}_m\}\in\mathcal{W}_{\mathrm{DMA}},\\ \{\boldsymbol{\Phi}_m\}\in\mathcal{C}^{N_d\times K}}} \sum_{m\in\mathcal{M}} \mathrm{tr}\Big\{2\Re\Big(\hat{\mathbf{G}}_m^H\mathbf{Q}_m^H\boldsymbol{\Phi}_m\Big) - \boldsymbol{\Phi}_m^H\left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\boldsymbol{\Upsilon}_m\mathbf{Q}_m^H\right)\boldsymbol{\Phi}_m\Big\}, \tag{14}$$

where $\{\boldsymbol{\Phi}_m\}\in\mathcal{C}^{N_d\times K}$ *are auxiliary matrices introduced for the matrix ratio term corresponding to frequency bin* m*.*

Proof: The first term in (13) is a constant with respect to $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$ and thus minimizing (13) is equivalent to

$$\max_{\{\mathbf{Q}_m\}\in\mathcal{W}_{\mathrm{DMA}}} \sum_{m\in\mathcal{M}} \mathrm{tr}\Big\{\hat{\mathbf{G}}_m^H\mathbf{Q}_m^H\left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\boldsymbol{\Upsilon}_m\mathbf{Q}_m^H\right)^{-1}\mathbf{Q}_m\hat{\mathbf{G}}_m\Big\}. \tag{15}$$

For a given $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$, the optimal $\boldsymbol{\Phi}_m$ maximizing (14) can be derived as

$$\boldsymbol{\Phi}_m = \left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\boldsymbol{\Upsilon}_m\mathbf{Q}_m^H\right)^{-1}\mathbf{Q}_m\hat{\mathbf{G}}_m. \tag{16}$$

Substituting (16) back to (14) leads to (15), which establishes the equivalence. ∎

The above proposition decouples the sum-of-functions-of-matrix-ratio problem in (13) into a more tractable problem as the first term in (14) is an affine function and the second term in (14) is a quadratic function of the matrix $\mathbf{Q}_m$. In this case, the equivalent problem (14) of tunnig the DMA weights for a fixed $\boldsymbol{\Phi}_m$ can be tackled using quadratic programming. Unfortunately, the structure constraints on $\mathbf{Q}_m$ induced by the feasible set of DMA weights, and particularly, the fact that some of its entries are restricted to zero, prevents direct application of quadratic programming solvers. Next, in an effort to eliminate the structure constraint of $\mathbf{Q}_m$, we transform (14) into a quadratic problem with respect to a vector concatenating the non-zero entries of $\mathbf{Q}_m^H$, as stated in the following proposition:

Proposition 2: The problem (14) *can be reformulated as an optimization with respect to the stacking vector of non-zero elements of* $\{\mathbf{Q}_m^H\}$*, i.e.,* $\bar{\mathbf{q}}_m=[\mathbf{q}_{m,1}^T,\mathbf{q}_{m,2}^T,\cdots,\mathbf{q}_{m,N_d}^T]^T$*, as follows*

$$\max_{\{\bar{\mathbf{q}}_m^*\}\in\mathcal{W}^{N\times 1}} \sum_{m\in\mathcal{M}} 2\Re\Big\{\sum_{k=1}^{K}\left(\boldsymbol{\phi}_{m,k}^T\otimes\hat{\mathbf{g}}_{m,k}^H\right)\bar{\mathbf{E}}\bar{\mathbf{q}}_m\Big\} - \bar{\mathbf{q}}_m^H\bar{\mathbf{E}}^H\left(\boldsymbol{\Phi}_m^*\boldsymbol{\Phi}_m^T\otimes\boldsymbol{\Upsilon}_m\right)\bar{\mathbf{E}}\bar{\mathbf{q}}_m, \tag{17}$$

where $\bar{\mathbf{E}}\triangleq\mathrm{blkdiag}(\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_{N_d})$*,* $\boldsymbol{\phi}_{m,k}$ *and* $\hat{\mathbf{g}}_{m,k}$ *denote the* k*th column of the matrices* $\boldsymbol{\Phi}_m$ *and* $\hat{\mathbf{G}}_m$*, respectively.*

Proof: See Appendix C. ∎

The equivalent representation in (17) optimizes only the non-zero entries of $\mathbf{Q}_m$. The remaining task is to solve the quadratic problem (17) considering the particular value constraints imposed on each of the entries of $\bar{\mathbf{q}}_m$, which are encapsulated in the feasible set $\mathcal{W}$. In the next subsection, we propose iterative DMA weight configuration procedures which accounts for different types of feasible set $\mathcal{W}$, based on matrix fractional programming (MFP) [46].

C. MFP-Based DMA Configuration

Applying MFP, we propose updating $\{\boldsymbol{\Phi}_m\}_{m\in\mathcal{M}}$, $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$ and γ^2 in an alternating fashion, setting each quantity in light of (14) while fixing the remaining ones. The setting of $\{\boldsymbol{\Phi}_m\}_{m\in\mathcal{M}}$ for fixed $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$ and γ^2, and the setting of γ^2 for fixed $\{\boldsymbol{\Phi}_m,\mathbf{Q}_m\}_{m\in\mathcal{M}}$

are given by (16) and (11), respectively. Thus, in the following, we elaborate on solving the problem (17) with respect to $\bar{\mathbf{q}}_m$ with fixed $\mathbf{\Phi}_m$ and γ^2, while considering different types of value constraints on $\bar{\mathbf{q}}_m$. For notational conciseness, we define $\boldsymbol{\xi}_m \triangleq \left(\sum_{k=1}^{K}\left(\boldsymbol{\phi}_{m,k}^T \otimes \hat{\mathbf{g}}_{m,k}^H\right)\bar{\mathbf{E}}\right)^H$ and $\mathbf{\Psi}_m \triangleq \bar{\mathbf{E}}^H\left(\mathbf{\Phi}_m^* \mathbf{\Phi}_m^T \otimes \mathbf{\Upsilon}_m\right)\bar{\mathbf{E}}$.

The strategy for recovering the weights depends on the feasible set $\mathcal{W}_{DMA}$. As detailed in Subsection II-A, two main models are considered for the DMA weights: a frequency-flat approximation which is accurate for narrowband signals, where the weights obey a relatively simple amplitude-only and Lorentzian-constrained phase form; and a frequency-selective model, accounting for the full frequency response of the metamateiral elements via the Lorentzian form (3). Therefore, in the following we divide our discussion into the configurations of frequency-flat elements (amplitude-only and Lorentzian-constrained weights) and frequency selective ones, presenting their corresponding alternating schemes.

1) Frequency-Flat Weights Design: In particular scenarios, such as low quality factor or narrowband transmissions, the frequency response of meatamterial elements are approximated as frequency flat [17]. This model is based on a narrowband signalling approximation for the frequency response (1), which also holds when the resonance frequency is far from the band of interest. In such cases, each element exhibits the same frequency response, which takes values in the set $\mathcal{W}$, for all considered frequency bins. Therefore, in this part, we remove frequency index 'm' from the weights, denoting $\bar{\mathbf{q}}_1 = \cdots = \bar{\mathbf{q}}_M \triangleq \bar{\mathbf{q}}$. In addition, we define $\boldsymbol{\xi} \triangleq \sum_{m\in\mathcal{M}} \boldsymbol{\xi}_m$ and $\mathbf{\Psi} \triangleq \sum_{m\in\mathcal{M}} \mathbf{\Psi}_m$. With the above notations, the general problem (17) can be simplified as

$$\max_{\{\bar{\mathbf{q}}^*\}\in\mathcal{W}^{N\times 1}} \quad 2\Re\left(\boldsymbol{\xi}^H\bar{\mathbf{q}}\right) - \bar{\mathbf{q}}^H\mathbf{\Psi}\bar{\mathbf{q}}. \tag{18}$$

For amplitude-only weights, all entries of $\bar{\mathbf{q}}$ have real values within the interval $[a_{\min}, a_{\max}]$, then problem (18) is equivalent to

$$\max_{\bar{\mathbf{q}}\in[a_{\min},a_{\max}]^{N\times 1}} \quad 2\Re\left(\boldsymbol{\xi}^H\right)\bar{\mathbf{q}} - \bar{\mathbf{q}}^T\mathbf{\Psi}\bar{\mathbf{q}}, \tag{19}$$

which yields a maximization of a concave function with linear constraints, i.e., it can be solved using conventional convex optimization methods.

For Lorentzian-constrained weights, $\bar{\mathbf{q}}$ takes the form $\bar{\mathbf{q}} = \frac{1}{2}(-j\mathbf{e}_N + \mathbf{p})$, where each entry of $\mathbf{p}$ is unit modulus, i.e., $\|(\mathbf{p})_i^*(\mathbf{p})_i\|^2 = 1$, and $\mathbf{e}_N$ denotes the N-dimensional all-one vector. Substituting this form into (18) yields the following constrained quadratic problem

$$\max_{\mathbf{p}} \quad \Re\left\{(2\boldsymbol{\xi} + j\mathbf{\Psi}\mathbf{e}_N)^H\mathbf{p}\right\} - \tfrac{1}{2}\mathbf{p}^H\mathbf{\Psi}\mathbf{p}, \tag{20a}$$

$$\text{s.t.} \quad \|(\mathbf{p})_i^*(\mathbf{p})_i\|^2 = 1 \text{ for } i = 1, 2, \cdots, N. \tag{20b}$$

Different from (19), this problem is non-convex due to the quadratic equality constraint (20b), which motivates us to derive a convex relaxation. A commonly used strategy is based on expressing (19) in terms of a positive semidefinite matrix. To this end, we define the following $(N+1)\times(N+1)$ symmetric matrix $\mathbf{T}$ and rank-one matrix $\mathbf{P}$ as follows:

$$\mathbf{T} = \frac{1}{2}\begin{bmatrix} -\mathbf{\Psi} & 2\boldsymbol{\xi} + j\mathbf{\Psi}\mathbf{e}_N \\ (2\boldsymbol{\xi} + j\mathbf{\Psi}\mathbf{e}_N)^H & 0 \end{bmatrix},$$
$$\mathbf{P} = \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix}\begin{bmatrix} \mathbf{p}^H & 1 \end{bmatrix}. \tag{21}$$

Then we reformulate the problem (20) with respect to the positive semidefinite matrix $\mathbf{P}$. The objective function (20a) can be rewritten as $[\mathbf{p}^H\ 1]\cdot\mathbf{T}\cdot[\mathbf{p}^H\ 1]^H = \mathrm{tr}(\mathbf{TP})$ by making use of the matrix formula $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BCA})$. Similarly, the constraint (20b) can be rewritten as $\mathrm{tr}(\mathbf{E}_{ii}\mathbf{P}) = 1$, where $\mathbf{E}_{ii}$ denotes a diagonal matrix with the i-th diagonal entry being one while others being zero. Unfortunately, the rank-one constraint on $\mathbf{P}$ is nonconvex. Nevertheless, semidefinite relaxation (SDR) can be applied to relax the reformulated problem by omitting the rank-one constraint, which results in the following SDR version:

$$\max_{\mathbf{P}} \ \mathrm{tr}(\mathbf{TP}),$$
$$\text{s.t. } \mathrm{tr}(\mathbf{E}_{ii}\mathbf{P}) = 1 \text{ for } i = 1, \cdots, N,\ \mathbf{P} \in \mathcal{S}^{(N+1)\times(N+1)}, \tag{22}$$

where $\mathcal{S}$ denotes the set of $(N+1)\times(N+1)$ positive semidefinite matrices. This problem is convex, which can be solved efficiently using well-developed techniques, such as interior point method. After solving (22), we need to extract a unit-modulus vector $\mathbf{p}$ from the solution matrix $\mathbf{P}$. One can obtain this vector via rank-one approximation to the matrix $\mathbf{P}$. Specifically, we first take the product of the maximum eigenvalue of $\mathbf{P}$ and its corresponding eigenvector, which yields $\mathbf{p}^{\text{appro}}$, then normalize the first N entries of $\mathbf{p}^{\text{appro}}$ such that the unit-modulus constraint (20b) holds, using the resultant vector to tune the Lorentzian-constrained phase weights.

In the above, we specialize the general form (17) into the solvable (19) and (22) for amplitude-only weights and Lorentzian-constrained weights, respectively. Finally we summarize the overall DMA configuration procedure in Algorithm 1. After random initialization of $\mathbf{Q}$, we successively update the auxiliary matrices $\{\mathbf{\Phi}_m\}_{m\in\mathcal{M}}$, DMA weights matrix $\mathbf{Q}$, and the ADC dynamic range γ, in an alternating manner over a predefined number of iterations $\mathrm{iter}_{\max}$.

2) Frequency-Selective Weights Design: While Algorithm 1 provides an effective approach to configure frequency flat DMA weights, it does not fully exploit the configurable frequency selective profile of the metamaterial elements. Following the frequency selective profile (3), and defining $\Omega_m \triangleq \Omega(\omega_m)$, we write $\bar{\mathbf{q}}_m = \mathbf{f}\odot\left(\mathbf{q}_m^{\sim F}\right)^*$, where $\mathbf{f} \triangleq [F_1, \cdots, F_N]^T$, and $\mathbf{q}_m^{\sim F} \triangleq [q_{m,1}^{\sim F}(\Omega_1^R, \chi_1), q_{m,2}^{\sim F}(\Omega_2^R, \cdots, q_{m,N}^{\sim F}(\Omega_N^R, \chi_N)]^T$, in which

$$q_{m,p}^{\sim F}(\Omega_p^R, \chi_p) \triangleq \frac{\Omega_m^2}{(\Omega_p^R)^2 - \Omega_m^2 - j\frac{\Omega_m\Omega_p^R}{\chi_p}}.$$

Note that the order of the entries of the vectors $\mathbf{f}$ and $\mathbf{q}_m^{\sim F}$ represents the stacking of the elements in the vector $\bar{\mathbf{q}}_m$, i.e., the p-th entry is associated with the microstrip index i

Algorithm 1 MFP-Based Frequency-Flat Weight setting

Input: Covariance matrices $\boldsymbol{\Upsilon}_m$ and equivalent channel matrices $\hat{\mathbf{G}}_m$ for $m \in \mathcal{M}$.

Init: DMA weight $\mathbf{Q}$ in the feasible and compute the corresponding γ^2 using (11).

for $\mathrm{iter} = 1, 2, \ldots, \mathrm{iter}_{\max}$ **do**
- Compute the auxiliary matrix $\boldsymbol{\Phi}_m$ using (16) for each $m \in \mathcal{M}$;
- Solve the problem (19) for amplitude-only weights, or (22) for Lorentzian-constrained weights, to obtain $\mathbf{Q} = \mathrm{blkdiag}(\mathbf{q}_1^H, \mathbf{q}_2^H, \cdots, \mathbf{q}_{N_d}^H)$;
- Set dynamic range γ^2 using (11) while setting $\sigma_q^2 = \frac{4\gamma^2}{3b^2}$;

end

Output: DMA weights $\mathbf{Q}$ and ADC dynamic range γ^2.

and the element index l in (3) via $p = N_d(i-1)+l$. Substituting the above formation of $\bar{\mathbf{q}}_m$ into (17) yields a complex optimization problem due to the non-convex structure of the objective function, being the sum of a quadratic function and the cross-products of several fractional terms. To tackle this challenging problem, we propose to update $\mathbf{f}$ as well as Ω_p^R and χ_p corresponding to each element of $\mathbf{q}_m^{\sim F}$ separately in an alternating manner.

When $\mathbf{q}_m^{\sim F}$ is fixed, the optimization over $\mathbf{f}$ is a quadratic problem with linear constraints:

$$\max_{\mathbf{f}} \; 2\Re\left\{\sum_{m\in\mathcal{M}} \boldsymbol{\xi}_m \mathsf{diag}\left(\left(\mathbf{q}_m^{\sim F}\right)^*\right)\right\}\mathbf{f} - \mathbf{f}^T\left\{\sum_{m\in\mathcal{M}} \mathsf{diag}\left(\mathbf{q}_m^{\sim F}\right)\boldsymbol{\Psi}_m \mathsf{diag}\left(\left(\mathbf{q}_m^{\sim F}\right)^*\right)\right\}\mathbf{f}, \quad (23)$$

which is convex and optimally solvable. However, when $\mathbf{f}$ is fixed, the optimization for each Ω_p^R and χ_p is analytically intractable. We thus propose a heuristic approach for the configuration of each Ω_p^R and χ_p. As numerically demonstrated in Section V, it leads to performance gains compared with frequency flat weights due to its ability to exploit the dynamic frequency selectivity of the DMA elements. Here, we sequentially design Ω_p^R and χ_p corresponding to each element of $\mathbf{q}_m^{\sim F}$ by searching over discrete grids, denoted by $\mathcal{F}$ and $\mathcal{X}$, respectively, which represent a broad range of typical frequency selective profiles. Specifically, for updating Ω_p^R and χ_p associated with the p-th element of $\mathbf{q}_m^{\sim F}$, we conditionally maximize the objective function in (17) while keeping the remaining elements of $\mathbf{q}_m^{\sim F}$ fixed. Accordingly, we define

$$\mathbf{q}_{m,\sim p}^{\sim F}(\Omega^R, \chi) \triangleq [\underbrace{q_{m,1}^{\sim F}(\Omega_1^R, \chi_1), \cdots, q_{m,p-1}^{\sim F}(\Omega_{p-1}^R, \chi_{p-1})}_{\text{fixed}}, \underbrace{q_{m,p}^{\sim F}(\Omega^R, \chi)}_{\text{to be updated}}, \underbrace{q_{m,p+1}^{\sim F}(\Omega_{p+1}^R, \chi_{p+1}), \cdots, q_{m,N}^{\sim F}(\Omega_N^R, \chi_N)}_{\text{fixed}}]^T.$$

Algorithm 2 Frequency-Selective Weights setting

Input: Covariance matrices $\boldsymbol{\Upsilon}_m$ and equivalent channel matrices $\hat{\mathbf{G}}_m$ for $m \in \mathcal{M}$.

Init: DMA weight $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$ in the feasible and compute the corresponding γ^2 using (11).

for $\mathrm{iter} = 1, 2, \ldots, \mathrm{iter}_{\max}$ **do**
- Compute the auxiliary matrix $\boldsymbol{\Phi}_m$ using (16) for each $m \in \mathcal{M}$;
- **for** $p = 1, 2, \ldots, N$ **do**
  - Set resonance frequency Ω_p^R and damping factor χ_p by solving the problem (24);
- **end**
- Set oscillator strength of each element $\mathbf{f}$ by solving the problem (23);
- Obtain $\bar{\mathbf{q}}_m = \mathbf{f} \odot \left(\mathbf{q}_m^{\sim F}\right)^*$ and rearrange $\{\bar{\mathbf{q}}_m\}_{m\in\mathcal{M}}$ into $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$;
- Set dynamic range γ^2 using (11) and parameter $\sigma_q^2 = \frac{4\gamma^2}{3b^2}$;

end

Output: DMA weight $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$ and ADC dynamic range γ^2.

Denoting $\boldsymbol{\Psi}'_m \triangleq \mathsf{diag}(\mathbf{f})\boldsymbol{\Psi}_m\mathsf{diag}(\mathbf{f})$, we express the sub-problem for configuring Ω_p^R and χ_p for fixed $\mathbf{f}$ as

$$\max_{\Omega^R\in\mathcal{F},\ \chi\in\mathcal{X}} \; 2\Re\left\{\sum_{m\in\mathcal{M}} \boldsymbol{\xi}_m \mathsf{diag}(\mathbf{f})\left(\mathbf{q}_{m,\sim p}^{\sim F}(\Omega^R,\chi)\right)^*\right\} - \left\{\sum_{m\in\mathcal{M}} \left(\mathbf{q}_{m,\sim p}^{\sim F}(\Omega^R,\chi)\right)^T \boldsymbol{\Psi}'_m \left(\mathbf{q}_{m,\sim p}^{\sim F}(\Omega^R,\chi)\right)^*\right\}. \quad (24)$$

Finally, we summarize the alternating approach in Algorithm 2. After random initialization of $\{\mathbf{Q}_m\}_{m\in\mathcal{M}}$, we successively update the auxiliary matrices $\{\boldsymbol{\Phi}_m\}_{m\in\mathcal{M}}$, the resonance frequency and the quality factor of each DMA element, the oscillator strength of each element and the ADC dynamic range γ in an alternating manner until a predefined maximum iteration number $\mathrm{iter}_{\max}$ is satisfied. Note that the gridded nature of the search set $\{\mathcal{F}, \mathcal{X}\}$ implies that the proposed algorithm is not guaranteed to recover the MSE-minimizing configuration. Nonetheless, it allows exploiting the configurable frequency selective profile of DMAs in light of the signal recovery objective while accounting for the bit constraints imposed on the receiver.

IV. Discussion

In this section, we discuss the proposed DMA-based receiver. In particular, we elaborate on its difference from conventional hybrid architectures in Subsection IV-A, present complexity analysis in Subsection IV-B, and discuss its associated possible extensions in Subsection IV-C.

A. Comparison to Conventional Hybrid Architectures

DMA-based receivers, as follows from our model in Subsection II-A, implement a type of hybrid beamforming. In particular, the received signal undergoes some analog

TABLE I
COMPUTATIONAL COMPLEXITY IN TERMS OF REAL-VALUED MULTIPLICATIONS FOR DIFFERENT OPERATIONS

	Frequency-Flat Weights	Frequency-Selective Weights
γ^2	$8N_dN_e^2$	$8MN_dN_e^2$
Grid search	-	$6N\left(N+1\right)+4N\left(\left\|\mathcal{X}\right\|\left\|\mathcal{F}\right\|-1\right)\left(2N+3\right)$
$\{\mathbf{\Phi}_m\}_{m\in\mathcal{M}}$	$M\left[4N_d\left(N_d+N_e\right)K+8N_d^2N_e^2+\frac{2}{3}N_d\left(N_d^2-1\right)\right]$	
$\{\boldsymbol{\xi}_m\}_{m\in\mathcal{M}}$	$4NKM$	
$\{\mathbf{\Psi}_m\}_{m\in\mathcal{M}}$	$2MN_d\left(N_e^2+K\right)$	

processing which reduces its dimensionality prior to being converted into a digital representation. Such hybrid architectures are commonly used in the massive MIMO literature, often as method to reduce the number of RF chains [10]–[12] but also to facilitate recovery under bit constraints [13]. Despite their similarities, there are several fundamental differences between DMAs and traditional hybrid receivers, both in terms of savings in cost and hardware as well as in terms of additional signal processing capabilities. In particular, conventional hybrid receivers utilize an additional dedicated hardware to implement the analog combining, typically consisting of phase shifter networks, i.e., an interconnection of phase shifters and adders. One of the benefits of using DMAs over conventional hybrid architectures, also noted in [28], is that the controllable power-efficient analog combining in DMAs is a natural byproduct of the antenna structure, and does not require additional dedicated hardware. The enhanced signal processing capabilities stem from the fact that conventional hybrid systems, such as phase shifter networks, are typically frequency flat, namely, the same analog mapping is implemented for all the spectral components of the input signal [12]. The inherent adjustable frequency selectivity of the DMA elements implies that they can be tuned to apply a frequency varying analog combining, which corresponds to one of the three profiles discussed in the previous subsection. This property is expected to improve the performance in wideband frequency selective setups, as numerically demonstrated in Section V.

B. Computational Complexity

Next, we provide a quantitative assessment of the computational complexity associated with tuning the proposed DMA-based receiver. The dynamic configurability of DMAs allows the receiver to adapt its analog mapping when the channel conditions change, e.g., as in block fading environments. The following analysis characterizes the complexity of the proposed algorithms configuring the DMA to a given channel.

In each iteration, MFP involves the computation of the auxiliary matrices $\{\mathbf{\Phi}_m\}_{m\in\mathcal{M}}$, vectors $\{\boldsymbol{\xi}_m\}_{m\in\mathcal{M}}$, matrices $\{\mathbf{\Psi}_m\}_{m\in\mathcal{M}}$ and dynamic range γ^2. We present the computational complexity of the above matrices in terms of the number of multiplication operations in Table I. We only elaborate the complexity analysis of the computation of auxiliary matrices $\{\mathbf{\Phi}_m\}_{m\in\mathcal{M}}$ here, and the complexity in computing the remaining quantities can be obtained in a similar way. Specifically, the computation of $\left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\mathbf{\Upsilon}_m\mathbf{Q}_m^H\right)^{-1}$ requires $8N_d^2N_e^2+\frac{2}{3}N_d\left(N_d^2-1\right)$ real-valued multiplications, due to the complexity of forming $\mathbf{Q}_m\mathbf{\Upsilon}_m\mathbf{Q}_m^H$ and computing the Cholesky factorization. In addition, the computation of the matrix multiplication $\left(\sigma_q^2\mathbf{I}_{N_d}+\mathbf{Q}_m\mathbf{\Upsilon}_m\mathbf{Q}_m^H\right)^{-1}\mathbf{Q}_m\hat{\mathbf{G}}_m$ requires $4N_d^2K+4N_dN_eK$ such operations. The setting of the dynamic range γ^2 also requires N_d-1 comparisons, and the grid search in the configuration of frequency-selective weight requires $6N\left(N+1\right)+4N\left(\left|\mathcal{X}\right|\left|\mathcal{F}\right|-1\right)\left(2N+3\right)$ multiplications and $N(\left|\mathcal{X}\right|\left|\mathcal{F}\right|-1)$ comparisons.

Finally, we note that the DMA configuration involves solving a relaxed convex optimization problem using the above multivariate quantities. In particular, tackling (19) for configuring frequency flat amplitude-only weights and (23) for frequency selective weights involve solving quadratic programming problems, whose complexity is in the order of $\mathcal{O}(N^2k)$ using conjugate gradient method, where k denotes the iteration number in implementing the conjugate gradient method. While this term dominates the complexity for the frequency flat case, the grid search carried out Algorithm 2 requires complexity of the order of $\mathcal{O}(N^2|\mathcal{F}||\mathcal{X}|)$, which is typically the dominant computational burden term. The problem (22), treated when optimizing frequency flat Lorentzian-constrained weights, requires solving a semidefinite programming problem, whose complexity is in the order of $\mathcal{O}(N^{4.5}\log(1/\varepsilon))$ [47], where ε denotes a given solution accuracy. We thus conclude that the computational complexity associated with the proposed algorithms for tuning DMAs to facilitate the operation of bit-constrained receivers is dominated by terms which grow with N^2 and $N^{4.5}$ when configuring frequency flat weights via Algorithm 1 under amplitude-only and Lorentzian-constrained restrictions, respectively, while the complexity of tuning frequency selective weights via Algorithm 2 grows with $N^2\left|\mathcal{X}\right|\left|\mathcal{F}\right|$. This implies that the complexity of Algorithm 2 can be balanced by limiting the cardinality of search grids, at the cost of possibly degrading the signal recovery accuracy of the resulting DMA configuration.

C. Future Extensions

The proposed methods are based on task-based quantization framework, exploiting the block diagonal structure of $\bar{\mathbf{Q}}$ to formulate a set of optimization problems which can be solved

in a alternating fashion. This structure of $\bar{\mathbf{Q}}$ stems from the DMA structure, which consists of several microstrips each feeding a different ADC. As the DMA operation is modeled as a form of analog beamforming, the proposed approach can also be utilized for conventional hybrid receivers in which the analog combiner exhibits a block diagonal structure, i.e., partially connected phase shifter networks [10], [34]. In these conventional architectures, the filter $\mathbf{H}(\omega)$, modeling the propagation inside the microstrips, is replaced with the identity matrix. Consequently, Algorithm 1, which is dedicated to frequency flat elements, can also be applied to design conventional partially connected frequency-invariant hybrid receivers by following a task-based quantization framework. Furthermore, the resulting model for DMAs used for transmission is similar to that used for reception here [29], [31]. We thus expect our proposed methods to facilitate the design of downlink DMA-assisted systems, and leave the detailed study of this application for future work.

The spectral profile of the DMA elements is exploited in Algorithm 2 to achieve various types of frequency selective analog combining. In particular, the Lorentzian form (3) allows to set each element to approach a controllable bandpass filter or an approximately spectrally-linear gain. While these profiles accommodate a broad family of spectral shapes, more complex profiles, such as multi-modal frequency responses, may not be closely approximated by DMAs. Furthermore, in Algorithm 2 we allowed to set the resonance frequencies $\{\Omega^R_{i,l}\}$ and the damping factors $\{\chi_{i,l}\}$ to take values in some predefined discrete sets $\mathcal{F}$ and $\mathcal{X}$. Setting the elements to have a large quality factor $\frac{\Omega^R_{i,l}}{\chi_{i,l}}$ may be difficult in practice, and this ratio is commonly restricted to be in the order of several tens [37]. These additional considerations can also be accounted for in the optimization of the frequency response parameters of the elements by setting the sets over which the search is conducted to account for these constraints. We leave the detailed analytical study of the effect of these considerations for future research.

The algorithms detailed in Section III provide methods for configuring the DMA weights along with the dynamic range of the ADCs and the digital processing to accurately recover OFDM signals, facilitating the demodulation by the BS. This procedure requires the receiver to know the multipath channel $\mathbf{G}[\tau]$, the noise covariance $\mathbf{C}_W$, and the DMA frequency selectivity profile $\mathbf{H}(\omega)$. While the latter can be obtained from the physics of the metasurface, the channel parameters must be estimated, which may be a challenging task in the presence of quantized channel outputs [13], [38], [48]. When channel estimation is carried out in a time division duplexing manner, the dynamic nature of DMAs can be exploited to assign different configurations during channel estimation and signal recovery. However, we leave this study for future investigation.

V. Numerical Evaluations

In this section, we numerically evaluate the performance of bit-constrained MIMO-OFDM receivers in which the BS is equipped with a DMA configured using the scheme proposed in Section III. We first detail the simulation setup in Subsection V-A, after which we present the numerical results in Subsection V-B.

A. Parameter Setup

We consider a single-cell uplink MU-MIMO setup in which a BS serves $K = 8$ users. The users simultaneously transmit OFDM signals with central frequency $f_c = 1.9$ GHz and bandwidth 40 MHz, which comprises $M = 128$ subcarriers. The data symbols are independently drawn from a QPSK constellation unless otherwise specified. The BS is equipped with a DMA consisting of a total $N = 100$ elements, divided into $N_d = 10$ microstrips with $N_e = 10$ elements in each. The frequency domain characterization of the propagation inside the microstrip, denoted as $h_{i,l}(\omega)$, is set as $h_{i,l}(\omega) = e^{-\alpha l - j\beta(\omega) l}$ for $i \in \mathcal{N}_e$, where $\alpha = 0.006$ $[\mathrm{m}^{-1}]$ and $\beta(\omega) = 1.592 \cdot \omega$ $[\mathrm{m}^{-1}]$. This setting represents microstrips with 50 ohm characteristic impedance made of Duroid 5880 operating at 1.9 GHz with element spacing of 0.2 wavelength (assuming free space wavelength) without mutual correlation between different microstrips [49, Ch. 3.8].

The wireless channel is generated based on the delay-d model [50]. Under this model, the channel between the user k and the BS is expressed by $\mathbf{g}_{k,d} = \sum_{s=1}^{L_k} \beta_{k,s} p(dT_S - \tau_{k,s}) \mathbf{a}(\theta_{k,s}, \varphi_{k,s})$, where L_k denotes the number of channel paths, which is set to 10 for all $k \in \mathcal{K}$; $\beta_{k,s}$ denotes the path gain of the path s for the user k, which are i.i.d. zero-mean complex Gaussian variables of unit variance; $\tau_{k,s}$ $\theta_{k,s}$ and $\varphi_{k,s}$ denote the path delay, the azimuth and elevation angles of arrival of the path s for the user k, respectively; $\mathbf{a}(\theta_{k,s},\varphi_{k,s})$ denotes the received steering vector, which is generated as the steering vector of a $N_d \times N_e$ uniform planar array; T_s denotes the symbol period; $p(\tau)$ denote the combination of transmit and receive filters, which is simulated as the raised-cosine filter with roll-off factor one. The quantities $\theta_{k,s}$ and $\varphi_{k,s}$ are generated as random variables uniformly distributed over $[-\pi, \pi]$ and $[-\frac{\pi}{2}, \frac{\pi}{2}]$, respectively, while $\tau_{k,s}$ is normalized to T_s and is generated from a discrete uniform distribution over the interval $[0, D-1]$, where D denotes the length of cyclic prefix. We transform the channel in the time domain $\mathbf{g}_{k,d}$ into the frequency domain via $\mathbf{g}_{k,m} = \sum_{d=0}^{D-1} \mathbf{g}_{k,d} e^{-\frac{j2\pi md}{M}}$, and the channel matrix in (6) is obtained by $\mathbf{G}_m = [\mathbf{g}_{1,m}, \mathbf{g}_{2,m}, \cdots, \mathbf{g}_{K,m}]$. In addition, the noise covariance $\mathbf{C}_W$ is set to be $\mathbf{C}_W = \sigma_z^2 \mathbf{I}_N$, where $\sigma_z^2 > 0$. While in principle the noise is likely to exhibit spatial correlation due to the sub-wavelength spacing of the elements, we consider conventional white Gaussian noise as commonly assumed in wireless receivers studies, while the spatial correlation is encapsulated in the channel coefficients. The instantaneous signal-to-noise ratio (SNR) is defined as $\frac{\sum_{m \in \mathcal{M}} \|\mathbf{G}_m \mathbf{s}_m\|^2}{MN\sigma_z^2}$.

In our simulations, we consider a bit-constrained BSs with a budget of overall B_{overall} bits, divided equally among the ADCs such that each ADC operates with a resolution of $b = \lfloor 2^{B_{\mathrm{overall}}/(2 \cdot N_d)} \rfloor$ output levels. For the task-based quantizers, we set $\eta = 2$. Unless stated otherwise, we carry out $\mathrm{iter}_{\max} = 20$ iterations when implementing Algorithms 1-2, while our numerical observation reported in the sequel shows

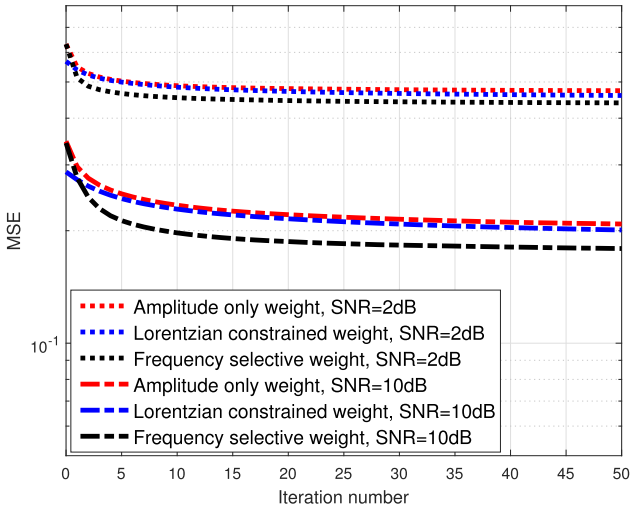
that convergence is typically observed after 15 iterations. The output of the ADCs is processed using the digital filter given by (12). We evaluate the uncoded bit error rate (BER) and MSE performance via Monte Carlo simulations over 100 independent channel realizations, where in each channel realization 100 MIMO-OFDM symbols are transmitted.

B. Simulation Results

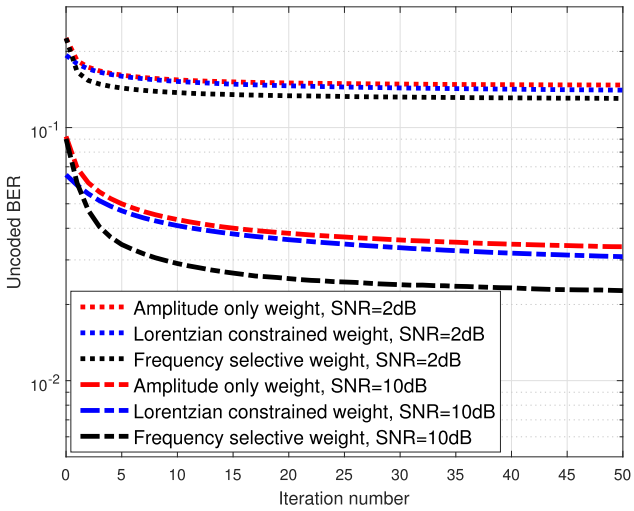
We begin by presenting the convergence of our proposed MFP-based algorithms for tunning the DMA weights (Algorithm 1 for frequency flat weight and Algorithm 2 for frequency selective weight). The motivation for comparing the narrowband setting of Algorithm 1 to the frequency selective configuration of Algorithm 2 is to empirically demonstrate the gains in accounting for the frequency selective nature of metamaterial elements over restricting them to operate in a frequency flat manner. To that aim, we numerically evaluated the uncoded BER and MSE versus the iteration number under $B_{\text{overall}} = 80$ bits in Fig. 6. As observed in Fig. 6, in all cases, the values of BER and MSE reduce sharply in the first few iterations and then remain stable with only minor fluctuations. After about 10 iterations, only minor decrease in BER and MSE is observed, and numerical stability can be regarded to be reached. From Fig. 6, we also notice that exploiting the frequency selective nature of the DMA elements allows achieving improved performance, especially under high SNR. These results indicate that while Algorithm 2 is based on heuristic arguments, it effectively extracts the potential performance gain of proper configuration of the dynamic frequency selective profile of DMAs.

We subsequently evaluate the signal recovery performance of the proposed receiver architecture versus the SNR in the range of $[-4, 16]$ dB under $B_{\text{overall}} = 80$ bits and depict the results in Fig. 7. In such, conventional conventional digital-only processing cannot be applied, and accurate symbol recovery is a challenging task, hence the relatively high BER values observed in Fig. 7a. As the measure evaluated is the uncoded BER, the error rate can be further reduced by using channel coding, as we show in the sequel. We compare the performance of our proposed DMA-based receivers to that achieved using frequency-flat DMA receivers designed using the greedy method proposed in our preliminary work [51], and to that of a partially-connected phase shifter network based hybrid receiver designed using the analog beamformer selection (ABFS) method proposed in [34].

Observing Fig. 7, we first note that accounting for the frequency selectivity of the weights allows achieving more accurate recovery compared to restricting the weights to be frequency flat. This indicates the improved reliability of configuring DMA weight with frequency selective profile and the effectiveness of the proposed strategy. Comparing the performance of two types of frequency flat weights, slight performance gain can be observed when using Lorentzian-constrained weights compared with amplitude-only weights. However, solving the SDR problem when configuring Lorentzian-constrained weights requires more computationally complexity than solving linearly constrained quadratic problem when configuring amplitude-only weights. Furthermore,



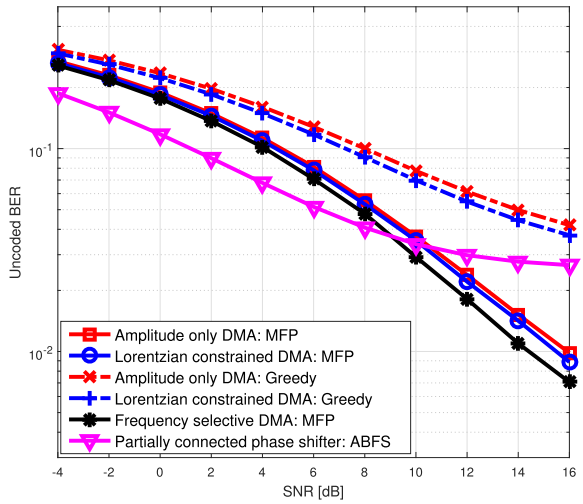
(a) MSE vs. iterations



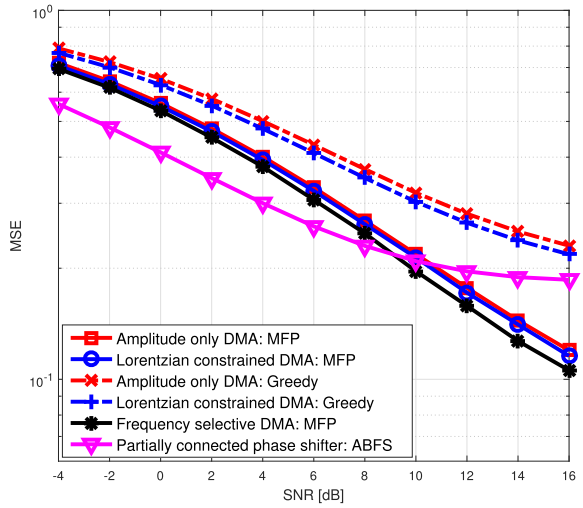
(b) Uncoded BER vs. iterations

Fig. 6. Signal recovery performance versus iteration number under $B_{\text{overall}} = 80$ bits and SNR = 2 dB and 10 dB.

the proposed MFP-based strategies apparently outperform greedy based strategies, showing the benefits of our proposed DMA configuration strategy based on task-based quantization and MFP framework. We then compare the performance of the proposed DMA-based hybrid receiver with the conventional partially-connected phase shifter network based hybrid receiver [34]. The phase shifter based receiver slightly outperforms the DMA based receiver in lower SNR region, which can be attributed to the frequency selective propagation in each microstrip, which induces a negative impact on the signal recovery capabilities. In higher SNR regions, the phase-shifter based architecture meets an error floor, and is outperformed by our proposed DMA receivers, whose performance continues decreasing as the SNR grows. These performance gains add to the fact that conventional hybrid receivers require additional dedicated analog combining circuitry, while DMA-based receivers implement their analog mapping as a natural byproduct of their antenna architecture, thus saving hardware cost

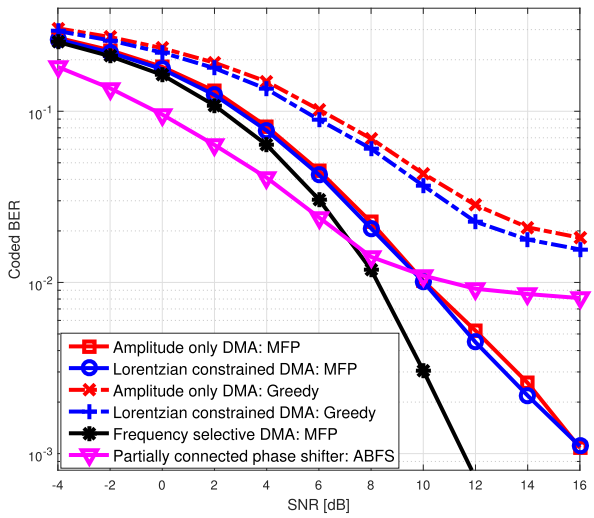


(a) Uncoded BER

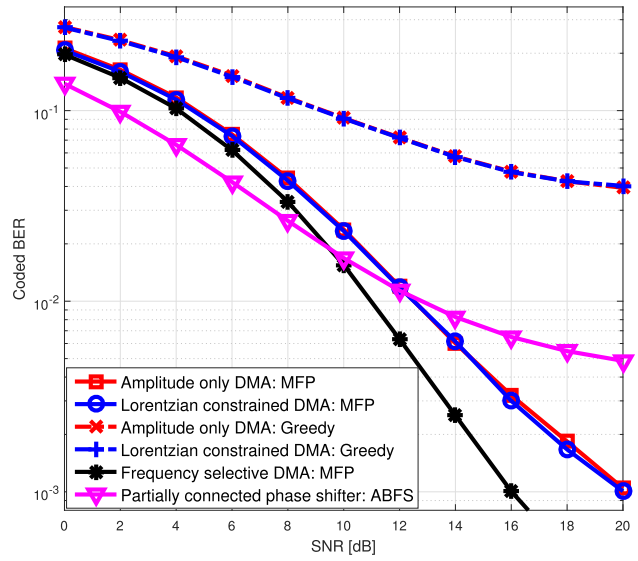


(b) MSE

Fig. 7. Uncoded BER and MSE versus SNR under $B_{\text{overall}} = 80$ bits.



(a) QPSK, $B_{\text{overall}} = 80$ bits.



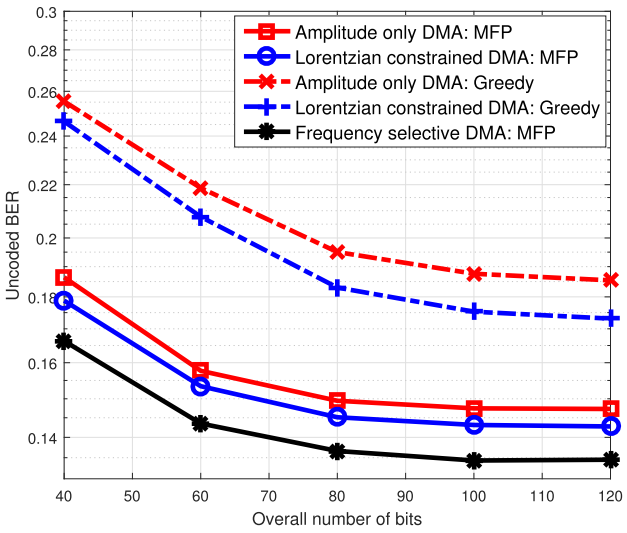
(b) 16QAM, $B_{\text{overall}} = 200$ bits.

Fig. 8. Coded BER and MSE versus SNR.

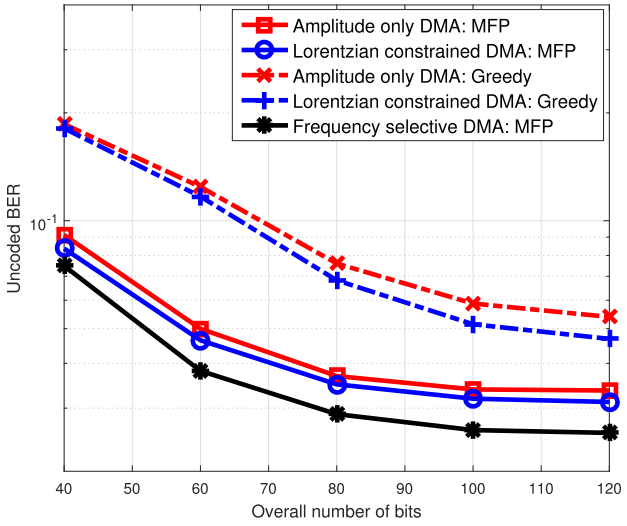
and size, as well as facilitating hardware miniaturization and integration of large-scale antenna arrays. The performance gains observed in high SNRs stem form their resulting model, which is not restricted to be frequency flat and to tune only the phases of the elements, as well as the task-based quantization tools used in our methods. While we believe that a similar analysis can improve the achievable performance of conventional hybrid receivers with bit constraints, we leave its analysis for future investigation.

We next demonstrate that the uncoded BER values observed in Fig. 7a, which are relatively high due to the tight bit budget, are translated into improved coded BER in the presence of channel coding. To that aim, we repeat the simulation study whose uncoded BER results are depicted in Fig. 7a, while adding a rate $1/3$ turbo code. Furthermore, to illustrate that the gains observed in Fig. 7 are maintained for modulations of higher order than QPSK, we also compute the coded BER when using 16QAM modulation. As higher order modulations typically require less noisy observations to be accurately detected compared to lower order ones, we increase the number of bits to $B_{\text{overall}} = 200$ and the number of RF chains to $N_d = 25$ while setting $N_e = 4$ to preserve the overall number of $N = 100$ elements. The coded BER results are depicted in Fig. 8, where it is observed that the MSE and uncoded BER gains achieved by the proposed DMA-aided receiver design noted in Fig. 7 are translated into similar gains of improved error rates when using channel coding, combined with both QPSK as well as 16QAM modulations.

Finally, we evaluate how the signal recovery performance of the DMAs-based receiver scales with respect to B_{overall}, providing guidance to budget in quantization bits in practical systems. To that aim, we depict in Figs. 9a and 9b the uncoded BER of the proposed receiver versus $B_{\text{overall}} \in [40, 120]$ under signal-to-noise ratios (SNRs) of 2 dB and 10 dB, respectively. Observing Fig. 9, we note that the MSE and uncoded BER gains of the proposed designs, observed in Fig. 7 for receivers operating utilizing 80 bits for representing their channel output, hold for all considered values



(a) SNR = 2 dB



(b) SNR = 10 dB

Fig. 9. Uncoded BER versus B_{overall} under SNR = 2 dB and 10 dB.

of B_{overall}. For $B_{\text{overall}} \leq 80$ bits, the MSE and uncoded BER both reduce sharply as the overall bit budget B_{overall} increases. However for $B_{\text{overall}} > 80$ bits, increasing B_{overall} only results in negligible performance gain, indicating that the proposed design effectively mitigates the distortion due to coarse quantization and the error term due to the noise induced by the channel in that B_{overall} region. Therefore, for the considered setup $B_{\text{overall}} = 80$ bits achieves a good balance between the hardware cost and power consumption induced by overall quantization bits and the signal recovery performance.

To summarize, DMAs typically use less power and cost less than standard antenna arrays to implement a configurable frequency selective hybrid architecture. The results presented in this section demonstrate that by using our proposed methods, one can implement a reliable DMA-based receiver which is particularly suitable for bit-constrained MU-MIMO-OFDM setups, achieving notable performance gains using previously proposed methods.

VI. Conclusion

In this work we studied the application of DMAs, which realize low cost and power efficient configurable antenna arrays, for bit-constrained MU-MIMO-OFDM systems. We formulated a model for the coarsely quantized DMA outputs which accounts for the adaptable frequency selective profile of the metamaterial elements, and characterized the achievable OFDM recovery accuracy for a given DMA configuration based on the task-based quatization framework. Next, we translated the joint design of the DMA weights and the digital filter into quadratic form based on the matrix fractional programming and eliminating the constraints imposed by the structure of the DMA weight matrix for tractability. Finally, we proposed alternating algorithms for DMAs configuration considering different types of DMA weights. Our numerical results demonstrate that by properly exploiting the frequency selective nature of DMAs, OFDM symbol detection is significantly facilitated when operating under strict bit constraints. We also demonstrated how the proposed method, which utilizes task-based quantization to determine the pre-quantization combining carried out in the DMA along with the dynamic range of the ADCs and the digital processing, can be utilized to design frequency flat hybrid receivers, such as conventional phase shifter networks, achieving improved performance compared to previously proposed hybrid architectures.

Appendix

A. Proof of Lemma 1

From the mathematical model of the DMA output (8), we have that

$$\begin{aligned}
&\mathbb{E}\left\{\mathbf{z}[t]\mathbf{z}^H[t]\right\} \\
&= \left((\mathbf{F}_M)_{:,t}^H \otimes \mathbf{I}_{N_d}\right) \bar{\mathbf{Q}} \cdot \mathbb{E}\left\{\bar{\mathbf{G}}\bar{\mathbf{G}}^H + \bar{\mathbf{H}}\left(\mathbf{I}_M \otimes \mathbf{C}_W\right)\bar{\mathbf{H}}^H\right\} \\
&\qquad \cdot \bar{\mathbf{Q}}^H \left((\mathbf{F}_M)_{:,t} \otimes \mathbf{I}_{N_d}\right) \\
&= \sum_{m\in\mathcal{M}} |((\mathbf{F}_M)_{:,t})_m|^2 \mathbf{Q}_m \left(\hat{\mathbf{G}}_m\hat{\mathbf{G}}_m^H + \mathbf{H}_m\mathbf{C}_W\mathbf{H}_m^H\right)\mathbf{Q}_m^H \\
&= \frac{1}{M}\sum_{m\in\mathcal{M}} \mathbf{Q}_m \boldsymbol{\Upsilon}_m \mathbf{Q}_m^H.
\end{aligned}$$

The above expression indicates that the covariance of $\mathbf{z}[t]$ is invariant to the time index t, and thus the maximization over t in (11) can be removed. Hence, $\mathbb{E}\{|(\mathbf{z}[t])_i|^2\} = \left(\mathbb{E}\left\{\mathbf{z}[t]\mathbf{z}^H[t]\right\}\right)_{i,i} = \frac{1}{M}\sum_{m\in\mathcal{M}} (\mathbf{Q}_m)_{i,:}\, \boldsymbol{\Upsilon}_m \left(\mathbf{Q}_m^H\right)_{:,i} = \frac{1}{M}\sum_{m\in\mathcal{M}} \mathbf{q}_{m,i}^H \mathbf{E}_i^H \boldsymbol{\Upsilon}_m \mathbf{E}_i \mathbf{q}_{m,i}$, which completes the proof.

B. Proof of Lemma 2

To prove the lemma, we recall that with the assumption of dithered quantization, the ADC output $\mathbf{R} = \mathcal{Q}_C(\mathbf{Z})$ can be written as the sum of its input $\mathbf{Z} = \mathbf{V}_1\bar{\mathbf{Q}}\bar{\mathbf{Y}}$ and an additive uncorrelated noise $\mathbf{n}_q$ with zero mean and covariance matrix $\sigma_q^2\mathbf{I}_{N_dM}$ and that $\mathbb{E}\{\|\mathbf{S}-\hat{\mathbf{S}}\|^2\} = e^{\text{o}} + \mathbb{E}\{\|\hat{\mathbf{S}}-\tilde{\mathbf{S}}\|^2\}$. Thus, the optimal digital filter is derived to minimize $\mathbb{E}\{\|\hat{\mathbf{S}}-\tilde{\mathbf{S}}\|^2\}$, which results in the linear MMSE estimator of $\tilde{\mathbf{S}}$ from $\mathbf{R} = \mathbf{Z} + \mathbf{n}_q$. Note that the asymptotic Gaussianity of OFDM signals [52] implies that the MMSE estimate of $\mathbf{S}$ from $\bar{\mathbf{Y}}$

is linear and is given by $\tilde{\mathbf{S}} = \mathbb{E}\{\mathbf{S}\bar{\mathbf{Y}}^H\}\mathbb{E}\{\bar{\mathbf{Y}}\bar{\mathbf{Y}}^H\}^{-1}\bar{\mathbf{Y}}$, when the number of subcarriers M is asymptotically large. Then the optimal digital filter and its corresponding MSE are given by $\mathbf{A}^*(\bar{\mathbf{Q}}) = \mathbb{E}\{\tilde{\mathbf{S}}\mathbf{R}^H\}\mathbb{E}\{\mathbf{R}\mathbf{R}^H\}^{-1} = \mathbf{V}_2^H\bar{\mathbf{G}}^H\bar{\mathbf{Q}}^H\big(\sigma_q^2\mathbf{I}_{MN_d}+\bar{\mathbf{Q}}\bar{\mathbf{\Upsilon}}\bar{\mathbf{Q}}^H\big)^{-1}\mathbf{V}_1^H$, and $\mathrm{EMSE}(\bar{\mathbf{Q}}) = \mathrm{tr}\left(\mathbb{E}\{\tilde{\mathbf{S}}\tilde{\mathbf{S}}^H\} - \mathbb{E}\{\tilde{\mathbf{S}}\mathbf{R}^H\}\mathbb{E}\{\mathbf{R}\mathbf{R}^H\}^{-1}\mathbb{E}\{\mathbf{R}\tilde{\mathbf{S}}^H\}\right)$, resulting in $\mathrm{EMSE}(\bar{\mathbf{Q}}) = \mathrm{tr}\left\{\bar{\mathbf{G}}^H\bar{\mathbf{\Upsilon}}^{-1}\bar{\mathbf{G}} - \bar{\mathbf{G}}^H\bar{\mathbf{Q}}^H\big(\sigma_q^2\mathbf{I}_{N_dM} + \bar{\mathbf{Q}}\bar{\mathbf{\Upsilon}}\bar{\mathbf{Q}}^H\big)^{-1}\bar{\mathbf{Q}}\bar{\mathbf{G}}\right\}$, which proves (12).

C. Proof of Proposition 2

It follows directly from the block diagonal structure of $\mathbf{Q}_m^H = \mathrm{blkdiag}(\mathbf{q}_{m,1}, \mathbf{q}_{m,2}, \cdots, \mathbf{q}_{m,N_d}) = [\mathbf{E}_1\mathbf{q}_{m,1}, \mathbf{E}_2\mathbf{q}_{m,2}, \cdots, \mathbf{E}_{N_d}\mathbf{q}_{m,N_d}]$ that $\mathrm{vec}\left(\mathbf{Q}_m^H\right) = \mathrm{blkdiag}(\mathbf{E}_1\mathbf{q}_{m,1}, \mathbf{E}_2\mathbf{q}_{m,2}, \cdots, \mathbf{E}_{N_d}\mathbf{q}_{m,N_d}) = \bar{\mathbf{E}}\bar{\mathbf{q}}_m$. We next reformulate the linear and quadratic terms in (14) with respect to $\bar{\mathbf{q}}_m$, noting that

$$\mathrm{tr}\left\{\Re\Big(\hat{\mathbf{G}}_m^H\mathbf{Q}_m^H\mathbf{\Phi}_m\Big)\right\} = \sum_{k=1}^{K}\Re\big(\hat{\mathbf{g}}_{m,k}^H\mathbf{Q}_m^H\boldsymbol{\phi}_{m,k}\big) \overset{(a)}{=} \Re\left\{\sum_{k=1}^{K}\Big(\boldsymbol{\phi}_{m,k}^T\otimes\hat{\mathbf{g}}_{m,k}^H\Big)\mathrm{vec}\left(\mathbf{Q}_m^H\right)\right\}, \tag{C.1}$$

where (a) follows the matrix vectorization rule for linear forms, $\mathrm{vec}(\mathbf{AXB}) = (\mathbf{B}^T\otimes\mathbf{A})\,\mathrm{vec}\,(\mathbf{X})$. Similarly, we have $\mathrm{tr}\left\{\mathbf{\Phi}_m^H\mathbf{Q}_m\mathbf{\Upsilon}_m\mathbf{Q}_m^H\mathbf{\Phi}_m\right\}\mathrm{vec}\left(\mathbf{Q}_m^H\right)^H\Big(\mathbf{\Phi}_m^*\mathbf{\Phi}_m^T\otimes\mathbf{\Upsilon}_m\Big)\mathrm{vec}\left(\mathbf{Q}_m^H\right)$ which follows from the formula $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD}$ and since $\boldsymbol{\phi}_{m,k}^H\mathbf{Q}_m = \mathrm{vec}\left(\mathbf{Q}_m^H\boldsymbol{\phi}_{m,k}\right)^H = \left\{\Big(\boldsymbol{\phi}_{m,k}^T\otimes\mathbf{I}_N\Big)\mathrm{vec}\left(\mathbf{Q}_m^H\right)\right\}^H = \mathrm{vec}\left(\mathbf{Q}_m^H\right)^H\left(\boldsymbol{\phi}_{m,k}^*\otimes\mathbf{I}_N\right)$ and $\mathbf{\Upsilon}_m\mathbf{Q}_m^H\boldsymbol{\phi}_{m,k} = \mathrm{vec}\left(\mathbf{\Upsilon}_m\mathbf{Q}_m^H\boldsymbol{\phi}_{m,k}\right) = \Big(\boldsymbol{\phi}_{m,k}^T\otimes\mathbf{\Upsilon}_m\Big)\mathrm{vec}\left(\mathbf{Q}_m^H\right)$. Substituting this along with (C.1) into (14) yields (17).

REFERENCES

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

[2] M. Jiang and L. Hanzo, "Multiuser MIMO-OFDM for next-generation wireless systems," *Proc. IEEE*, vol. 95, no. 7, pp. 1430–1469, Jul. 2007.

[3] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[4] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[5] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[6] N. Shlezinger, Y. C. Eldar, and M. R. D. Rodrigues, "Hardware-limited task-based quantization," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5223–5238, Oct. 2019.

[7] S. Salamatian, N. Shlezinger, Y. C. Eldar, and M. Médard, "Task-based quantization for recovering quadratic functions using principal inertia components," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 390–394.

[8] N. Shlezinger and Y. C. Eldar, "Deep task-based quantization," 2019, *arXiv:1908.06845*. [Online]. Available: http://arxiv.org/abs/1908.06845

[9] N. Shlezinger and Y. C. Eldar, "Task-based quantization with application to MIMO receivers," 2020, *arXiv:2002.04290*. [Online]. Available: http://arxiv.org/abs/2002.04290

[10] 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.

[11] 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.

[12] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432–1443, Jul. 2017.

[13] N. Shlezinger, Y. C. Eldar, and M. R. D. Rodrigues, "Asymptotic task-based quantization with application to massive MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 3995–4012, Aug. 2019.

[14] 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.

[15] 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. Appl.*, vol. 8, no. 5, Nov. 2017, Art. no. 054048.

[16] D. R. Smith *et al.*, "An analysis of beamed wireless power transfer in the fresnel zone using a dynamic, metasurface aperture," *J. Appl. Phys.*, vol. 121, no. 1, Jan. 2017, Art. no. 014901.

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

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

[19] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

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

[21] P. del Hougne, M. Fink, and G. Lerosey, "Optimally diverse communication channels in disordered environments with tuned randomness," *Nature Electron.*, vol. 2, no. 1, p. 36, 2019.

[22] W. Tang *et al.*, "Wireless communications with programmable metasurface: Transceiver design and experimental results," *China Commun.*, vol. 16, no. 5, pp. 46–61, 2019.

[23] J. Y. Dai *et al.*, "Wireless communications through a simplified architecture based on time-domain digital coding metasurface," *Adv. Mater. Technol.*, vol. 4, no. 7, 2019, Art. no. 1900044.

[24] H. Li, R. Liu, M. Liy, Q. Liu, and X. Li, "IRS-enhanced wideband MU-MISO-OFDM communication systems," in *Proc. IEEE WCNC*, May 2020, pp. 1–6.

[25] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, Aug. 2020.

[26] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, Aug. 2020.

[27] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, "Energy efficient multi-user MISO communication using low resolution large intelligent surfaces," in *Proc. IEEE Globecom*, Dec. 2018, pp. 1–6.

[28] 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.

[29] H. Wang *et al.*, "Dynamic metasurface antennas based downlink massive MIMO systems," in *Proc. IEEE SPAWC*, Jul. 2019, pp. 1–5.

[30] 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.

[31] N. Shlezinger, G. C. Alexandropoulos, M. F. Imani, Y. C. Eldar, and D. R. Smith, "Dynamic metasurface antennas for 6G extreme massive MIMO communications," 2020, *arXiv:2006.07838*. [Online]. Available: http://arxiv.org/abs/2006.07838

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

[33] 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.

[34] K. Roth, H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek, "A comparison of hybrid beamforming and digital beamforming with low-resolution ADCs for multiple users and imperfect CSI," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 484–498, Jun. 2018.
[35] D. R. Smith, J. B. Pendry, and M. C. K. Wiltshire, "Metamaterials and negative refractive index," *Science*, vol. 305, no. 5685, pp. 788–792, Aug. 2004.
[36] C. L. Holloway, E. F. Kuester, J. A. Gordon, J. O'Hara, J. Booth, and D. R. Smith, "An overview of the theory and applications of metasurfaces: The two-dimensional equivalents of metamaterials," *IEEE Antennas Propag. Mag.*, vol. 54, no. 2, pp. 10–35, Apr. 2012.
[37] J. Hunt *et al.*, "Metamaterial apertures for computational imaging," *Science*, vol. 339, no. 6117, pp. 310–313, Jan. 2013.
[38] J. Mo, P. Schniter, and R. W. Heath, Jr., "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, Mar. 2018.
[39] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.
[40] J. Choi, J. Mo, and R. W. Heath, Jr., "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, May 2016.
[41] R. M. Gray and T. G. Stockham, "Dithered quantizers," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 805–812, May 1993.
[42] B. Widrow, I. Kollar, and M. C. Liu, "Statistical theory of quantization," *IEEE Trans. Inst. Measure.*, vol. 45, no. 2, pp. 353–361, Apr. 1996.
[43] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.
[44] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
[45] 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.
[46] K. Shen, W. Yu, L. Zhao, and D. P. Palomar, "Optimization of MIMO device-to-device networks via matrix fractional programming: A minorization-maximization approach," *IEEE/ACM Trans. Netw.*, vol. 27, no. 5, pp. 2164–2177, Oct. 2019.
[47] Z.-Q. Luo, W.-K. Ma, A. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
[48] H. Wang, W.-T. Shih, C.-K. Wen, and S. Jin, "Reliable OFDM receiver with ultra-low resolution ADC," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3566–3579, May 2019.
[49] D. M. Pozar, *Microwave Engineering*. Hoboken, NJ, USA: Wiley, 2009.
[50] A. Alkhateeb and R. Heath, Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.
[51] H. Wang *et al.*, "Dynamic metasurface antennas for bit-constrained MIMO-OFDM receivers," in *Proc. ICASSP*, May 2020, pp. 9155–9159.
[52] S. Wei, D. L. Goeckel, and P. A. Kelly, "Convergence of the complex envelope of bandlimited OFDM signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4893–4904, Oct. 2010.

Hanqing Wang (Member, IEEE) received the B.S. degree in communications engineering from Anhui University, Hefei, China, in 2013, and the M.S. and Ph.D. degrees in information and communications engineering from Southeast University, Nanjing, China, in 2016 and 2020, respectively. From 2018 to 2019, he was a Visiting Student with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa, Israel. His research interests include detection and estimation theory, compressive sensing, and their application to communication systems with hardware imperfection and nonlinear distortion.

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

Yonina C. Eldar (Fellow, IEEE) received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering from Tel-Aviv University (TAU), Tel-Aviv, Israel, in 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, and an Adjunct Professor with Duke University and was a Visiting Professor at Stanford. She is a member of the Israel Academy of Sciences and Humanities (elected 2017), an IEEE Fellow and a EURASIP Fellow. Her research interests include in the broad areas of statistical signal processing, sampling theory and compressed sensing, learning and optimization methods, and their applications to biology and optics. She has received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award in 2013, the IEEE/AESS Fred Nathanson Memorial Radar Award in 2014, and the IEEE Kiyo Tomiyasu Award in 2016. She was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. 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 at the Technion, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technion's 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, and was selected as one of the 50 most influential women in Israel. She was a member of the Young Israel Academy of Science and Humanities and the Israel Committee for Higher Education. She is the Editor-in-Chief of *Foundations and Trends in Signal Processing*, a member of the IEEE Sensor Array and Multichannel Technical Committee and serves on several other IEEE committees. 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. She served as an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING, the *EURASIP Journal of Signal Processing*, the *SIAM Journal on Matrix Analysis and Applications*, and the *SIAM Journal on Imaging Sciences*. She was the Co-Chair and the Technical Co-Chair of several international conferences and workshops.

Shi Jin (Senior Member, IEEE) received the B.S. degree in communications engineering from the Guilin University of Electronic Technology, Guilin, China, in 1996, the M.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, and the Ph.D. degree in information and communications engineering from Southeast University, Nanjing, in 2007. From June 2007 to October 2009, he was a Research Fellow with the Adastral Park Research Campus, University College London, London, U.K. He is currently with the faculty of the National Mobile Communications Research Laboratory, Southeast University. His research interests include space time wireless communications, random matrix theory, and information theory. He and his coauthors have been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory and the 2010 Young Author Best Paper Award by the IEEE Signal Processing Society. He serves as an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE COMMUNICATIONS LETTERS, and *IET Communications*.

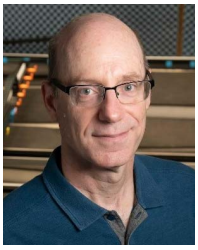
Mohammadreza F. Imani (Member, IEEE) received the B.S.E. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2007, and the M.S.E. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2010 and 2013, respectively.

From 2014 to 2018, he has served as a Post-Doctoral Associate with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA, where he was a Research Scientist, from 2018 to 2020. In August 2020, he joined the School of Electrical, Computer, and Energy Engineering, Arizona State University, as an Assistant Professor. His research interests include metamaterials and metasurfaces, microwave imaging and sensing, wireless power transfer, and communication systems.

Insang Yoo (Graduate Student Member, IEEE) received the B.S. (Hons.) and M.S. degrees in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2009 and 2011, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA.

From 2011 to 2013, he was with LG Uplus Corporation, Seoul. From 2013 to 2014, he was with the Natural Science Research Institute, Yonsei University. Since 2015, he has been with the Center for Metamaterials and Integrated Plasmonics, Duke University. His current research interests include metamaterials, and metasurface antennas and its applications in wireless communication systems.

Mr. Yoo received the Fulbright Foreign Student Program Fellowship from the U.S. Department of State, Bureau of Educational and Cultural Affairs, from 2015 to 2017.



David R. Smith (Senior Member, IEEE) received the B.S. and Ph.D. degrees in physics from the University of California at San Diego, San Diego, CA, USA, in 1988 and 1994, respectively. He is currently the James B. Duke Professor of electrical and computer engineering with Duke University and the Director of the Center for Metamaterials and Integrated Plasmonics. Since 2009, he has been listed as a highly cited researcher by Clarivate Analytics in the field of physics. His research interests include the theory, the simulation, and the characterization of unique electromagnetic structures, including photonic crystals and metamaterials, and the applications of such materials. He has provided key experimental demonstrations in the metamaterials field, including the first demonstration of a negative index metamaterial in 2000 and the first demonstration of a metamaterial invisibility cloak in 2006. He was elected as a fellow of the National Academy of Inventors in 2016. He was a co-recipient of the Descartes Scientific Research Prize from the European Union in 2005 and the James C. McGroddy Prize for New Materials from the American Physical Society in 2013.