A Framework on Complex Matrix Derivatives with Special Structure Constraints for Wireless Systems

Xin Ju, Shiqi Gong, Nan Zhao, *Senior Member, IEEE*, Chengwen Xing, Arumugam Nallanathan, *Fellow, IEEE*, and Dusit Niyato, *Fellow, IEEE*

Abstract—**Matrix-variate optimization plays a central role in advanced wireless system designs. In this paper, we aim to explore optimal solutions of matrix variables under two special structure constraints using complex matrix derivatives, including diagonal structure constraints and constant modulus constraints, both of which are closely related to the state-of-the-art wireless applications. Specifically, for diagonal structure constraints mostly considered in the uplink multi-user single-input multiple-output (MU-SIMO) system and the amplitude-adjustable intelligent reflecting surface (IRS)-aided multiple-input multiple-output (MIMO) system, the capacity maximization problem, the mean-squared error (MSE) minimization problem and their variants are rigorously investigated. By leveraging complex matrix derivatives, the optimal solutions of these problems are directly obtained in closed forms. Nevertheless, for constant modulus constraints with the intrinsic nature of element-wise decomposability, which are often seen in the hybrid analog-digital MIMO system and the fully-passive IRS-aided MIMO system, we firstly explore inherent structures of the element-wise phase derivatives associated with different optimization problems. Then, we propose a novel alternating optimization (AO) algorithm with the aid of several arbitrary feasible solutions, which avoids the complicated matrix inversion and matrix factorization involved in conventional element-wise iterative algorithms. Numerical simulations reveal that the proposed algorithm can dramatically reduce the computational complexity without loss of system performance.**

Index Terms—**Complex matrix derivatives, special structure constraints, matrix-variate optimization, hybrid analog-digital system, intelligent reflecting surface.**

I. Introduction

Multi-antenna technology opens a new era for wireless communications due to its effective utilization of limited spatial resources [1]–[3]. From the mathematical viewpoint, the deployment of multi-antenna arrays at transceivers generally leads to matrix-variate optimization problems [4]–[6]. Specifically, in the typical multiple-input multiple-output (MIMO) communication systems, the transmit beamformer optimization and the receive equalizer optimization can be both modeled as matrix-variate optimization problems [7]–[9]. Compared to scalar-variate optimization, matrix-variate optimization is generally more challenging to tackle because it inherently involves complex matrix operations, including matrix determinant, inversion, matrix decomposition and so on. In fact, with the development of wireless communications, many matrix-variate optimization problems with special structure constraints such as symmetric, diagonal and constant modulus structure constraints are emerging, which are closely related to the state-of-the-art wireless systems equipped with multi-antenna transceiver antenna arrays.

In general, a structure of the matrix variable strongly depends on three factors, namely, network architectures, frequency bands and communication demands. First, the distributed network architecture has been studied, since its involved distributed antenna arrays are capable of increasing spatial diversity gain and extending communication coverage, as compared with the centralized counterpart [10]. In this distributed network, the corresponding matrix variable usually has a diagonal structure. Second, for high-frequency millimeter wave (mmWave) and terahertz (THz) communications, a hybrid analog-digital transceiver structure has been regarded as an economic and effective way to achieve a large array gain [11], [12], in which the analog beamforming matrix is usually subject to the nonconvex and intractable constant modulus constraints. Third, for smartly reconfiguring the wireless environment in a cost-effective manner, intelligent reflecting surfaces (IRSs) composed of a large number of passive reflecting elements have attracted a lot of attention recently [13]–[15]. Considering different levels of hardware implementation, there are two main types of IRS structures that are widely studied, i.e., the amplitude-adjustable IRS and the fully-passive IRS. Note that the reflection matrices of these two IRSs can be mathematically modeled as diagonal matrices. In particular, for the fully-passive IRS, the corresponding reflection matrix is additionally subject to the nonconvex constant modulus constraint. Building upon the above discussions, it is clear that matrix-variate optimization problems with special structure constraints have been widely considered in the state-of-the-art wireless systems. Therefore, it is essential to develop a framework for optimization algorithms with guaranteed performance and low complexity for the matrix-variate optimization.

Currently, there have been many common popular algorithms for solving matrix-variate optimization problems, such as the Karush-Kuhn-Tucker (KKT)-based algorithm [16]–[19], the block coordinate descent (BCD) algorithm [20] and the majorization-minimization (MM)-based algorithm [21]. It is well-known that for convex matrix-variate problems, the KKT-based algorithm is able to directly derive the optimal structures of matrices. Generally, the symmetric structure constraints of matrix variables can be implicitly satisfied by the derived optimal closed-form solutions [22]. Moreover,

X. Ju and C. Xing are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (E-mails: xinjubit@gmail.com; chengwenxing@ieee.org).

S. Gong is with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China (E-mail: gsqyx@163.com).

N. Zhao is with the School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China (E-mail: zhaonan@dlut.edu.cn).

Arumugam Nallanathan is with the QueenMary University of London, E14NS London, U.K. (E-mail: a.nallanathan@qmul.ac.uk).

Dusit Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (E-mail: dniyato@ntu.edu.sg).

TABLE I

COMPLEX MATRIX DERIVATIVES UNDER DIAGONAL STRUCTURE CONSTRAINTS

Function Type	**Derivative w.r.t. $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}$**	**Derivative w.r.t. $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}^*$**
$f_{\mathrm{D,TL}} = \mathrm{Tr}(\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}^{\mathrm{H}}\boldsymbol{M}) + \mathrm{Tr}(\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\boldsymbol{M}^{\mathrm{H}})$	$\mathrm{Diag}\{\boldsymbol{M}^{\mathrm{H}}\}$	$\mathrm{Diag}\{\boldsymbol{M}\}$
$f_{\mathrm{D,TQ}} = \mathrm{Tr}(\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}^{\mathrm{H}}\boldsymbol{W}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}})$	$\mathrm{Diag}\{\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}^{\mathrm{H}}\boldsymbol{W}\}$	$\mathrm{Diag}\{\boldsymbol{W}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\}$
$f_{\mathrm{D,TI}} = \mathrm{Tr}\left((\boldsymbol{I}_M + \boldsymbol{\Phi}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}})^{-1}\right)$	$-\mathrm{Diag}\left\{\boldsymbol{\Phi}^{\frac{1}{2}}(\boldsymbol{I}_M + \boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\boldsymbol{\Phi}^{\frac{1}{2}})^{-2}\boldsymbol{\Phi}^{\frac{1}{2}}\right\}$	0
$f_{\mathrm{D,LD}} = \log\lvert\boldsymbol{I}_M + \boldsymbol{\Phi}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\rvert$	$\mathrm{Diag}\left\{\boldsymbol{\Phi}^{\frac{1}{2}}(\boldsymbol{I}_M + \boldsymbol{\Phi}^{\frac{1}{2}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\boldsymbol{\Phi}^{\frac{1}{2}})^{-1}\boldsymbol{\Phi}^{\frac{1}{2}}\right\}$	0

for diagonal structure constraints, this algorithm considers applying the first-order derivative to each diagonal element to obtain the optimal solution. Furthermore, in terms of the intractable constant modulus constraints, dual variables are usually introduced and iteratively optimized by the subgradient method to satisfy complementary slackness conditions, thereby potentially suffering from high iteration complexity [23]. In contrast, the BCD algorithm is always adopted to solve highly nonconvex problems caused by strongly-coupled matrix variables. Specifically, under the BCD framework, the original matrix-variate optimization problem can be decomposed into multiple low-dimensional subproblems, each of which needs to be iteratively optimized until convergence. In order to ensure at least local convergence of the BCD algorithm [20], each subproblem is required to have a unique optimal solution. Nevertheless, considering that subproblems may be nonconvex and thus hard to globally solve, the MM-based algorithm has attained extensive attention, whose core idea is to construct a tractable surrogate function to locally approximate the original nonconvex subproblem. Unfortunately, the derivation of the surrogate function usually involves high-complexity matrix manipulations and also needs to be iteratively carried out to achieve a close approximation.

Obviously, the KKT-based algorithm based on complex matrix derivatives generally achieves the lowest complexity among the three types of algorithms [24]. Nonetheless, its application range is relatively limited as compared to the BCD and MM-based algorithms. Note that the implementation of the latter two algorithms depends on the specific wireless system and may have high computational complexity, especially for large-scale arrays. To circumvent these issues, in this paper, we aim to develop a unified framework for matrix-variate optimization with two special structure constraints, namely, diagonal structure and constant modulus constraints. For each considered case, the novel low-complexity algorithm with guaranteed performance is proposed. The main contributions of our work are further summarized as follows.

- Firstly, we consider the diagonal structure constraints often seen in the uplink multi-user single-input multiple-output (MU-SIMO) system and the amplitude-adjustable IRS-aided MIMO system, which are always involved in the capacity maximization problem, mean squared error (MSE) minimization problem and their variants. We propose complex matrix derivatives associated with diagonal structures, based on which the optimal solutions of these matrix-variate problems are directly obtained in closed forms. Furthermore, the above study is extended to the case of block-diagonal structure constraints.
- Secondly, in terms of constant modulus constraints mostly adopted in the hybrid analog-digital MIMO system and the fully-passive IRS-aided MIMO system, we propose the element-wise phase derivatives inspired by their element-wise decomposability nature. For different classical matrix-variate optimization problems, it is revealed that the element-wise phase derivatives can be classified into the following two general forms, i.e., the linear form and the conjugate linear form.
- Finally, by exploring inherent structures of the element-wise phase derivatives, we develop a novel alternating optimization (AO) algorithm with the aid of several arbitrary feasible solutions for the matrix-variate optimization under constant modulus constraints. Note that the computational complexity of the proposed AO algorithm sharply decreases, since it avoids the complicated matrix inversion and matrix factorization involved in the conventional element-wise iterative algorithm. Moreover, we demonstrate that the proposed algorithm is able to achieve almost the same performance as the existing benchmark schemes.

Notation: Scalars, vectors and matrices are represented by non-bold, bold lowercase, and bold uppercase letters, respectively. The notations $\boldsymbol{A}^{\mathrm{T}}$, $\boldsymbol{A}^*$, $\boldsymbol{A}^{\mathrm{H}}$, $\boldsymbol{A}^{-1}$, $\mathrm{Tr}(\boldsymbol{A})$ and $|\boldsymbol{A}|$ denote the transpose, conjugate, hermitian, inversion, trace and determinant of the complex matrix $\boldsymbol{A}$, respectively. $\mathrm{Diag}\{\boldsymbol{A}\}$ denotes a vector whose elements are diagonal elements of matrix $\boldsymbol{A}$, and $\mathrm{Blockdiag}(\{\boldsymbol{A}_k\}_{k=1}^K)$ is a block diagonal matrix with diagonal sub-matrices of $\boldsymbol{A}_k$'s. Moreover, the ith row and the jth column of $\boldsymbol{A}$ are denoted as $[\boldsymbol{A}]_{i,:}$ and $[\boldsymbol{A}]_{:,j}$, respectively, the element in the ith row and the jth column is denoted as $[\boldsymbol{A}]_{i,j}$. $\frac{\mathrm{d}f}{\mathrm{d}a}$ and $\frac{\partial f}{\partial a}$ denote the differential and the partial derivative of f with respect to a, respectively. $\odot$ denotes the Hadamard product and $(a)^+ = \max\{0, a\}$. $\Re\{a\}$ and $\Im\{a\}$ denote the real and imaginary parts of a complex variable a, respectively. The symbol $\mathrm{Phase}\{a\}$ denotes the phase of a complex scalar a and $-\pi < \mathrm{Phase}\{a\} \leq \pi$. Lastly, the word "with respect to" is abbreviated as "w.r.t.".

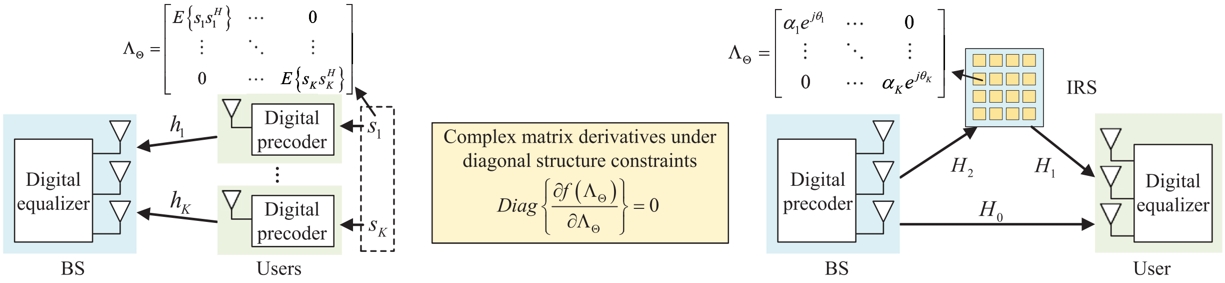
II. DIAGONAL STRUCTURE CONSTRAINTS

In this section, we firstly provide some fundamental properties of complex matrix derivatives associated with diagonal structures for several types of objective functions. Based on these properties, we then obtain the optimal solutions of a series of optimization problems in the uplink MU-SIMO system and the amplitude-adjustable IRS-aided MIMO system in closed forms. Moreover, the above study is extended to the case of block-diagonal matrix variables.

A. Mathematical Preliminaries

At the beginning, some fundamental definitions for diagonal matrices are provided, which are the basis of the following analysis. For any two diagonal matrices $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_1} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_2} \in \mathbb{C}^{M\times M}$, we have

$$\mathrm{Diag}\{\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_1} + \boldsymbol{\Lambda}_{\boldsymbol{\Theta}_2}\} = \mathrm{Diag}\{\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_1}\} + \mathrm{Diag}\{\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_2}\}. \quad (1)$$


$\Lambda_\Theta = \begin{bmatrix} E\{s_1 s_1^H\} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & E\{s_K s_K^H\} \end{bmatrix}$
Digital equalizer
h_1
h_K
Digital precoder
Digital precoder
s_1
s_K
BS
Users
Complex matrix derivatives under diagonal structure constraints
$Diag\left\{\frac{\partial f(\Lambda_\Theta)}{\partial \Lambda_\Theta}\right\} = 0$
$\Lambda_\Theta = \begin{bmatrix} \alpha_1 e^{j\theta_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_K e^{j\theta_K} \end{bmatrix}$
IRS
Digital precoder
H_2
H_1
H_0
Digital equalizer
BS
User

(a) Uplink MU-MISO system

(b) Amplitude-adjustable IRS-aided MIMO system

Fig. 1. A diagram of application scenarios associated with diagonal matrix variables.

For an arbitrary square matrix $\mathbf{\Phi} \in \mathbb{C}^{M\times M}$, we have

$$\mathrm{Diag}\{\mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{\Theta}}\} = \mathrm{Diag}\{\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Phi}\}. \tag{2}$$

Moreover, the following equality holds for any complex matrices $\boldsymbol{N} \in \mathbb{C}^{N\times M}$ and $\boldsymbol{M} \in \mathbb{C}^{M\times N}$, i.e.,

$$\mathrm{Diag}\{\boldsymbol{N}\mathbf{\Lambda}_{\mathbf{\Theta}}\boldsymbol{M}\} = \mathbf{\Phi}\mathrm{Diag}\{\mathbf{\Lambda}_{\mathbf{\Theta}}\} \text{ with } \mathbf{\Phi} = \boldsymbol{N} \odot \boldsymbol{M}^{\mathrm{T}}. \tag{3}$$

Together with the operation of $\mathrm{Diag}\{\cdot\}$, the vectorization of the first-order derivative of a scalar-valued function $f(\mathbf{\Lambda}_{\mathbf{\Theta}})$ w.r.t the complex diagonal matrix $\mathbf{\Lambda}_{\mathbf{\Theta}}$ can be defined as [24]

$$\mathrm{Diag}\left\{\frac{\partial f(\mathbf{\Lambda}_{\mathbf{\Theta}})}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} = \left[\frac{\partial f(\mathbf{\Lambda}_{\mathbf{\Theta}})}{\partial [\mathbf{\Lambda}_{\mathbf{\Theta}}]_{1,1}}, \cdots, \frac{f(\mathbf{\Lambda}_{\mathbf{\Theta}})}{\partial [\mathbf{\Lambda}_{\mathbf{\Theta}}]_{N,N}}\right]^{\mathrm{T}},$$
$$\mathrm{Diag}\left\{\frac{\partial f(\mathbf{\Lambda}_{\mathbf{\Theta}})}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}^*}\right\} = \left[\frac{\partial f(\mathbf{\Lambda}_{\mathbf{\Theta}})}{\partial [\mathbf{\Lambda}_{\mathbf{\Theta}}]_{1,1}^*}, \cdots, \frac{f(\mathbf{\Lambda}_{\mathbf{\Theta}})}{\partial [\mathbf{\Lambda}_{\mathbf{\Theta}}]_{N,N}^*}\right]^{\mathrm{T}}. \tag{4}$$

In the sequel, we mainly concern about complex matrix derivatives w.r.t. diagonal matrices for four common objective functions, including the trace-linear function, the trace-quadratic function, the trace-inverse function and the log-determinant function.

1) Trace-Linear Function: For an arbitrary complex matrix $\boldsymbol{M} \in \mathbb{C}^{M\times M}$, the differential of a trace-linear function $f_{\mathrm{D,TL}} = \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}\boldsymbol{M}) + \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\boldsymbol{M}^{\mathrm{H}})$ w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ can be obtained as $\mathrm{d}\left(f_{\mathrm{D,TL}}\right) = \mathrm{Tr}\left(\mathrm{d}\left(\mathbf{\Lambda}_{\mathbf{\Theta}}\right)\boldsymbol{M}^{\mathrm{H}}\right)$. Based on this, the corresponding first-order derivative w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ can be obtained as

$$\mathrm{Diag}\left\{\frac{\partial f_{\mathrm{D,TL}}}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} = \mathrm{Diag}\{\boldsymbol{M}^{\mathrm{H}}\}. \tag{5}$$

The first-order derivative of $f_{\mathrm{D,TL}}$ w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}^*$ is also given by

$$\mathrm{Diag}\left\{\frac{\partial f_{\mathrm{D,TL}}}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}^*}\right\} = \mathrm{Diag}\{\boldsymbol{M}\}. \tag{6}$$

2) Trace-Quadratic Function: For a Hermitian matrix $\boldsymbol{W} \in \mathbb{C}^{M\times M}$, the differential of a trace-quadratic function $f_{\mathrm{D,TQ}} = \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}\boldsymbol{W}\mathbf{\Lambda}_{\mathbf{\Theta}})$ w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ and $\mathbf{\Lambda}_{\mathbf{\Theta}}^*$ are respectively calculated as $\mathrm{d}\left(f_{\mathrm{D,TQ}}\right) = \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}\boldsymbol{W}\mathrm{d}(\mathbf{\Lambda}_{\mathbf{\Theta}})), \mathrm{d}\left(f_{\mathrm{D,TQ}}\right) = \mathrm{Tr}(\mathrm{d}(\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}})\boldsymbol{W}\mathbf{\Lambda}_{\mathbf{\Theta}})$. Then, the corresponding first-order derivatives w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ and $\mathbf{\Lambda}_{\mathbf{\Theta}}^*$ can be obtained as

$$\mathrm{Diag}\left\{\frac{\partial f_{\mathrm{D,TQ}}}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} = \mathrm{Diag}\{\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}\boldsymbol{W}\},$$
$$\mathrm{Diag}\left\{\frac{\partial f_{\mathrm{D,TQ}}}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}^*}\right\} = \mathrm{Diag}\{\boldsymbol{W}\mathbf{\Lambda}_{\mathbf{\Theta}}\}. \tag{7}$$

3) Trace-Inverse Function: For a positive semi-definite matrix $\mathbf{\Phi} \in \mathbb{C}^{M\times M}$, the differential of a trace-inverse function $f_{\mathrm{D,TI}} = \mathrm{Tr}\left((\boldsymbol{I}_M + \mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{\Theta}})^{-1}\right)$ w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ is given by $\mathrm{d}\left(f_{\mathrm{D,TI}}\right) = -\mathrm{Tr}\left((\boldsymbol{I}_M + \mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{\Theta}})^{-2}\mathbf{\Phi}\mathrm{d}\mathbf{\Lambda}_{\mathbf{\Theta}}\right)$. The corresponding first-order derivative w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ is then given by

$$\mathrm{Diag}\left\{\frac{\partial f_{\mathrm{D,TI}}}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} = -\,\mathrm{Diag}\left\{(\boldsymbol{I}_M + \mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{\Theta}})^{-2}\mathbf{\Phi}\right\} \tag{8}$$
$$= -\,\mathrm{Diag}\left\{\mathbf{\Phi}^{\frac{1}{2}}(\boldsymbol{I}_M + \mathbf{\Phi}^{\frac{1}{2}}\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Phi}^{\frac{1}{2}})^{-2}\mathbf{\Phi}^{\frac{1}{2}}\right\}.$$

4) Log-Determinant Function: Considering a log-determinant function $f_{\mathrm{D,LD}} = \log|\boldsymbol{I}_M + \mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{\Theta}}|$, its differential w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ can be derived as $\mathrm{d}\left(f_{\mathrm{D,LD}}\right) = \mathrm{Tr}\left((\boldsymbol{I}_M + \mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{\Theta}})^{-1}\mathbf{\Phi}\mathrm{d}\mathbf{\Lambda}_{\mathbf{\Theta}}\right)$, based on which the following first-order derivative w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ holds.

$$\mathrm{Diag}\left\{\frac{\partial f_{\mathrm{D,LD}}}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} = \mathrm{Diag}\left\{(\boldsymbol{I}_M + \mathbf{\Phi}\mathbf{\Lambda}_{\mathbf{\Theta}})^{-1}\mathbf{\Phi}\right\} \tag{9}$$
$$= \mathrm{Diag}\left\{\mathbf{\Phi}^{\frac{1}{2}}(\boldsymbol{I}_M + \mathbf{\Phi}^{\frac{1}{2}}\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Phi}^{\frac{1}{2}})^{-1}\mathbf{\Phi}^{\frac{1}{2}}\right\}.$$

In conclusion, the complex matrix derivatives of the above four types of objective functions are summarized in Table I. Exploiting these fundamental properties, some classical wireless applications are investigated in the following subsection.

B. Specific Wireless Applications

1) Uplink MU-SIMO System: As shown in Fig. 1(a), we firstly consider an uplink distributed MU-SIMO system, where K single-antenna users transmit independent data streams to the BS equipped with N_t antennas [10]. The received signal at the BS can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n} \quad \text{with} \quad \boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K] \in \mathbb{C}^{N_{\mathrm{t}}\times K}, \tag{10}$$

where $\boldsymbol{h}_k \in \mathbb{C}^{N_{\mathrm{t}}\times 1}$ denotes the channel between the BS and the kth user, and $\boldsymbol{s} = [s_1, \cdots, s_K]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ is the transmit signal, $\boldsymbol{n}$ is the additive noise obeying Gaussian distribution with zero mean and covariance matrix $\mathbb{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}\} = \mathbf{\Sigma}$. It is worth noting that the covariance matrix of $\boldsymbol{s}$ is diagonal, since the transmit data streams of K users are independent of each other, i.e., $\mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}\} = \mathbf{\Lambda}_{\mathbf{\Theta}}$. Then, the uplink capacity maximization problem is formulated as

$$\textbf{Prob.1:} \max_{\mathbf{\Lambda}_{\mathbf{\Theta}}} \log\left|\boldsymbol{I}_{N_t} + \mathbf{\Sigma}^{-1}\boldsymbol{H}\mathbf{\Lambda}_{\mathbf{\Theta}}\boldsymbol{H}^{\mathrm{H}}\right|$$
$$\text{s.t. } \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}) \le P, \; 0 \le [\mathbf{\Lambda}_{\mathbf{\Theta}}]_{k,k} \le P_k, \; \forall k. \tag{11}$$

By recalling (9), the first-order derivative of the objective function of **Prob.1** w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ is given by

$$\mathrm{Diag}\left\{\frac{\partial \log|\boldsymbol{I}_{N_t} + \mathbf{\Sigma}^{-1}\boldsymbol{H}\mathbf{\Lambda}_{\mathbf{\Theta}}\boldsymbol{H}^{\mathrm{H}}|}{\partial \mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} \tag{12}$$
$$= \mathrm{Diag}\left\{\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{I}_{N_t} + \mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\mathbf{\Lambda}_{\mathbf{\Theta}}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\right)^{-1}\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\right\}.$$

In order to derive the optimal $\mathbf{\Lambda_\Theta}$, we present the KKT optimality conditions of **Prob.1** as follows [25]:

$$\begin{cases} \mathrm{Diag}\Big\{\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\mathbf{\Lambda_\Theta}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\right)^{-1} \\ \qquad \times\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\Big\}=\mu\mathrm{Diag}\left\{\boldsymbol{I}_K\right\}-[\psi_1,\cdots,\psi_K]^{\mathrm{T}}, & \text{(13a)} \\ \mu\left(\mathrm{Tr}(\mathbf{\Lambda_\Theta})-P\right)=0, & \text{(13b)} \\ \psi_k\left([\mathbf{\Lambda_\Theta}]_{k,k}-P_k\right)=0,\ \forall k, & \text{(13c)} \end{cases}$$

where μ is the Lagrange multiplier associated with the sum power constraint and ψ_k's correspond to the power constraints imposed on each user. We firstly define $\boldsymbol{J}_k=\boldsymbol{I}_K+\left(\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}-\overline{\boldsymbol{H}}_k\right)\left(\mathbf{\Lambda_\Theta}-\overline{\mathbf{\Lambda_\Theta}}_k\right)$, where $\overline{\boldsymbol{H}}_k\in\mathbb{C}^{K\times K}$ is an all-zero matrix except for its kth column being $\left[\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right]_{:,k}$ and $\overline{\mathbf{\Lambda_\Theta}}_k\in\mathbb{C}^{K\times K}$ is an all-zero matrix except for its (k,k)th element being $[\mathbf{\Lambda_\Theta}]_{k,k}$. Then, the left-hand side of (13a) can be rewritten as

$$\begin{aligned} &\mathrm{Diag}\left\{\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\mathbf{\Lambda_\Theta}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\right)^{-1}\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\right\} \\ \overset{(a_1)}{=}&\mathrm{Diag}\left\{\left(\boldsymbol{I}_K+\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\mathbf{\Lambda_\Theta}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\} \\ \overset{(a_2)}{=}&-\frac{[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}}{1+[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)} \\ &+\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}, \end{aligned} \tag{14}$$

where (a_1) holds based on the matrix inversion lemma $\boldsymbol{N}(\boldsymbol{I}+\boldsymbol{M}\boldsymbol{N})^{-1}=(\boldsymbol{I}+\boldsymbol{N}\boldsymbol{M})^{-1}\boldsymbol{N}$ and (a_2) is attained using the Sherman Morrison formula, i.e., $(\boldsymbol{M}+\boldsymbol{N})^{-1}=\boldsymbol{M}^{-1}-\frac{\boldsymbol{M}^{-1}\boldsymbol{N}\boldsymbol{M}^{-1}}{1+\mathrm{Tr}(\boldsymbol{M}^{-1}\boldsymbol{N})}$ for a full-rank matrix $\boldsymbol{M}$ and a rank-one matrix $\boldsymbol{N}$. By substituting (14) into (13a), the optimal $[\mathbf{\Lambda_\Theta}]_{k,k}$'s is derived in the following water-filling form [26].

$$[\mathbf{\Lambda_\Theta}]_{k,k}=\begin{cases}\left(\frac{y_k}{\left[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}\right]_k+\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)y_k}\right)^{+}, \\ \qquad\qquad \text{if } [\mathbf{\Lambda_\Theta}]_{k,k}\leq P_k, \\ P_k, \qquad \text{if } [\mathbf{\Lambda_\Theta}]_{k,k}>P_k,\ \forall k, \end{cases} \tag{15}$$

where $y_k=\left[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}\right]_k-\mu$. Moreover, since $\mathrm{Tr}(\mathbf{\Lambda_\Theta})$ is monotonically decreasing w.r.t $\mu>0$, the optimal μ satisfying (13b) can be found via the bisection search.

In addition, MSE is a widely used performance metric, which reflects the accuracy of the desired signals that can be recovered from the noise corrupted observations. Accordingly, the MSE minimization problem is formulated as

$$\begin{aligned} \textbf{Prob.2:}\ \min_{\mathbf{\Lambda_\Theta}}\ &\mathrm{Tr}\left[\left(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-1}\boldsymbol{H}\mathbf{\Lambda_\Theta}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\right] \\ \text{s.t. }&\mathrm{Tr}(\mathbf{\Lambda_\Theta})\leq P,\ 0\leq[\mathbf{\Lambda_\Theta}]_{k,k}\leq P_k,\ \forall k. \end{aligned} \tag{16}$$

By recalling (8), the first-order derivative of the objective function of **Prob.2** w.r.t. $\mathbf{\Lambda_\Theta}$ is given by

$$\begin{aligned} &\mathrm{Diag}\left\{\frac{\partial\mathrm{Tr}\left[(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-1}\boldsymbol{H}\mathbf{\Lambda_\Theta}\boldsymbol{H}^{\mathrm{H}})^{-1}\right]}{\partial\mathbf{\Lambda_\Theta}}\right\} \qquad (17) \\ =&-\mathrm{Diag}\left\{\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\mathbf{\Lambda_\Theta}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\right)^{-2}\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\right\}. \end{aligned}$$

Based on (17), the KKT optimality conditions of **Prob.2** can be formulated as

$$\begin{cases} \mathrm{Diag}\left\{\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\mathbf{\Lambda_\Theta}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\right)^{-2}\right. \\ \qquad \left.\times\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\right\}=\mu\mathrm{Diag}\left\{\boldsymbol{I}_K\right\}-[\psi_1,\cdots,\psi_K]^{\mathrm{T}}, & \text{(18a)} \\ \mu\left(\mathrm{Tr}(\mathbf{\Lambda_\Theta})-P\right)=0, & \text{(18b)} \\ \psi_k\left([\mathbf{\Lambda_\Theta}]_{k,k}-P_k\right)=0,\ \forall k, & \text{(18c)} \end{cases}$$

where μ and ψ_k's are defined similarly to **Prob.1**. The left-hand side of (18a) can be rewritten as

$$\begin{aligned} &\mathrm{Diag}\left\{\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\mathbf{\Lambda_\Theta}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\right)^{-2}\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\right\} \\ \overset{(b)}{=}&\mathrm{Diag}\left\{\left(\boldsymbol{J}_k^{-1}-\frac{[\mathbf{\Lambda_\Theta}]_{k,k}\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}}{1+[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)}\right)^2\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\} \\ =&\frac{[\mathbf{\Lambda_\Theta}]_{k,k}^2\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-2}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}}{\left(1+[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)\right)^2} \\ &-\frac{[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Diag}\left\{\boldsymbol{J}_k^{-2}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}}{1+[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)} \\ &-\frac{[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-2}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}}{1+[\mathbf{\Lambda_\Theta}]_{k,k}\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)} \\ &+\mathrm{Diag}\left\{\boldsymbol{J}_k^{-2}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}, \end{aligned} \tag{19}$$

where (b) holds similarly to (14). Substituting (19) into (18a), we have

$$\widetilde{a}_k[\mathbf{\Lambda_\Theta}]_{k,k}^2+\widetilde{b}_k[\mathbf{\Lambda_\Theta}]_{k,k}+x_k=0,\ \forall k, \tag{20}$$

where

$$\begin{aligned} \widetilde{a}_k=&\mathrm{Tr}^2(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)x_k+\Big[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-2}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\right. \\ &\left.\boldsymbol{H}\right\}\Big]_k-\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-2}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}]_k \\ &-\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-2}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}]_k, \\ \widetilde{b}_k=&2\mathrm{Tr}(\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k)x_k-[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-2}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-1}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}]_k \\ &-[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-1}\overline{\boldsymbol{H}}_k\boldsymbol{J}_k^{-2}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}]_k, \\ x_k=&\left[\mathrm{Diag}\left\{\boldsymbol{J}_k^{-2}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-1}\boldsymbol{H}\right\}\right]_k-\mu,\ \forall k. \end{aligned} \tag{21}$$

Then, the optimal $[\mathbf{\Lambda_\Theta}]_{k,k}$'s of **Prob.2** can be attained using quadratic formula as follows:

$$[\mathbf{\Lambda_\Theta}]_{k,k}=\begin{cases}\left(\frac{-\widetilde{b}\pm\sqrt{\widetilde{b}^2-4[\boldsymbol{x}_k]_k\widetilde{a}}}{2\widetilde{a}}\right)^{+}, & \text{if } [\mathbf{\Lambda_\Theta}]_{k,k}\leq P_k, \\ P_k, & \text{if } [\mathbf{\Lambda_\Theta}]_{k,k}>P_k, \end{cases}\ \forall k. \tag{22}$$

Similarly, μ can be obtained using the bisection search.

2) Amplitude-Adjustable IRS-aided MIMO System: Hereafter, we consider the state-of-the-art amplitude-adjustable IRS-aided point-to-point MIMO system as shown in Fig. 1(b), where a N_t-antenna BS serves a N_r-antenna user with the aid of a K-element amplitude-adjustable IRS. The received signal at the user can be expressed as

$$\boldsymbol{y}=\boldsymbol{H}\boldsymbol{s}+\boldsymbol{n}\ \text{ with }\ \boldsymbol{H}=\boldsymbol{H}_0+\boldsymbol{H}_1\mathbf{\Lambda_\Theta}\boldsymbol{H}_2, \tag{23}$$

where $\boldsymbol{H}_0\in\mathbb{C}^{N_r\times N_t}$, $\boldsymbol{H}_1\in\mathbb{C}^{N_r\times K}$ and $\boldsymbol{H}_2\in\mathbb{C}^{K\times N_t}$ represent the BS-user direct channel, the IRS-user channel and the BS-IRS channel, respectively. $\mathbf{\Lambda_\Theta}\in\mathbb{C}^{K\times K}$ denotes the diagonal

IRS reflection matrix whose each diagonal element represents the adjustable amplitude and phase of the corresponding reflecting element, which usually satisfies $\left|[\mathbf{\Lambda}_{\mathbf{\Theta}}]_{i,i}\right| \leq 1$ or $\mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}) \leq K$ [27]. Similar to Sec. II-B1, we firstly consider the following capacity maximization problem.

$$\begin{aligned}\textbf{Prob.3:}\ \max_{\mathbf{\Lambda}_{\mathbf{\Theta}}}\ & \log\left|\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right| \\ \text{s.t. } & \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}) \leq K. \end{aligned} \tag{24}$$

By recalling (7) and (9), the first-order derivative of the objective function of **Prob.3** w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ is given by

$$\begin{aligned}&\mathrm{Diag}\left\{\frac{\partial \log\left|\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right|}{\partial\mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} \\ =&\mathrm{Diag}\left\{\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}\boldsymbol{H}_1^{\mathrm{H}}\left[\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right]^{-1}\mathbf{\Sigma}^{-1}\boldsymbol{H}_1\right\} \\ &+\mathrm{Diag}\left\{\boldsymbol{H}_2\boldsymbol{H}_0^{\mathrm{H}}\left[\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right]^{-1}\mathbf{\Sigma}^{-1}\boldsymbol{H}_1\right\}. \end{aligned} \tag{25}$$

Unfortunately, even though the first-order derivative is derived, it is still difficult to derive the optimal closed-form solution from (25), since its involved quadratic term w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ appears in an inverse form. As a remedy, we intend to solve it based on problem transformation. Specifically, via introducing a series of auxiliary variables, **Prob.3** is equivalently transformed into

$$\begin{aligned}\textbf{Prob.4:}\ \min_{\boldsymbol{G}_{\mathbf{\Lambda}},\mathbf{\Lambda}_{\mathbf{\Theta}},\boldsymbol{W}}\ & \mathrm{Tr}\left(\boldsymbol{W}\left[\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}-\boldsymbol{I}_{N_t}\right]\left[\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}-\boldsymbol{I}_{N_t}\right]^{\mathrm{H}}\right) \\ &+\mathrm{Tr}(\boldsymbol{W}\boldsymbol{G}_{\mathbf{\Lambda}}\mathbf{\Sigma}\boldsymbol{G}_{\mathbf{\Lambda}}^{\mathrm{H}})-\log|\boldsymbol{W}| \\ \text{s.t. } & \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}) \leq K. \end{aligned} \tag{26}$$

The equivalence between **Prob.3** and **Prob.4** is built based on the idea of weighted MSE minimization (WMMSE) [28], [29]. Then, **Prob.4** can be efficiently solved via the AO among $\boldsymbol{G}_{\mathbf{\Lambda}}$, $\mathbf{\Lambda}_{\mathbf{\Theta}}$ and $\boldsymbol{W}$. Specifically, both optimal $\boldsymbol{G}_{\mathbf{\Lambda}}$ and $\boldsymbol{W}$ can be directly derived by taking the first-order derivatives of the objective function of **Prob.4** w.r.t. $\boldsymbol{G}_{\mathbf{\Lambda}}$ and $\boldsymbol{W}$ to zeros, i.e., $\boldsymbol{G}_{\mathbf{\Lambda}}=\boldsymbol{H}^{\mathrm{H}}\left(\mathbf{\Sigma}+\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}, \boldsymbol{W}=\left(\boldsymbol{I}_{N_t}-\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}\right)^{-1}$. Then, the optimization problem w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ can be written as

$$\begin{aligned}\textbf{Prob.5:}\ \min_{\mathbf{\Lambda}_{\mathbf{\Theta}}}\ & \mathrm{Tr}\left(\boldsymbol{W}\left[\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}-\boldsymbol{I}_{N_t}\right]\left[\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}-\boldsymbol{I}_{N_t}\right]^{\mathrm{H}}\right) \\ \text{s.t. } & \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}) \leq K. \end{aligned} \tag{27}$$

The first-order derivative of the objective function of **Prob.5** w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ is given by

$$\begin{aligned}&\mathrm{Diag}\left\{\frac{\partial\mathrm{Tr}\left(\boldsymbol{W}\left[\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}-\boldsymbol{I}_{N_t}\right]\left[\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}-\boldsymbol{I}_{N_t}\right]^{\mathrm{H}}\right)}{\partial\mathbf{\Lambda}_{\mathbf{\Theta}}}\right\} \\ =&\mathrm{Diag}\{\boldsymbol{H}_2\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}_{\mathbf{\Lambda}}^{\mathrm{H}}-\boldsymbol{H}_2\right)\boldsymbol{W}\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}_1\}, \end{aligned} \tag{28}$$

based on which the KKT optimality conditions of **Prob.5** can be formulated as

$$\begin{cases}\mathrm{Diag}\{\boldsymbol{H}_2\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}_{\mathbf{\Lambda}}^{\mathrm{H}}-\boldsymbol{H}_2\right)\boldsymbol{W}\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}_1\}=-\mu\mathrm{Diag}\{\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}\}, & \text{(29a)} \\ \mu\left(\mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}})-K\right)=0, & \text{(29b)}\end{cases}$$

where μ is the dual variable associated with the amplitude constraint. Based on (29a), the optimal $\mathbf{\Lambda}_{\mathbf{\Theta}}$ is derived as

$$\mathrm{Diag}\{\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}\}=(\mu\boldsymbol{I}_K+\mathbf{\Phi})^{-1}\boldsymbol{a}, \tag{30}$$

where $\mathbf{\Phi}=\boldsymbol{H}_2\boldsymbol{H}_2^{\mathrm{H}}\odot\left(\boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{W}\boldsymbol{G}\boldsymbol{H}_1\right)^{\mathrm{T}}$ and $\boldsymbol{a}=\mathrm{Diag}\{\boldsymbol{H}_2\boldsymbol{W}\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}_1-\boldsymbol{H}_2\boldsymbol{H}_0^{\mathrm{H}}\boldsymbol{G}_{\mathbf{\Lambda}}^{\mathrm{H}}\boldsymbol{W}\boldsymbol{G}_{\mathbf{\Lambda}}\boldsymbol{H}_1\}$. Moreover, since $\mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}})$ is monotonically decreasing w.r.t. μ, the optimal μ satisfying (29b) is found via the bisection search.

In addition, we formulate the MSE minimization problem for the amplitude-adjustable IRS-aided MIMO system as

$$\begin{aligned}\textbf{Prob.6:}\ \min_{\mathbf{\Lambda}_{\mathbf{\Theta}}}\ & \mathrm{Tr}\left(\left[\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right]^{-1}\right) \\ \text{s.t. } & \mathrm{Tr}(\mathbf{\Lambda}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{\Theta}}^{\mathrm{H}}) \leq K. \end{aligned} \tag{31}$$

Similar to **Prob.3**, **Prob.6** is difficult to solve since it involves a quadratic term w.r.t. $\mathbf{\Lambda}_{\mathbf{\Theta}}$ appears in an inverse form. Fortunately, it can also be equivalently transformed into the WMMSE minimization problem **Prob.4** by setting $\boldsymbol{W}=\boldsymbol{I}_{N_t}$.

C. Extension to Block-Diagonal Structure Constraints

In this subsection, we extend the complex matrix derivative to the block-diagonal matrix, which is essentially a kind of bidiagonal matrix. In a general multi-antenna MU-MIMO uplink system, the received signal from K users at the BS can be written as [30]

$$\boldsymbol{y}=\boldsymbol{H}\boldsymbol{s}+\boldsymbol{n}\ \text{ with }\ \boldsymbol{H}=[\boldsymbol{H}_1,\cdots,\boldsymbol{H}_K], \tag{32}$$

where $\boldsymbol{H}_k\in\mathbb{C}^{N_t\times N_r}$ denotes the channel between the BS and the kth user, $\boldsymbol{s}_k\in\mathbb{C}^{N_r\times 1}$ denotes the transmitted data stream of the kth user, and all data streams $\boldsymbol{s}_k$'s are stacked into the vector $\boldsymbol{s}\in\mathbb{C}^{N_rK\times 1}$, i.e., $\boldsymbol{s}=[\boldsymbol{s}_1^{\mathrm{T}},\cdots,\boldsymbol{s}_K^{\mathrm{T}}]^{\mathrm{T}}$. Accordingly, the covariance matrix of $\boldsymbol{s}$ is a block-diagonal matrix, i.e.,

$$\boldsymbol{Q}=\mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}\}=\mathrm{Blockdiag}\left(\{\boldsymbol{Q}_k\}_{k=1}^{K}\right), \tag{33}$$

where $\boldsymbol{Q}_k\!=\mathbb{E}\{\boldsymbol{s}_k\boldsymbol{s}_k^{\mathrm{H}}\}$ is the transmit covariance matrix of $\boldsymbol{s}_k$. Hereafter, we mainly consider the capacity maximization problem under the general user grouping power constraints, which is formulated as

$$\begin{aligned}\textbf{Prob.7:}\ \max_{\boldsymbol{Q}}\ & \log\left|\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{H}}\right| \\ \text{s.t. } & \sum_{k\in\phi_n}\mathrm{Tr}(\boldsymbol{Q}_k)\leq P_n, \boldsymbol{Q}_k\succeq\boldsymbol{0},\ \forall k,n, \end{aligned} \tag{34}$$

where ϕ_n is the set of user indices in the n-th user group and P_n is the corresponding available transmit power. Generally, we have $\bigcup_{n=1}^{N}\phi_n=\{1,2,\cdots,K\}$ and $\phi_n\cap\phi_m=\varnothing,\forall m\neq n$. Specifically, we set $n=1,\cdots,K$, $\phi_n=\{k\}$ in the MU-MIMO uplink system, and $n=1$, $\phi_n=\{1,2,\cdots,K\}$ in the virtual MU-MIMO uplink system based on the uplink-downlink duality [17]. The differential of the objective function of **Prob.7** w.r.t. $\boldsymbol{Q}$ is given by $\mathrm{d}\left(\log\left|\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{H}}\right|\right)=\mathrm{Tr}\big(\big(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\big)^{-1}\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\mathrm{d}\big(\boldsymbol{Q}\big)\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\big)$. The corresponding first-order derivative w.r.t $\boldsymbol{Q}$ is then derived as

$$\begin{aligned}&\frac{\partial\log\left|\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{H}}\right|}{\partial\boldsymbol{Q}} \\ =&\mathrm{Blockdiag}\Big(\Big\{\boldsymbol{H}_k^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\big(\boldsymbol{I}_{N_t}+\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{H}}\mathbf{\Sigma}^{-\frac{1}{2}}\big)^{-1} \\ &\times\mathbf{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_k\Big\}_{k=1}^{K}\Big). \end{aligned} \tag{35}$$

TABLE II
ELEMENT-WISE PHASE DERIVATIVES UNDER THE CONSTANT MODULUS CONSTRAINT

Function Type	**Element-Wise Phase Derivative w.r.t.** $[\boldsymbol{\Theta}]_{i,j}, \forall i, j$
$f_{\mathrm{C,TL}} = \mathrm{Tr}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{X}) + \mathrm{Tr}(\boldsymbol{B}\boldsymbol{X}^{\mathrm{H}})$	$-2\Im\{[\boldsymbol{B}^*]_{i,j}[\boldsymbol{X}]_{i,j}\}$
$f_{\mathrm{C,TQ}} = \mathrm{Tr}(\boldsymbol{X}\boldsymbol{\Pi}\boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Phi})$	$-2\Im\{[(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Pi})^*]_{i,j}[\boldsymbol{X}]_{i,j}\}$
$f_{\mathrm{C,TI}} = \mathrm{Tr}\left((\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right)$	$2\Im\left\{\left[\left(\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-2}\right)^*\right]_{i,j}[\boldsymbol{X}]_{i,j}\right\}$
$f_{\mathrm{C,LD}} = \log\lvert\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X}\rvert$	$-2\Im\left\{\left[\left(\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right)^*\right]_{i,j}[\boldsymbol{X}]_{i,j}\right\}$

Following that, the KKT optimality conditions are given by

$$\begin{cases} \mathrm{Blockdiag}\Big(\Big\{\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\big(\boldsymbol{I}_{N_t}+\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\big)^{-1} \\ \quad\times\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_k\Big\}_{k=1}^{K}\Big)=\mathrm{Blockdiag}\Big(\{\mu_k\boldsymbol{I}_{N_r}-\boldsymbol{\Psi}_k\}_{k=1}^{K}\Big), & \text{(36a)} \\ \mu_n\Big(\sum_{k\in\phi_n}\mathrm{Tr}(\boldsymbol{Q}_k)-P_n\Big)=0, \mu_k=\mu_n, \forall k\in\phi_n, & \text{(36b)} \\ \mathrm{Tr}(\boldsymbol{\Psi}_k\boldsymbol{Q}_k)=0,\ \forall k, & \text{(36c)} \end{cases}$$

where μ_k and $\boldsymbol{\Psi}_k$ are the Lagrange multipliers associated with the transmit power constraint and the positive semi-definite constraint at the kth user, respectively. Then, (36a) can be rewritten in terms of $\boldsymbol{Q}_k$ as follows:

$$\begin{aligned} &\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\big(\boldsymbol{I}_{N_t}+\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\big)^{-1}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_k \\ =&\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{L}_k^{-\frac{1}{2}}\big(\boldsymbol{I}_{N_t}+\boldsymbol{L}_k^{-\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_k\boldsymbol{Q}_k\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{L}_k^{-\frac{1}{2}}\big)^{-1} \\ &\times\boldsymbol{L}_k^{-\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_k=\mu_k\boldsymbol{I}_{N_r}-\boldsymbol{\Psi}_k,\ \forall k, \end{aligned} \tag{37}$$

where $\boldsymbol{L}_k = \boldsymbol{I}_{N_t} + \sum_{j\neq k}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_j\boldsymbol{Q}_j\boldsymbol{H}_j^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}$. Thus, based on eigenspace alignment, the optimal $\boldsymbol{Q}_k$'s can be derived as that in [16, Theorem 1], i.e., $\boldsymbol{Q}_k = \boldsymbol{V}_{\boldsymbol{\mathcal{H}}_k}\boldsymbol{\Lambda}_{\boldsymbol{Q}_k}\boldsymbol{V}_{\boldsymbol{\mathcal{H}}_k}^{\mathrm{H}}$, $\forall k$, where $\boldsymbol{\Lambda}_{\boldsymbol{Q}_k}$ is a diagonal matrix, each diagonal element of which has a water-filling form, and $\boldsymbol{V}_{\boldsymbol{\mathcal{H}}_k}$ is an unitary matrix coming from the singular value decomposition (SVD) represented as $\boldsymbol{L}_k^{-\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_k = \boldsymbol{U}_{\boldsymbol{\mathcal{H}}_k}\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{H}}_k}\boldsymbol{V}_{\boldsymbol{\mathcal{H}}_k}^{\mathrm{H}}$ with $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{H}}_k}\searrow$, where $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{H}}_k}\searrow$ implies that the diagonal elements of $\boldsymbol{\Lambda}_{\boldsymbol{\mathcal{H}}_k}$ are arranged in descending order. Similarly, μ_k that satisfying (36b) can be obtained by the bisection search.

Remark 1: Based on the above discussions, we can conclude that globally optimal solutions of several classical optimization problems in the state-of-the-art wireless systems can be directly obtained with low complexity by the proposed complex matrix derivatives under diagonal structure constraints. In addition, for optimization problems that not satisfy diagonal structure constraints directly, the proposed algorithm is also able to obtain an approximate solution by further exploring the inherent structure of the optimal solution. For example, the optimal matrix variables in the point-to-point MIMO system operating at high SNR conditions and the MU-MISO downlink system employing the BD-ZF strategy [31] are both validated to be approximately diagonal.

III. CONSTANT MODULUS CONSTRAINTS

Different from diagonal structure constraints, constant modulus constraints are imposed on matrix variables in an element-wise manner, which makes the optimization problem challenging to directly solve using complex matrix derivatives. Motivated by this fact, we firstly provide some mathematical preliminaries for the element-wise phase derivatives of several widely adopted objective functions. Then, we investigate specific optimization problems in both the hybrid analog-digital MIMO system and the fully-passive IRS-aided MIMO system. In order to avoid complicated matrix inversion and matrix factorization, a novel AO algorithm with the aid of several arbitrary feasible solutions is proposed.

A. Mathematical Preliminaries

We firstly introduce a complex matrix variable $\boldsymbol{X} \in \mathbb{C}^{N\times M}$ subject to constant modulus constraints as follows:

$$[\boldsymbol{X}]_{i,j} = e^{\jmath\theta_{i,j}},\ \forall i,j \text{ and } \mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}) = NM, \tag{38}$$

where $\theta_{i,j} \in [0, 2\pi]$ denotes the phase of $[\boldsymbol{X}]_{i,j}$. Based on (38), the first-order derivative w.r.t. the constant modulus constrained $\boldsymbol{X}$ can be replaced by the first-order derivative w.r.t. the corresponding unconstrained phase matrix $\boldsymbol{\Theta}$, where $[\boldsymbol{\Theta}]_{i,j} = \theta_{i,j}, \forall i,j$. Accordingly, the element-wise phase derivatives of the function $f(\boldsymbol{X})$ w.r.t. $\boldsymbol{\Theta}$ can be defined as

$$\left[\frac{\partial f(\boldsymbol{X})}{\partial\boldsymbol{\Theta}}\right]_{i,j} = \frac{\partial f(\boldsymbol{X})}{\partial[\boldsymbol{\Theta}]_{i,j}},\ \forall i,j. \tag{39}$$

Similar to Sec. II-A, we also consider the element-wise phase derivatives for four common objective functions, i.e.,

1) Trace-Linear Function: Since the phase $[\boldsymbol{\Theta}]_{i,j}$'s are real scalar, for arbitrary complex matrix $\boldsymbol{B} \in \mathbb{C}^{N\times M}$, the element-wise phase derivatives of a trace-linear function $f_{\mathrm{C,TL}} = \mathrm{Tr}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{X}) + \mathrm{Tr}(\boldsymbol{B}\boldsymbol{X}^{\mathrm{H}})$ w.r.t. $[\boldsymbol{\Theta}]_{i,j}$'s can be obtained as

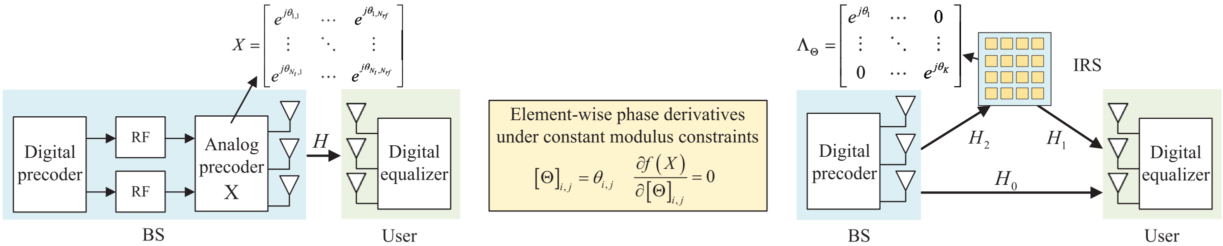
$$\begin{aligned} \left[\frac{\partial f_{\mathrm{C,TL}}}{\partial\boldsymbol{\Theta}}\right]_{i,j} =& \left[\frac{\partial\sum_m\sum_n[\boldsymbol{B}^*]_{n,m}[\boldsymbol{X}]_{n,m}}{\partial\boldsymbol{\Theta}}\right]_{i,j} \\ &+\left[\partial\frac{\sum_m\sum_n[\boldsymbol{B}]_{n,m}[\boldsymbol{X}^*]_{n,m}}{\partial\boldsymbol{\Theta}}\right]_{i,j} \\ =&\jmath[\boldsymbol{B}^*]_{i,j}[\boldsymbol{X}]_{i,j} - \jmath[\boldsymbol{B}]_{i,j}[\boldsymbol{X}^*]_{i,j} \\ =&-2\Im\{[\boldsymbol{B}^*]_{i,j}[\boldsymbol{X}]_{i,j}\},\ \forall i,j. \end{aligned} \tag{40}$$

2) Trace-Quadratic Function: For arbitrary Hermitian matrices $\boldsymbol{\Pi} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{\Phi} \in \mathbb{C}^{N\times N}$, the element-wise phase derivatives of a trace-quadratic function $f_{\mathrm{C,TQ}} = \mathrm{Tr}(\boldsymbol{X}\boldsymbol{\Pi}\boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Phi})$ w.r.t. $[\boldsymbol{\Theta}]_{i,j}$'s are given by

$$\begin{aligned} \left[\frac{\partial f_{\mathrm{C,TQ}}}{\partial\boldsymbol{\Theta}}\right]_{i,j} &= \jmath[(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Pi})^*]_{i,j}[\boldsymbol{X}]_{i,j} - \jmath[(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Pi})]_{i,j}[\boldsymbol{X}^*]_{i,j} \\ &= -2\Im\{[(\boldsymbol{\Phi}\boldsymbol{X}\boldsymbol{\Pi})^*]_{i,j}[\boldsymbol{X}]_{i,j}\},\ \forall i,j. \end{aligned} \tag{41}$$

3) Trace-Inverse Function: Regarding a trace-inverse function $f_{\mathrm{C,TI}} = \mathrm{Tr}\left((\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right)$, we have the following element-wise phase derivatives w.r.t. $[\boldsymbol{\Theta}]_{i,j}$'s.

$$\begin{aligned} \frac{\partial f_{\mathrm{C,TI}}}{\partial\boldsymbol{\Theta}} =& -\jmath\left[\left(\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X}\right)^{-2}\boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\right]_{i,j}[\boldsymbol{X}]_{i,j} \\ &+\jmath\left[\left(\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-2}\right)\right]_{i,j}[\boldsymbol{X}^*]_{i,j} \qquad (42) \\ =&2\Im\left\{\left[\left(\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-2}\right)^*\right]_{i,j}[\boldsymbol{X}]_{i,j}\right\},\ \forall i,j. \end{aligned}$$


$X = \begin{bmatrix} e^{j\theta_{1,1}} & \cdots & e^{j\theta_{1,N_{rf}}} \\ \vdots & \ddots & \vdots \\ e^{j\theta_{N_t,1}} & \cdots & e^{j\theta_{N_t,N_{rf}}} \end{bmatrix}$
Digital precoder
RF
RF
Analog precoder X
H
Digital equalizer
BS
User
Element-wise phase derivatives under constant modulus constraints
$[\Theta]_{i,j} = \theta_{i,j}$ $\frac{\partial f(X)}{\partial [\Theta]_{i,j}} = 0$
$\Lambda_\Theta = \begin{bmatrix} e^{j\theta_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{j\theta_K} \end{bmatrix}$
IRS
Digital precoder
H_2
H_1
H_0
Digital equalizer
BS
User

(a) Hybrid analog-digital MIMO system

(b) Fully-passive IRS-aided MIMO system

Fig. 2. A diagram of application scenarios associated with constant modulus matrix variables.

4) Log-Determinant Function: Similarly, the element-wise phase derivatives of a log-determinant function $f_{\mathrm{C,LD}} = \log|\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X}|$ w.r.t. $[\boldsymbol{\Theta}]_{i,j}$'s are given by

$$\left[\frac{\partial f_{\mathrm{C,LD}}}{\partial \boldsymbol{\Theta}}\right]_{i,j} = -2\Im\Bigg\{\left[\left(\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{\Phi} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right)^*\right]_{i,j} \times [\boldsymbol{X}]_{i,j}\Bigg\},\ \forall i,j. \tag{43}$$

The element-wise phase derivatives of the above four types of objective functions are summarized in Table II. In the following subsection, several state-of-the-art wireless applications will be investigated in detail based on the above fundamental properties.

B. Specific Wireless Applications

In Fig. 2, there are two typical wireless applications associated with constant modulus constraints, i.e., the analog beamforming optimization in the hybrid analog-digital MIMO system and the phase shift optimization in the fully-passive IRS-aided MIMO system, which are elaborated as follows.

1) Hybrid Analog-Digital MIMO System: As shown in Fig. 2(a), we firstly consider the hybrid analog-digital beamforming design in the downlink point-to-point MIMO system, where the BS equipped with N_t antennas and N_{rf} radio-frequency (RF) chains transfers N_s data streams to a N_r-antenna user with the fully-digital equalizer [12]. Then, the received signal at the user can be expressed as

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{H}\boldsymbol{X}\boldsymbol{F}_{\mathrm{D}}\boldsymbol{s} + \boldsymbol{G}\boldsymbol{n}, \tag{44}$$

where $\boldsymbol{G} \in \mathbb{C}^{N_s \times N_r}$ denotes the fully-digital receive equalizer, $\boldsymbol{H} \in \mathbb{C}^{N_r \times N_t}$ denotes the channel between the BS and the user, $\boldsymbol{X} \in \mathbb{C}^{N_t \times N_{rf}}$ and $\boldsymbol{F}_{\mathrm{D}} \in \mathbb{C}^{N_{rf} \times N_s}$ are the constant modulus analog beamformer and the digital beamformer, respectively. $\boldsymbol{s} \in \mathbb{C}^{N_s \times 1}$ is the transmit data streams with unit covariance matrix, i.e., $\mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}\} = \boldsymbol{I}_{N_s}$. $\boldsymbol{n} \in \mathbb{C}^{N_r \times 1}$ is the additive Gaussian noise with zero mean and covariance matrix $\mathbb{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}\} = \boldsymbol{\Sigma}$. Based on (44), the MSE matrix is given by

$$\begin{aligned} \boldsymbol{E}_{\mathrm{MSE}} &= \mathbb{E}\left[(\hat{\boldsymbol{s}} - \boldsymbol{s})(\hat{\boldsymbol{s}} - \boldsymbol{s})^{\mathrm{H}}\right] \\ &= (\boldsymbol{G}\boldsymbol{H}\boldsymbol{X}\boldsymbol{F}_{\mathrm{D}} - \boldsymbol{I}_{N_s})(\boldsymbol{G}\boldsymbol{H}\boldsymbol{X}\boldsymbol{F}_{\mathrm{D}} - \boldsymbol{I}_{N_s})^{\mathrm{H}} + \boldsymbol{G}\boldsymbol{\Sigma}\boldsymbol{G}^{\mathrm{H}} \\ &\overset{(c)}{=} \left[\boldsymbol{I}_{N_s} + \boldsymbol{F}_{\mathrm{D}}^{\mathrm{H}}\boldsymbol{X}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{X}\boldsymbol{F}_{\mathrm{D}}\right]^{-1}, \end{aligned} \tag{45}$$

where $\hat{\boldsymbol{s}}$ denotes the estimated signal and (c) holds based on the optimal unconstrained Wiener filter $\boldsymbol{G} = \boldsymbol{F}_{\mathrm{D}}^{\mathrm{H}}\boldsymbol{X}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{X}\boldsymbol{F}_{\mathrm{D}}\boldsymbol{F}_{\mathrm{D}}^{\mathrm{H}}\boldsymbol{X}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_{\mathrm{n}}\right)^{-1}$ [32]. Without loss of generality, we usually assume $\boldsymbol{F}_{\mathrm{D}}\boldsymbol{F}_{\mathrm{D}}^{\mathrm{H}} \approx \gamma^2\boldsymbol{I}_{N_{rf}}$ for large-scale MIMO systems [20]. Under this assumption, the capacity maximization problem of the hybrid analog-digital MIMO system can be formulated as

$$\begin{aligned} \textbf{Prob.8:}\ &\max_{\boldsymbol{X}}\ \log|\boldsymbol{I}_{N_{rf}} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X}| \\ &\ \mathrm{s.t.}\ |[\boldsymbol{X}]_{i,j}| = 1,\ \forall i,j, \end{aligned} \tag{46}$$

where $\boldsymbol{\Pi} = \gamma^2\boldsymbol{H}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{H}$ represents the effective signal-to-noise ratio (SNR). According to the KKT optimality conditions, the element-wise phase derivatives of the objective function of **Prob.8** w.r.t. $[\boldsymbol{X}]_{i,j}$'s must equal zeros at the optimal $[\boldsymbol{X}]_{i,j}$'s. Specifically, we recall (43) to obtain

$$\underbrace{\Im\left\{\left[\left(\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{I}_{N_{rf}} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right)^*\right]_{i,j}[\boldsymbol{X}]_{i,j}\right\}}_{\triangleq g_{i,j}(\boldsymbol{X})} = 0,\ \forall i,j. \tag{47}$$

Define $\boldsymbol{A}_j = \boldsymbol{I}_{N_t} + \widetilde{\boldsymbol{X}}_j\widetilde{\boldsymbol{X}}_j^{\mathrm{H}}\boldsymbol{\Pi}$ and $n_j = 1 + [\boldsymbol{X}]_{:,j}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}[\boldsymbol{X}]_{:,j}$, where $\widetilde{\boldsymbol{X}}_j$ denotes the sub-matrix of $\boldsymbol{X}$ with the jth column removed. Then, $\left[\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{I}_{N_{rf}} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right]_{i,j}$'s can be rewritten as

$$\begin{aligned} &\left[\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{I}_{N_{rf}} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right]_{i,j} \\ &\overset{(d)}{=} \left[\boldsymbol{\Pi}\left(\boldsymbol{A}_j^{-1} - \frac{\boldsymbol{A}_j^{-1}[\boldsymbol{X}]_{:,j}[\boldsymbol{X}]_{:,j}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}}{n_j}\right)\boldsymbol{X}\right]_{i,j} \\ &= \frac{1}{n_j}\sum\nolimits_{l \neq i}[\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}]_{l,i}^{\mathrm{H}}[\boldsymbol{X}]_{l,j} + \frac{1}{n_j}[\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}]_{i,i}^{\mathrm{H}}[\boldsymbol{X}]_{i,j},\ \forall i,j, \end{aligned} \tag{48}$$

where (d) holds similarly to (14). By substituting (48) into (47), we have

$$\Im\left\{\left[\left(\boldsymbol{\Pi}\boldsymbol{X}(\boldsymbol{I}_{N_{rf}} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X})^{-1}\right)^*\right]_{i,j}[\boldsymbol{X}]_{i,j}\right\} \tag{49a}$$

$$\begin{aligned} =&\Im\Bigg\{\left(\frac{1}{n_j}\sum\nolimits_{l \neq i}[\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}]_{l,i}^{\mathrm{H}}[\boldsymbol{X}]_{l,j}\right)^*[\boldsymbol{X}]_{i,j} \\ &+ \left(\frac{1}{n_j}[\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}]_{i,i}^{\mathrm{H}}\right)^*\Bigg\} = 0,\ \forall i,j. \end{aligned} \tag{49b}$$

Since the following equality holds, i.e.,

$$\begin{aligned} \boldsymbol{\Pi}\boldsymbol{A}_j^{-1} =& \boldsymbol{\Pi}(\boldsymbol{I}_{N_t} + \widetilde{\boldsymbol{X}}_j\widetilde{\boldsymbol{X}}_j^{\mathrm{H}}\boldsymbol{\Pi})^{-1} \\ =& (\boldsymbol{I}_{N_t} + \boldsymbol{\Pi}\widetilde{\boldsymbol{X}}_j\widetilde{\boldsymbol{X}}_j^{\mathrm{H}})^{-1}\boldsymbol{\Pi} = (\boldsymbol{\Pi}\boldsymbol{A}_j^{-1})^{\mathrm{H}},\ \forall j, \end{aligned} \tag{50}$$

we can conclude that $\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}$ is a Hermitian matrix. As such, it is readily inferred that n_j and $\frac{1}{n_j}[\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}]_{i,i}^{\mathrm{H}}$ are both real scalars. Recall the definition of $[\boldsymbol{X}]_{i,j} = e^{j\theta_{i,j}}, \forall i,j$, it follows from (49b) that the optimal $[\boldsymbol{\Theta}]_{i,j}$'s to **Prob.8** are derived as

$$\begin{aligned} [\boldsymbol{\Theta}]_{i,j} =& \mathrm{Phase}\left\{\sum\nolimits_{l \neq i}[\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}]_{l,i}^{\mathrm{H}}[\boldsymbol{X}]_{l,j}\right\} \\ &\text{or } \pi + \mathrm{Phase}\left\{\sum\nolimits_{l \neq i}[\boldsymbol{\Pi}\boldsymbol{A}_j^{-1}]_{l,i}^{\mathrm{H}}[\boldsymbol{X}]_{l,j}\right\},\ \forall i,j. \end{aligned} \tag{51}$$

In addition, the MSE minimization problem is expressed as

$$\begin{aligned} \textbf{Prob.9:}\ &\min_{\boldsymbol{X}}\ \mathrm{Tr}\left(\left(\boldsymbol{I}_{N_{rf}} + \boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi}\boldsymbol{X}\right)^{-1}\right) \\ &\ \mathrm{s.t.}\ |[\boldsymbol{X}]_{i,j}| = 1,\ \forall i,j. \end{aligned} \tag{52}$$

Similarly, to find the optimal solution of **Prob.9**, the element-wise phase derivatives of the objective function w.r.t. $[\boldsymbol{X}]_{i,j}$'s must equal zeros, that is,

$$\underbrace{\Im\left\{\left[\left(\boldsymbol{\Pi X}\left(\boldsymbol{I}_{N_{rf}}+\boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi X}\right)^{-2}\right)^{*}\right]_{i,j}[\boldsymbol{X}]_{i,j}\right\}}_{\triangleq g_{i,j}(\boldsymbol{X})}=0,\ \forall i,j. \quad (53)$$

The terms $\left[\boldsymbol{\Pi X}\left(\boldsymbol{I}_{N_{rf}}+\boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi X}\right)^{-2}\right]_{i,j}$'s can be rewritten as

$$\begin{aligned}&\left[\boldsymbol{\Pi X}(\boldsymbol{I}_{N_{rf}}+\boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi X})^{-2}\right]_{i,j}\\ \overset{(e_1)}{=}&\left[\boldsymbol{\Pi}\left(\boldsymbol{A}_j^{-1}-\frac{\boldsymbol{A}_j^{-1}[\boldsymbol{X}]_{:,j}[\boldsymbol{X}]_{:,j}^{\mathrm{H}}\boldsymbol{\Pi A}_j^{-1}}{n_j}\right)^2\boldsymbol{X}\right]_{i,j}\\ \overset{(e_2)}{=}&\frac{1}{n_j}[\boldsymbol{\Pi A}_j^{-2}]_{:,i}^{\mathrm{H}}[\boldsymbol{X}]_{:,j}-\frac{m_j}{n_j^2}[\boldsymbol{\Pi A}_j^{-1}]_{:,i}^{\mathrm{H}}[\boldsymbol{X}]_{:,j}\\ \overset{(e_3)}{=}&\frac{A_{1,i,j}^{*}+B_{1,i,j}^{*}[\boldsymbol{X}]_{i,j}^{2}+C_{1,i,j}^{*}[\boldsymbol{X}]_{i,j}}{n_j^2},\ \forall i,j, \qquad (54)\end{aligned}$$

where (e_1) holds due to the same reasons as (14), (e_2) holds by defining $m_j=[\boldsymbol{X}]_{:,j}^{\mathrm{H}}\boldsymbol{\Pi A}_j^{-2}[\boldsymbol{X}]_{:,j}$, which is a real scalar and this can be proved similarly to (50). (e_3) is obtained by rewriting n_j and m_j in terms of $[\boldsymbol{X}]_{i,j}$ as

$$n_j=\zeta_j^n+2\Re\left\{\eta_j^n[\boldsymbol{X}]_{i,j}^{*}\right\},m_j=\zeta_j^m+2\Re\left\{\eta_j^m[\boldsymbol{X}]_{i,j}^{*}\right\}, \quad (55)$$

where $\zeta_j^n=1+[\boldsymbol{\Pi A}_j^{-1}]_{i,i}+\Re\left\{\sum_{p\neq i,q\neq i}[\boldsymbol{X}]_{p,j}^{*}[\boldsymbol{\Pi A}_j^{-1}]_{p,q}[\boldsymbol{X}]_{q,j}\right\}$, $\zeta_j^m=[\boldsymbol{\Pi A}_j^{-2}]_{i,i}+\Re\left\{\sum_{p\neq i,q\neq i}[\boldsymbol{X}]_{p,j}^{*}[\boldsymbol{\Pi A}_j^{-2}]_{p,q}[\boldsymbol{X}]_{q,j}\right\}$, $\eta_j^n=\sum_{l\neq i}[\boldsymbol{\Pi A}_j^{-1}]_{l,i}[\boldsymbol{X}]_{l,j}^{*}$, $\eta_j^m=\sum_{l\neq i}[\boldsymbol{\Pi A}_j^{-2}]_{l,i}[\boldsymbol{X}]_{l,j}^{*}$, $\forall i,j$. Moreover, $A_{1,i,j}$, $B_{1,i,j}$ and $C_{1,i,j}$ in (54) are defined as

$$\begin{aligned}A_{1,i,j}=&(\eta_j^n)^{*}[\boldsymbol{\Pi A}_j^{-2}]_{i,i}^{*}-(\eta_j^m)^{*}[\boldsymbol{\Pi A}_j^{-1}]_{i,i}^{*}\\&+(\zeta_j^n)^{*}(\eta_j^m)^{*}-(\zeta_j^m)^{*}(\eta_j^n)^{*},\\ B_{1,i,j}=&\eta_j^n[\boldsymbol{\Pi A}_j^{-2}]_{i,i}^{*}-\eta_j^m[\boldsymbol{\Pi A}_j^{-1}]_{i,i}^{*},\\ C_{1,i,j}=&(\zeta_j^n)^{*}[\boldsymbol{\Pi A}_j^{-2}]_{i,i}^{*}-(\zeta_j^m)^{*}[\boldsymbol{\Pi A}_j^{-1}]_{i,i}^{*}\\&+\eta_j^n(\eta_j^m)^{*}-\eta_j^m(\eta_j^n)^{*},\ \forall i,j. \qquad (56)\end{aligned}$$

By substituting (54) into (53), we have

$$\begin{aligned}&\Im\left\{\left[\left(\boldsymbol{\Pi X}(\boldsymbol{I}_{N_{rf}}+\boldsymbol{X}^{\mathrm{H}}\boldsymbol{\Pi X})^{-2}\right)^{*}\right]_{i,j}[\boldsymbol{X}]_{i,j}\right\} \qquad (57a)\\ =&\Im\left\{\frac{A_{1,i,j}[\boldsymbol{X}]_{i,j}+B_{1,i,j}[\boldsymbol{X}]_{i,j}^{*}+C_{1,i,j}}{n_j^2}\right\}=0,\ \forall i,j. \quad (57b)\end{aligned}$$

Then, since $\Im\{\frac{a}{b}\}=0$ is equal to $\Im\{a\}=0$ for a real scalar b, (57b) can be further simplified as

$$\begin{aligned}&\Im\left\{A_{1,i,j}e^{\jmath\theta_{i,j}}+B_{1,i,j}e^{-\jmath\theta_{i,j}}+C_{1,i,j}\right\} \qquad (58a)\\ \overset{(f_1)}{=}&|A_{1,i,j}|\sin(\theta_{i,j}+\alpha_{i,j})-|B_{1,i,j}|\sin(\theta_{i,j}-\beta_{i,j})+\Im\{C_{1,i,j}\}\\ \overset{(f_2)}{=}&\sqrt{z_{1,i,j}^2+z_{2,i,j}^2}\sin(\theta_{i,j}+\phi_{i,j})+\Im\{C_{1,i,j}\}=0,\ \forall i,j, \quad (58b)\end{aligned}$$

where (f_1) is obtained using Euler's formula with $A_{1,i,j}=|A_{1,i,j}|e^{\jmath\alpha_{i,j}}$ and $B_{1,i,j}=|B_{1,i,j}|e^{\jmath\beta_{i,j}}$, (f_2) holds due to the sum-to-product trigonometric identity with $z_{1,i,j}=|A_{1,i,j}|\cos(\alpha_{i,j})+|B_{1,i,j}|\cos(\beta_{i,j})$, $z_{2,i,j}=|A_{1,i,j}|\sin(\alpha_{i,j})+|B_{1,i,j}|\sin(\beta_{i,j})$ and

$$\phi_{i,j}=\begin{cases}\arctan\left(\frac{z_{2,i,j}}{z_{1,i,j}}\right), & \text{if } z_{1,i,j}\geq 0,\\ \pi-\arctan\left(\frac{z_{2,i,j}}{z_{1,i,j}}\right), & \text{if } z_{1,i,j}<0,\end{cases}\ \forall i,j. \quad (59)$$

Based on (58b), the optimal $[\boldsymbol{\Theta}]_{i,j}$'s to **Prob.9** are obtained as

$$\begin{aligned}[\boldsymbol{\Theta}]_{i,j}=&-\phi_{i,j}+\arcsin\left(-\frac{\Im\{C_{1,i,j}\}}{\sqrt{z_{1,i,j}^2+z_{2,i,j}^2}}\right) \qquad (60)\\ &\text{or } \pi-\phi_{i,j}+\arcsin\left(-\frac{\Im\{C_{1,i,j}\}}{\sqrt{z_{1,i,j}^2+z_{2,i,j}^2}}\right),\ \forall i,j.\end{aligned}$$

Remark 2: Since the capacity maximization problem w.r.t. each user's analog beamformer $\boldsymbol{X}_u$ in the uplink hybrid analog-digital MU-MIMO system can be formulated in a similar form to **Prob.8** that for the single-user case by separating $\boldsymbol{X}_u$ from other $\boldsymbol{X}_i,\forall i\neq u$ [33], its optimal solution can still be obtained according to (51) by modifying $\boldsymbol{\Pi}$ as $\boldsymbol{H}_u^H\left(\frac{1}{\gamma^2}\boldsymbol{\Sigma}+\sum_{l\neq u}\boldsymbol{H}_l\boldsymbol{X}_l\boldsymbol{X}_l^{\mathrm{H}}\boldsymbol{H}_l^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_u$. Moreover, for the MSE minimization problem in the uplink hybrid analog-digital MU-MIMO system, the corresponding element-wise phase derivative also has a similar form to (53) for the single-user counterpart, thereby leading to the optimal solution obtained by (60).

Next, we consider a general WMMSE minimization problem often studied in the downlink hybrid analog-digital MU-MIMO system for both capacity maximization and MSE minimization problems, which can be formulated as [21]

$$\begin{aligned}\textbf{Prob.10:}\ \min_{\{\boldsymbol{X}_u\}}\ &\sum_{u=1}^{U}\left(\mathrm{Tr}(\boldsymbol{\Phi}_u\boldsymbol{X}_u\boldsymbol{\Pi}_u\boldsymbol{X}_u^{\mathrm{H}})-\mathrm{Tr}(\boldsymbol{B}_u^{\mathrm{H}}\boldsymbol{X}_u)\right.\\ &\left.-\mathrm{Tr}(\boldsymbol{B}_u\boldsymbol{X}_u^{\mathrm{H}})\right)\\ \text{s.t. }&|[\boldsymbol{X}_u]_{i,j}|=1,\ \forall u,i,j, \qquad (61)\end{aligned}$$

where $\boldsymbol{X}_u$'s are the analog beamformer for the uth user. $\boldsymbol{\Phi}_u\in\mathbb{C}^{N_t\times N_t}$'s, $\boldsymbol{\Pi}_u\in\mathbb{C}^{N_{rf}\times N_{rf}}$'s and $\boldsymbol{B}_u\in\mathbb{C}^{N_t\times N_{rf}}$'s denote the corresponding effective channel covariance matrix, the digital beamforming covariance matrix and the cascade channel, respectively, which are mathematically modeled as

$$\begin{aligned}\boldsymbol{\Phi}_u&=\begin{cases}\boldsymbol{H}_{u,u}^{\mathrm{H}}\boldsymbol{G}_{\boldsymbol{\Theta},u}^{\mathrm{H}}\boldsymbol{W}_u\boldsymbol{G}_{\boldsymbol{\Theta},u}\boldsymbol{H}_{u,u}, & \text{for capacity max problem},\\ \boldsymbol{H}_{u,u}^{\mathrm{H}}\boldsymbol{G}_{\boldsymbol{\Theta},u}^{\mathrm{H}}\boldsymbol{G}_{\boldsymbol{\Theta},u}\boldsymbol{H}_{u,u}, & \text{for MSE min problem},\end{cases}\\ \boldsymbol{\Pi}_u&=\boldsymbol{F}_{\mathrm{D,u}}\boldsymbol{F}_{\mathrm{D,u}}^{\mathrm{H}}, \qquad (62)\\ \boldsymbol{B}_u&=\begin{cases}\boldsymbol{H}_{u,u}^{\mathrm{H}}\boldsymbol{G}_{\boldsymbol{\Theta},u}^{\mathrm{H}}\boldsymbol{W}_u\boldsymbol{F}_{\mathrm{D,u}}^{\mathrm{H}}, & \text{for capacity max problem},\\ \boldsymbol{H}_{u,u}^{\mathrm{H}}\boldsymbol{G}_{\boldsymbol{\Theta},u}^{\mathrm{H}}\boldsymbol{F}_{\mathrm{D,u}}^{\mathrm{H}}, & \text{for MSE min problem},\end{cases}\ \forall u.\end{aligned}$$

Generally, the optimal $[\boldsymbol{X}_u]_{i,j}$'s can be obtained when the element-wise phase derivatives of the objective function of **Prob.10** w.r.t. $[\boldsymbol{X}_u]_{i,j}$'s equal zeros, i.e.,

$$\underbrace{\Im\left\{\left([(\boldsymbol{\Phi X}_u\boldsymbol{\Pi})^{*}]_{i,j}-[\boldsymbol{B}^{*}]_{i,j}\right)[\boldsymbol{X}_u]_{i,j}\right\}}_{\triangleq g_{i,j}(\boldsymbol{X}_u)}=0,\ \forall u,i,j, \quad (63)$$

where $[(\boldsymbol{\Phi X}_u\boldsymbol{\Pi})]_{i,j}$'s can be rewritten as

$$[(\boldsymbol{\Phi X}_u\boldsymbol{\Pi})]_{i,j}=\mathrm{Tr}\left(\boldsymbol{X}_u[\boldsymbol{\Pi}]_{:,j}[\boldsymbol{\Phi}]_{i,:}\right)$$

$$
\begin{aligned}
=&[\boldsymbol{X}_u]_{i,j}\left[([\boldsymbol{\Pi}]_{:,j}[\boldsymbol{\Phi}]_{i,:})\right]_{j,i} \qquad (64)\\
&+\underbrace{\sum\nolimits_{m\neq i}\sum\nolimits_{n\neq j}[\boldsymbol{X}_u]_{m,n}\left[([\boldsymbol{\Pi}]_{:,j}[\boldsymbol{\Phi}]_{i,:})\right]_{n,m}}_{\triangleq s_{u,i,j}},\ \forall u,i,j.
\end{aligned}
$$

By substituting (64) into (63), we have

$$
\begin{aligned}
&\Im\left\{\left([(\boldsymbol{\Phi}\boldsymbol{X}_u\boldsymbol{\Pi})^*]_{i,j}-[\boldsymbol{B}^*]_{i,j}\right)[\boldsymbol{X}_u]_{i,j}\right\} \qquad (65a)\\
=&\Im\left\{(s_{i,j}-[\boldsymbol{B}]_{i,j})^*\,[\boldsymbol{X}_u]_{i,j}+\left[([\boldsymbol{\Pi}]_{:,j}[\boldsymbol{\Phi}]_{i,:})\right]^*_{j,i}\right\}=0,\ \forall u,i,j. \qquad (65b)
\end{aligned}
$$

The last term $\left[([\boldsymbol{\Pi}]_{:,j}[\boldsymbol{\Phi}]_{i,:})\right]^*_{j,i}$ is a real scalar since it satisfies $\left[([\boldsymbol{\Pi}]_{:,j}[\boldsymbol{\Phi}]_{i,:})\right]^*_{j,i}=[\boldsymbol{\Pi}]_{j,j}[\boldsymbol{\Phi}]_{i,i}=([\boldsymbol{\Pi}]_{j,j}[\boldsymbol{\Phi}]_{i,i})^*$. Thus, the optimal $[\boldsymbol{\Theta}_u]_{i,j}$'s satisfying (65b) for **Prob.10** are given by

$$
\begin{aligned}
[\boldsymbol{\Theta}_u]_{i,j}=&\mathrm{Phase}\left\{s_{i,j}-[\boldsymbol{B}]_{i,j}\right\}\\
&\text{or } \pi+\mathrm{Phase}\left\{s_{i,j}-[\boldsymbol{B}]_{i,j}\right\},\ \forall u,i,j. \qquad (66)
\end{aligned}
$$

2) Fully-Passive IRS-aided MIMO System: In the fully-passive IRS-aided point-to-point MIMO system as shown in Fig. 2(b), the received signal at the user can be written as

$$
\boldsymbol{y}=\boldsymbol{H}\boldsymbol{s}+\boldsymbol{n}=(\boldsymbol{H}_0+\boldsymbol{H}_1\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\boldsymbol{H}_2)\,\boldsymbol{s}+\boldsymbol{n}, \qquad (67)
$$

where $\boldsymbol{H}_0\in\mathbb{C}^{N_r\times N_t}$, $\boldsymbol{H}_1\in\mathbb{C}^{N_r\times K}$ and $\boldsymbol{H}_2\in\mathbb{C}^{K\times N_t}$ represent the BS-user direct channel, the IRS-user channel and the BS-IRS channel, respectively. $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\in\mathbb{C}^{K\times K}$ is the diagonal IRS reflection matrix subject to both diagonal structure constraints and constant modulus constraints [13]. The phase shift vector $\boldsymbol{\theta}$ corresponding to the IRS reflection matrix $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}$ is then defined as

$$
[\boldsymbol{\theta}]_i=\theta_i,\ [\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}=e^{\jmath\theta_i},\ \forall i. \qquad (68)
$$

Firstly, we consider the classical capacity maximization problem in the fully-passive IRS-aided MIMO system as follows:

$$
\begin{aligned}
\textbf{Prob.11:}\ \max_{\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}}\ &\log\left|\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right|\\
\mathrm{s.t.}\ &|[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}|=1,\ \forall i. \qquad (69)
\end{aligned}
$$

By leveraging the KKT optimality conditions, the element-wise phase derivatives of the objective function of **Prob.11** w.r.t. $[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}$'s equal zeros at the optimal $[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}$'s, i.e.,

$$
\underbrace{\Im\left\{\left[\boldsymbol{H}_2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-1}\boldsymbol{H}_1\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\}}_{\triangleq g_{i,i}(\boldsymbol{\Lambda}_{\boldsymbol{\Theta}})}=0,\ \forall i. \qquad (70)
$$

The left-hand side of (70) can be further rewritten as

$$
\begin{aligned}
&\Im\left\{\left[\boldsymbol{H}_2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-1}\boldsymbol{H}_1\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\}\\
=&\Im\left\{\left[\boldsymbol{H}_2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-\frac{1}{2}}+\boldsymbol{I}_{N_r}\right)^{-1}\right.\right.\\
&\left.\left.\times\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_1\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\}\\
\overset{(g_1)}{=}&\Im\left\{\mathrm{Tr}\left[\boldsymbol{\Gamma}_i\left(\boldsymbol{M}_i^{\mathrm{H}}+e^{-\jmath\theta_i}\boldsymbol{\Gamma}_i^{\mathrm{H}}\right)\left(\boldsymbol{M}_i\boldsymbol{M}_i^{\mathrm{H}}+e^{-\jmath\theta_i}\boldsymbol{M}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}\right.\right.\right.\\
&\left.\left.\left.+e^{\jmath\theta_i}\boldsymbol{\Gamma}_i\boldsymbol{M}_i^{\mathrm{H}}+\boldsymbol{\Gamma}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-1}\right]e^{\jmath\theta_i}\right\}\\
\overset{(g_2)}{=}&\Im\left\{\mathrm{Tr}\left[e^{\jmath\theta_i}\boldsymbol{\Gamma}_i\boldsymbol{M}_i^{\mathrm{H}}\left(\boldsymbol{\Phi}_i+e^{-\jmath\theta_i}\boldsymbol{M}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}e^{\jmath\theta_i}\boldsymbol{\Gamma}_i\boldsymbol{M}_i^{\mathrm{H}}\right)^{-1}\right]+c_i\right\}\\
\overset{(g_3)}{=}&\Im\left\{\mathrm{Tr}\left[e^{\jmath\theta_i}\boldsymbol{u}_i\boldsymbol{v}_i^{\mathrm{H}}\left(\boldsymbol{\Psi}_i-\boldsymbol{a}_i\boldsymbol{a}_i^{\mathrm{H}}\right)^{-1}\right]+c_i\right\}\\
\overset{(g_4)}{=}&\Im\left\{e^{\jmath\theta_i}\boldsymbol{v}_i^{\mathrm{H}}\left(\boldsymbol{\Psi}_i^{-1}+\frac{\boldsymbol{\Psi}_i^{-1}\boldsymbol{a}_i\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}}{1-\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{a}_i}\right)\boldsymbol{u}_i+c_i\right\},\ \forall i, \qquad (71)
\end{aligned}
$$

where (g_1) holds by rewriting $\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}$ in terms of $[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}$ as

$$
\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}=[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\,\boldsymbol{\Gamma}_i+\underbrace{\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_0+\sum\nolimits_{n\neq i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{n,n}\,\boldsymbol{\Gamma}_n}_{\triangleq\boldsymbol{M}_i}, \qquad (72)
$$

where $\boldsymbol{\Gamma}_i=[\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}_1]_{:,i}[\boldsymbol{H}_2]_{i,:}$. The equality (g_2) in (71) is obtained by defining $\boldsymbol{\Phi}_i=\boldsymbol{M}_i\boldsymbol{M}_i^{\mathrm{H}}+\boldsymbol{\Gamma}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}+\boldsymbol{I}_{N_r}$, $c_i=\mathrm{Tr}\left[\boldsymbol{\Gamma}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}\left(\boldsymbol{\Phi}_i+e^{-\jmath\theta_i}\boldsymbol{M}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}+e^{\jmath\theta_i}\boldsymbol{\Gamma}_i\boldsymbol{M}_i^{\mathrm{H}}\right)^{-1}\right]$, where c_i is a real scalar independent of the optimal $[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}$. The equality (g_3) holds based on $\boldsymbol{a}_i=e^{\jmath\theta_i}\boldsymbol{v}_i-\boldsymbol{u}_i$, $\boldsymbol{\Psi}_i=\boldsymbol{\Phi}_i+\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{H}}+\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}$, where $\boldsymbol{v}_i$ and $\boldsymbol{u}_i$ come from the SVD of the rank-one matrix $\boldsymbol{M}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}$, i.e., $\boldsymbol{M}_i\boldsymbol{\Gamma}_i^{\mathrm{H}}=\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}$. The equality (g_4) holds similarly to (a_2) in (14). Then, by substituting (71) into (70), we have

$$
\begin{aligned}
&\Im\left\{\left[\boldsymbol{H}_2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-1}\boldsymbol{H}_1\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\} \qquad (73a)\\
=&\Im\left\{\frac{A_{2,i}e^{\jmath\theta_i}+C_{2,i}}{D_{2,i}}+c_i\right\}=0,\ \forall i, \qquad (73b)
\end{aligned}
$$

where

$$
\begin{aligned}
&A_{2,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i, \qquad (74)\\
&C_{2,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{v}_i-\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i,\\
&D_{2,i}=2\mathrm{Re}\left\{\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_ie^{\jmath\theta_i}\right\}+1-\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{v}_i-\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i,\ \forall i.
\end{aligned}
$$

Similar to (49b), since $C_{2,i}$'s, $D_{2,i}$'s and c_i's are real scalars, the optimal $[\boldsymbol{\theta}]_i$'s satisfying (73b) for **Prob.11** are given by

$$
[\boldsymbol{\theta}]_i=\mathrm{Phase}\left\{A^*_{2,i}\right\}\ \text{ or }\pi+\mathrm{Phase}\left\{A^*_{2,i}\right\},\ \forall i. \qquad (75)
$$

Additionally, we consider the MSE minimization problem in the fully-passive IRS-aided point-to-point MIMO system, which is formulated as

$$
\begin{aligned}
\textbf{Prob.12:}\ \min_{\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}}\ &\mathrm{Tr}\left[\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-1}\right]\\
\mathrm{s.t.}\ &|[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}|=1,\ \forall i. \qquad (76)
\end{aligned}
$$

Since the element-wise phase derivatives of the objective function of **Prob.12** w.r.t. $[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}$'s equal zeros at the optimal solution, we have

$$
\underbrace{\Im\left\{\left[\boldsymbol{H}_2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-2}\boldsymbol{H}_1\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\}}_{\triangleq g_{i,i}(\boldsymbol{\Lambda}_{\boldsymbol{\Theta}})}=0,\ \forall i. \qquad (77)
$$

The left-hand side of (77) can be further rewritten as

$$
\begin{aligned}
&\Im\left\{\left[\boldsymbol{H}_2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-2}\boldsymbol{H}_1\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\}\\
\overset{(h)}{=}&\Im\left\{e^{\jmath\theta_i}\boldsymbol{v}_i^{\mathrm{H}}\left(\boldsymbol{\Psi}_i^{-1}+\frac{\boldsymbol{\Psi}_i^{-1}\boldsymbol{a}_i\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}}{1-\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{a}_i}\right)\right.\\
&\left.\times\left(\boldsymbol{\Psi}_i^{-1}+\frac{\boldsymbol{\Psi}_i^{-1}\boldsymbol{a}_i\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}}{1-\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{a}_i}\right)\boldsymbol{u}_i+c_i\right\}\\
=&\Im\left\{\frac{e^{\jmath\theta_i}\left(d_i\eta_{1,i}+d_i\eta_{2,i}+d_i^2\varepsilon_i+b_{1,i}-r_{0,i}\lambda_i\right)}{\left(2\mathrm{Re}\left\{\lambda_ie^{\jmath\theta_i}\right\}+d\right)^2}\right.
\end{aligned}
$$

$$
+\frac{e^{-\jmath\theta_i}\left((\lambda_i^*)^2\varepsilon_i+b_{2,i}-r_{0,i}\lambda_i^*\right)}{\left(2\mathrm{Re}\left\{\lambda_i e^{\jmath\theta_i}\right\}+d\right)^2} \tag{78}
$$

$$
\left.+\frac{2d_i\lambda_i^*\varepsilon_i+\lambda_i^*\eta_{1,i}+\lambda_i^*\eta_{2,i}-b_{0,i}-r_{0,i}d_i}{\left(2\mathrm{Re}\left\{\lambda_i e^{\jmath\theta_i}\right\}+d\right)^2}+c_i\right\},\ \forall i,
$$

where (h) holds similarly to (71) and we have

$$
\begin{aligned}
&\lambda_i=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i, \varepsilon_i=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\boldsymbol{u}_i, d_i=1-\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{v}_i-\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i,\\
&\eta_{1,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\left(\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}+\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{H}}\right)\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i,\\
&\eta_{2,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\left(\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}+\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{H}}\right)\boldsymbol{\Psi}_i^{-2}\boldsymbol{u}_i,\\
&b_{0,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\left(\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}+\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{H}}\right)\boldsymbol{\Psi}_i^{-2}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i\\
&\qquad+\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\left(\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}+\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{H}}\right)\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i,\\
&b_{1,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\left(\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}+\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{H}}\right)\boldsymbol{\Psi}_i^{-2}\left(\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}+\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{H}}\right)\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i\\
&\quad+\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\left(\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{u}_i+\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{v}_i\right)\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i,\\
&b_{2,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i, \qquad (79)\\
&r_{0,i}=\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\boldsymbol{u}_i+\boldsymbol{v}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-2}\boldsymbol{v}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{u}_i,\ \ \forall i.
\end{aligned}
$$

By substituting (78) into (77), we have

$$
\Im\left\{\left[\boldsymbol{H}_2\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{I}_{N_r}\right)^{-2}\boldsymbol{H}_1\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\} \tag{80a}
$$

$$
=\Im\left\{\frac{A_{3,i}e^{\jmath\theta_i}+B_{3,i}e^{-\jmath\theta_i}+C_{3,i}}{D_{3,i}}+c_i\right\}=0,\ \forall i, \tag{80b}
$$

where

$$
\begin{aligned}
A_{3,i}=&d_i\eta_{1,i}+d_i\eta_{2,i}+d_i^2\varepsilon_i+b_{1,i}-r_{0,i}\lambda_i,\\
B_{3,i}=&(\lambda_i^*)^2\varepsilon_i+b_{2,i}-r_{0,i}\lambda_i^*,\\
C_{3,i}=&2d_i\lambda_i^*\varepsilon_i+\lambda_i^*\eta_{1,i}+\lambda_i^*\eta_{2,i}-b_{0,i}-r_{0,i}d_i,\\
D_{3,i}=&\left(2\mathrm{Re}\left\{\lambda_i e^{\jmath\theta_i}\right\}+d_i\right)^2,\ \forall i.
\end{aligned} \tag{81}
$$

It is noted that (80b) has the same form as (57b) for **Prob.9**. Thus, the optimal $[\boldsymbol{\theta}]_i$'s to **Prob.12** can be obtained similarly.

Similarly, the capacity maximization and MSE minimization problems in the uplink fully-passive IRS-aided MU-MIMO system have the same forms as these in the single-user case, thus their corresponding optimal solutions can be attained directly. Moreover, it is essential to investigate the general WMMSE minimization problem mostly considered in the downlink fully-passive IRS-aided MU-MIMO system, which is formulated as

$$
\begin{aligned}
\textbf{Prob.13:}\ &\min_{\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}}\ \mathrm{Tr}(\boldsymbol{\Phi}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\boldsymbol{\Pi}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}^{\mathrm{H}})-\mathrm{Tr}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}})-\mathrm{Tr}(\boldsymbol{B}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}^{\mathrm{H}})\\
&\ \mathrm{s.t.}\ |[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}|=1,\ \forall i.
\end{aligned} \tag{82}
$$

Since the element-wise phase derivatives of the objective function w.r.t. $[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}$'s are zeros at the optimal solution, i.e.,

$$
\underbrace{\Im\left\{\left[(\boldsymbol{\Phi}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}\boldsymbol{\Pi})^*\right]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\}-\Im\left\{[\boldsymbol{B}^*]_{i,i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{i,i}\right\}}_{\triangleq g_{i,i}(\boldsymbol{\Lambda}_{\boldsymbol{\Theta}})}=0,\ \forall i, \tag{83}
$$

referring to (54), the optimal solution to **Prob.13** can be easily derived as

$$
\begin{aligned}
[\boldsymbol{\theta}]_i=&\mathrm{Phase}\left\{s_i-[\boldsymbol{B}]_{i,i}\right\}\\
&\text{or }\pi+\mathrm{Phase}\left\{s_i-[\boldsymbol{B}]_{i,i}\right\},\ \forall i,
\end{aligned} \tag{84}
$$

where $s_i=\sum_{m\neq i}[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}}]_{m,m}\left[([\boldsymbol{\Pi}]_{:,i}[\boldsymbol{\Phi}]_{i,:})\right]_{m,m}$.

Algorithm 1 A Novel AO Algorithm for Solving Problems under Constant Modulus Constraints

```
Initialize: Arbitrary five feasible solutions $\widehat{\boldsymbol{X}}_m^{(0)}$, $m = 1,\cdots,5$; iteration index $t=0$; convergence threshold $\epsilon$.
repeat
    for $i=1$ to $N_t$, $j=1$ to $N_{rf}$ do
        Calculate $g_{i,j}(\widehat{\boldsymbol{X}}_m^{(t)}), m=1,\cdots,5$ and the auxiliary vector $\boldsymbol{w}_{i,j}^{(t)}$ as in (87).
        Update $\left[\boldsymbol{\Theta}^{(t)}\right]_{i,j}$ as in (89) and obtain $[\boldsymbol{X}^{(t)}]_{i,j}$ according to (38).
    end for
    Update $\widehat{\boldsymbol{X}}_m^{(t+1)}=\boldsymbol{X}^{(t)}$ for an arbitrary $m\in[1,5]$.
    $t=t+1$.
until The increment/decrement of the objective function value between two consecutive iterations is less than $\epsilon$.
return $\boldsymbol{X}$.
```

In a nutshell, a series of optimization problems in the wireless systems associated with constant modulus constraints are investigated in this subsection, whose optimal solutions are available using the proposed element-wise phase derivatives.

C. A Novel AO Algorithm

It follows from Sec. III-B that these element-wise phase derivatives associated with **Prob.8∼Prob.13** can be mainly classified into two forms. Moreover, each form always has two zero-derivative points. The optimal one in these two zero-derivative points can be determined by an elegant function, as summarized in the following proposition.

Proposition 1. *For different types of objective functions, the element-wise phase derivatives under constant modulus constraints can be mainly summarized as the following two general forms, i.e., the linear form and the conjugate linear form, which are shown as*

$$
g_{i,j}(\boldsymbol{X})=\begin{cases}\Im\left\{A_{i,j}^{\mathrm{L}}[\boldsymbol{X}]_{i,j}+C_{i,j}^{\mathrm{L}}\right\},\forall i,j,\textit{for trace-linear,}\\ \textit{trace-quadratic, log-determinant functions}, \qquad (85a)\\ \Im\left\{A_{i,j}^{\mathrm{CL}}[\boldsymbol{X}]_{i,j}+B_{i,j}^{\mathrm{CL}}[\boldsymbol{X}]_{i,j}^*+C_{i,j}^{\mathrm{CL}}\right\},\forall i,j,\\ \qquad\qquad \textit{for trace-inverse function}, \qquad (85b)\end{cases}
$$

where $A_{i,j}^{\mathrm{L}}$'s, $A_{i,j}^{\mathrm{CL}}$'s, $B_{i,j}^{\mathrm{CL}}$'s and $C_{i,j}^{\mathrm{CL}}$'s are all complex scalars and $C_{i,j}^{\mathrm{L}}$'s are real scalars. In particular, the linear element-wise phase derivative in (85a) can be regarded as a simplified case of its conjugate linear counterpart by setting $B_{i,j}^{\mathrm{CL}}=0$ and $\Im\left\{C_{i,j}^{\mathrm{CL}}\right\}=0,\forall i,j$ in (85b). Moreover, there are two points $\left\{[\boldsymbol{X}_1]_{i,j},[\boldsymbol{X}_2]_{i,j}\right\}$ satisfying $g_{i,j}(\boldsymbol{X})=0,\forall i,j$, from which the optimal solution can be determined by the following conjugate linear function

$$
[\boldsymbol{X}_{\mathrm{opt}}]_{i,j}=\arg_{\left\{[\boldsymbol{X}_1]_{i,j},[\boldsymbol{X}_2]_{i,j}\right\}} \tag{86}
$$

$$
\begin{cases}\Re\{A_{i,j}^{\mathrm{CL}}[\boldsymbol{X}]_{i,j}-B_{i,j}^{\mathrm{CL}}[\boldsymbol{X}]_{i,j}^*\}\geq 0,\ \textit{for min problem},\\ \Re\{A_{i,j}^{\mathrm{CL}}[\boldsymbol{X}]_{i,j}-B_{i,j}^{\mathrm{CL}}[\boldsymbol{X}]_{i,j}^*\}<0,\ \textit{for max problem},\end{cases}\forall i,j.
$$

Proof. The detailed proof is shown in Appendix A.

Based on **proposition 1**, it is seen that the optimal $[\boldsymbol{X}]_{i,j}$'s are obtained by aligning their phase-shifts with the corresponding counterparts jointly determined by $A_{i,j}^{\mathrm{CL}}$'s, $B_{i,j}^{\mathrm{CL}}$'s and $C_{i,j}^{\mathrm{CL}}$'s, which are all related to $[\boldsymbol{X}]_{m,n}$'s, $m\neq i, n\neq j$

and need to be frequently calculated in each iteration of updating $[\boldsymbol{X}]_{i,j}$'s. In addition, the calculations of $A_{i,j}^{\mathrm{CL}}$'s, $B_{i,j}^{\mathrm{CL}}$'s and $C_{i,j}^{\mathrm{CL}}$'s all involve complicated matrix inversion with complexity of $\mathcal{O}\left(N_r^3\right)$ and SVD with complexity of $\mathcal{O}\left(2N_rN_t^2+N_r^3\right)$. It is evident that this complexity will become enormous as N_r and N_t increases. In order to avoid the frequent calculations of $A_{i,j}^{\mathrm{CL}}$'s, $B_{i,j}^{\mathrm{CL}}$'s and $C_{i,j}^{\mathrm{CL}}$'s, we next derive the optimal solution directly based on the functions $g_{i,j}(\boldsymbol{X})$'s associated with the original element-wise phase derivatives, which is shown in **Proposition 2**.

Proposition 2. *Define arbitrary five feasible solutions satisfying constant modulus constraints, i.e., $\widehat{\boldsymbol{X}}_m$, $m=1,\cdots,5$ and calculate their corresponding $g_{i,j}(\widehat{\boldsymbol{X}}_m)$'s, we have*

$$\boldsymbol{w}_{i,j}=\boldsymbol{R}_{i,j}^{-1}\boldsymbol{t}_{i,j} \text{ with } \boldsymbol{R}_{i,j}\in\mathbb{C}^{5\times5},\boldsymbol{t}_{i,j}\in\mathbb{C}^{5\times1},\ \forall i,j, \tag{87}$$

where

$$\begin{aligned}
[\boldsymbol{R}_{i,j}]_{m,:}=&\Big[\Im\{[\widehat{\boldsymbol{X}}_m]_{i,j}\}-[\boldsymbol{t}_{i,j}]_m\Re\{[\widehat{\boldsymbol{X}}_m]_{i,j}\},\Re\{[\widehat{\boldsymbol{X}}_m]_{i,j}\}\\
&+[\boldsymbol{t}_{i,j}]_m\Im\{[\widehat{\boldsymbol{X}}_m]_{i,j}\},-\Im\{[\widehat{\boldsymbol{X}}_m]_{i,j}\}-[\boldsymbol{t}_{i,j}]_m\Re\{[\widehat{\boldsymbol{X}}_m]_{i,j}\},\\
&\Re\{[\widehat{\boldsymbol{X}}_m]_{i,j}\}-[\boldsymbol{t}_{i,j}]_m\Im\{[\widehat{\boldsymbol{X}}_m]_{i,j}\},1\Big],\\
[\boldsymbol{t}_{i,j}]_m=&\tan(\angle g_{i,j}(\widehat{\boldsymbol{X}}_m)),\ \forall i,j,m.
\end{aligned}\tag{88}$$

Then, we obtain the optimal solutions of optimization problems with conjugate linear element-wise phase derivatives as

$$[\boldsymbol{\Theta}]_{i,j}=\begin{cases}
-\arctan\left(\frac{\widehat{z}_{2,i,j}}{\widehat{z}_{1,i,j}}\right)+\arcsin\left(-\frac{[\boldsymbol{w}_{i,j}]_5}{\sqrt{\widehat{z}_{1,i,j}^2+\widehat{z}_{2,i,j}^2}}\right)\\
\quad \textit{or } \pi-\arctan\left(\frac{\widehat{z}_{2,i,j}}{\widehat{z}_{1,i,j}}\right)+\arcsin\left(-\frac{[\boldsymbol{w}_{i,j}]_5}{\sqrt{\widehat{z}_{1,i,j}^2+\widehat{z}_{2,i,j}^2}}\right),\\
\qquad\qquad\qquad \textit{if } \widehat{z}_{1,i,j}\geq 0,\ \forall i,j,\\
-\pi+\arctan\left(\frac{\widehat{z}_{2,i,j}}{\widehat{z}_{1,i,j}}\right)+\arcsin\left(-\frac{[\boldsymbol{w}_{i,j}]_5}{\sqrt{\widehat{z}_{1,i,j}^2+\widehat{z}_{2,i,j}^2}}\right)\\
\quad \textit{or } \arctan\left(\frac{\widehat{z}_{2,i,j}}{\widehat{z}_{1,i,j}}\right)+\arcsin\left(-\frac{[\boldsymbol{w}_{i,j}]_5}{\sqrt{\widehat{z}_{1,i,j}^2+\widehat{z}_{2,i,j}^2}}\right),\\
\qquad\qquad\qquad \textit{if } \widehat{z}_{1,i,j}< 0,\ \forall i,j,
\end{cases}\tag{89}$$

where $\widehat{z}_{1,i,j}=[\boldsymbol{w}_{i,j}]_1-[\boldsymbol{w}_{i,j}]_3$ and $\widehat{z}_{2,i,j}=[\boldsymbol{w}_{i,j}]_2+[\boldsymbol{w}_{i,j}]_4,\forall i,j$. In particular, for the special case of the linear element-wise phase derivatives, we have $\boldsymbol{w}_{i,j}=[[\boldsymbol{w}_{i,j}]_1,[\boldsymbol{w}_{i,j}]_2,0,0,0]^{\mathrm{T}}$ and $\widehat{z}_{1,i,j}$ does not affect the optimal $[\boldsymbol{\Theta}]_{i,j}$.

Proof. The detailed proof is shown in Appendix B.

Based on **Proposition 2**, we next aim to develop a novel AO algorithm with the aid of five arbitrary feasible solutions to determine the optimal solutions of all above optimization problems under constant modulus constraints, which is summarized in **Algorithm 1**.

Remark 3: In fact, the element-wise phase derivatives for different optimization problems with constant modulus constraints can be roughly summarized as a general (conjugate) linear form, and the corresponding optimal solutions can be determined from two potential solutions based on a concise conjugate linear function. Based on this, a novel AO algorithm with the advantages of the low complexity and guaranteed performance is developed. Moreover, for the optimization problems that have other nonconvex constraints such as the fractional functions (e.g., signal-to-interference-plus-noise ratio) [34] or the difference of convex functions (e.g., secrecy rate) [35], the proposed novel AO algorithm can also be used to derive the closed-form solutions. Specifically, we simultaneously employ the successive convex approximation technique to handle the nonconvex constraints and the Lagrange multiplier method to reformulate the objective function. The corresponding Lagrangian belongs to one of the four types of the considered functions in Sec. II-A and thus can be efficiently solved by the proposed algorithms. However, dual variables are introduced and need to be further determined by the subgradient method.

TABLE III
COMPUTATIONAL COMPLEXITIES OF ALL STUDIED ALGORITHMS

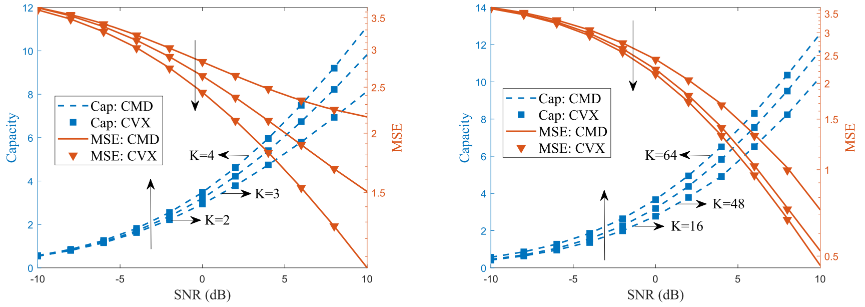
Algorithm	Computational complexity
Proposed novel AO algorithm	$\mathcal{O}\left(KN_r^3\right)$
Element-wise BCD algorithm	$\mathcal{O}\left(2KN_r^2N_t+3KN_r^3\right)$
MM-based algorithm	$\mathcal{O}\left(T_{\mathrm{MM}}K^2\right)$
RCG-based algorithm	$\mathcal{O}\left(K^2+T_{\mathrm{RCG}}\left(K^3\right)\right)$

D. Convergence and Complexity Analysis of the Novel AO Algorithm

In this subsection, we firstly demonstrate the convergence of the proposed novel AO algorithm to a locally optimal solution. By leveraging the element-wise phase derivative, the proposed algorithm firstly decouples the original nonconvex problem under constant modulus constraints into multiple unconstrained subproblems associated with each matrix element, and then derives their optimal closed-form solutions since it satisfies the first-order KKT optimality condition in each subproblem. Thus, the objective function is monotonically non-increasing over the iterations. Moreover, the objective function is lower bounded since the feasible region of the original problem is closed. Therefore, the proposed algorithm is guaranteed to converge to a stationary solution for the original problem. That is to say, its local optimality is thus ensured [36].

Then, taking the capacity maximization problem in the fully-passive IRS-aided MIMO system as an example, we analyze and compare the computational complexities of the proposed novel AO algorithm and the advanced algorithms for constant modulus constrained problem, which is summarized in Table III. It is observed that the computational complexity of the proposed novel AO algorithm is dominated by the matrix inversion for calculating $g_{i,j}(\widehat{\boldsymbol{X}}_m),\forall m,i,j$, which is computed as $\mathcal{O}\left(KN_r^3\right)$.

Currently, three effective algorithms are widely utilized to optimize the constant modulus constrained problems, i.e., the element-wise BCD algorithm [18], the MM-based algorithm [21] and the Riemannian conjugate gradient (RCG)-based algorithm [38]. Similarly, the computational complexities of these advanced algorithms can be elaborated as below. Specifically, the element-wise BCD algorithm has a high complexity of $\mathcal{O}\left(2KN_r^2N_t+3KN_r^3\right)$, which mainly comes from operations of matrix inversion and SVD. Moreover, the computational complexity of the MM-based algorithm is given by $\mathcal{O}\left(T_{\mathrm{MM}}K^2\right)$ with T_{MM} denoting the number of iterations. The RCG-based algorithm primarily comprises three components, i.e., the computation of Riemannian gradient, the retraction operator and the Armijo backtracking line search,



(a) Uplink MU-SIMO system

(b) Amplitude-adjustable IRS-aided MIMO system

Fig. 3. The capacity and MSE performance comparison in the two systems considered in Sec. II.

whose complexity is calculated as $\mathcal{O}\left(K^2+T_{\text{RCG}}\left(K^3\right)\right)$ with T_{RCG} being the line search times. In conclusion, since $K \gg \{N_r, N_t\}$, the proposed novel AO algorithm achieves a significant reduction in the computational complexity as compared to the advanced algorithms. Therefore, it is more suitable for practical applications.

IV. Simulations and Discussions

In this section, numerical simulation results are provided to evaluate the performance of the derived optimal closed-form solutions based on complex matrix derivatives in Sec. II-B (also referred to as CMD-based algorithm), and the novel AO algorithm in **Algorithm 1**, which are respectively proposed for tackling the optimization problems under diagonal structure constraints and constant modulus constraints.

A. Diagonal Structure Constraints

We firstly consider the uplink MU-SIMO system, where $K=2,3$ and 4 single-antenna users transmit signals to the BS equipped with $N_t=6$ antennas, respectively. Moreover, the maximum transmit power of each user is assumed to be $P_k=28$ dBm, 26 dBm and 25 dBm for $K=2$, 3 and 4 users, respectively, and the maximum sum power is set as $P=30$ dBm. Under the assumption that the channel follows the circularly symmetric complex Gaussian distribution with unit noise variance, i.e. $\boldsymbol{H} \sim \mathcal{CN}\left(\mathbf{0}, \boldsymbol{I}_{N_tK}\right)$, the SNR is defined as $\text{SNR} = 10\log_{10}(\frac{P}{\sigma^2})$, where noise power σ^2 varies with SNR. All the results are obtained by averaging over 100 channel realizations. Firstly, Fig. 3(a) compares the capacity and MSE performance achieved by the CMD-based algorithm and the CVX toolbox [37] versus SNR under different numbers of users. For each considered number of users, it is clearly observed that the CMD-based algorithm achieves almost the same capacity and MSE performance as the numerical CVX optimization with a lower complexity. Specifically, the complexity of the CMD-based algorithm is dominated by matrix inversion with the complexity of $\mathcal{O}\left(K^3\right)$, while the numerical CVX optimization adopting the interior-point method has a complexity of $\mathcal{O}\left((K+N_tK)^{3.5}+K^3\right)$. Moreover, as expected that with the expansion of the number of users, the capacity and MSE performance are further enhanced due to the increasing system degrees of freedom. Thus, the global optimality and low complexity of the CMD-based algorithm are demonstrated.

Then, the amplitude-adjustable IRS-aided MIMO system is taken into account, where the BS equipped with $N_t=6$ antennas and $N_{rf}=4$ RF chains communicates with the user equipped with $N_r=4$ antennas, while $K=16,48$ and 64-element IRSs are deployed to enhance the point-to-point communication, respectively. The path loss setting is the same as that in [17] and other parameter settings are the same as those in the uplink MU-SIMO system. Fig. 3(b) illustrates that the capacity and MSE performance attained by the CMD-based algorithm and the numerical CVX optimization as the function of SNR. We find that the CMD-based algorithm attains almost the same optimal performance as the numerical CVX toolbox. Also, the increase in the number of IRS elements results in the improved capacity and MSE performance.

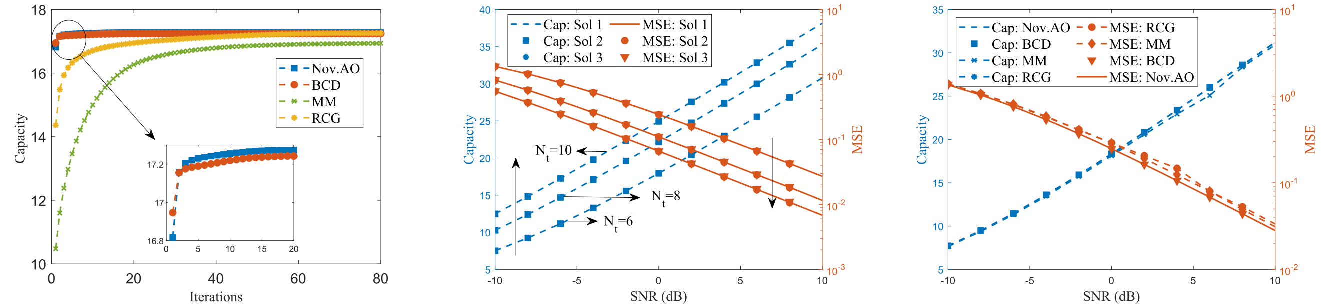
B. Constant Modulus Constraints

In the point-to-point hybrid analog-digital MIMO system, the BS equipped with $N_t=6$ antennas and $N_{rf}=4$ RF chains serves a user equipped with $N_r=4$ antennas. Other parameter settings are the same as those of the uplink MU-SIMO system. Moreover, we compare the proposed novel AO algorithm with the following benchmark schemes: **BCD**: The authors of [20] propose an element-wise BCD algorithm for optimizing the analog beamforming matrix. **MM**: The authors of [21] equivalently reformulate the original problem into the WMMSE minimization problem, and then optimize the analog beamforming matrix using the MM-based algorithm. **RCG**: The RCG-based algorithm is applied to effectively solve the corresponding optimization problem [38].

Taking the capacity maximization problem as an example, we firstly depict Fig. 4(a) to show the convergence behavior of all studied algorithms, where SNR$=-2$ dB. As can be seen, the proposed novel AO algorithm obviously outperforms the MM-based and RCG-based algorithms in terms of the convergence speed and capacity performance. Moreover, both the proposed algorithm and the element-wise BCD algorithm converge within 20 iterations, while the proposed algorithm shows a little higher capacity and a faster speed than the element-wise BCD algorithm.

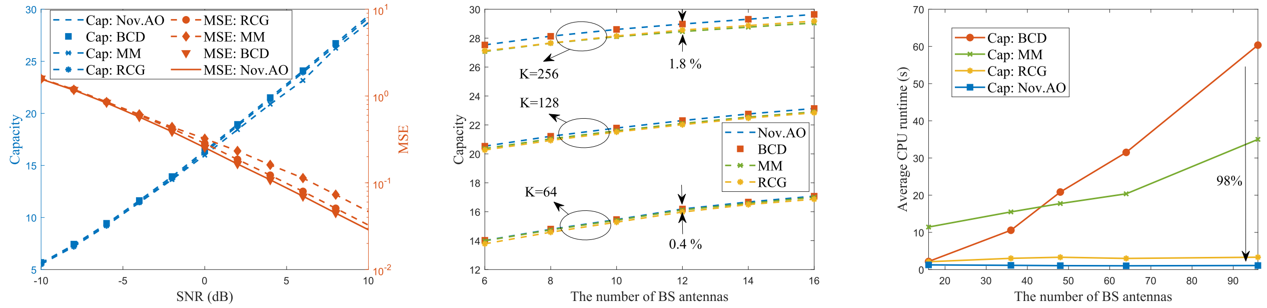
Next, we demonstrate the effectiveness of the proposed novel AO algorithm in Fig. 4(b), where three sets of different feasible solutions satisfying constant modulus constraints are generated randomly to calculate the optimal solutions of the capacity maximization and MSE minimization problems for different numbers of BS antennas, respectively. The same capacity and MSE performance are attained for different feasible solutions under each antenna number setting and these two performance are enhanced with the increasing number of BS antennas, implying the stability and effectiveness of the proposed novel AO algorithm.

Fig. 4(c) compares the capacity and MSE performance of the proposed novel AO algorithm and benchmark schemes



(a) Convergence behaviors of all studied algorithms (b) For different feasible solutions (c) For all studied algorithms

Fig. 4. The capacity and MSE performance comparison in the hybrid analog-digital MIMO system.



(a) For all studied algorithms (b) For different system sizes (c) Average CPU runtime comparison

Fig. 5. The capacity and MSE performance comparison in the fully-passive IRS-aided MIMO system.

versus SNR. Obviously, the proposed novel AO algorithm consistently performs better than the MM-based and RCG-based algorithms, since both benchmark schemes do not directly act on the original objective function. To be specific, the MM-based and RCG-based algorithms respectively adopt the surrogate function and the gradient projection to tackle constant modulus constraints, respectively, whose performance depends on the choice of the initial point and the descending step size. Moreover, the proposed novel AO algorithm attains nearly identical performance to the element-wise BCD algorithm, since both the two algorithms derive the optimal closed-form solution associated with each matrix element.

Furthermore, we consider the fully-passive IRS-aided MIMO system with the same parameter settings as the amplitude-adjustable IRS-aided MIMO system. In Fig. 5(a), the capacity and MSE performance of the proposed novel AO algorithm are respectively compared with that of benchmark schemes in the hybrid analog-digital MIMO system. Notice that the above three benchmark schemes, i.e., the element-wise BCD algorithm [18], [19], the MM-based algorithm and the RCG-based algorithm, have also been widely adopted in most existing IRS related works and thus are still considered here. The observed trend in Fig. 5(a) indicates that the proposed novel AO algorithm is able to achieve better performance than the MM-based and RCG-based algorithms, and almost the same performance as the element-wise BCD algorithm while maintaining a lower complexity.

Then, taking the capability maximization problem as an example, we compare the proposed novel AO algorithm and benchmark schemes versus the numbers of BS antennas and IRS elements in Fig. 5(b), where SNR$= -2$ dB. It is naturally seen that the capacity performance becomes better with the increasing sizes of BS antennas and IRS elements. Moreover, the performance gap between the proposed novel AO algorithm and benchmark schemes becomes more significant with the increase of the system size. For example, for $N_t = 12$, when the number of IRS elements increases from $K = 64$ to $K = 256$, there is 1.4% capacity increment achieved by the proposed algorithm relative to the MM-based algorithm. Moreover, the proposed algorithm always achieves almost the same performance as the element-wise BCD algorithm while maintaining a lower complexity. Therefore, the proposed algorithm's scalability and optimality are ensured.

Finally, in order to demonstrate the low-complexity advantage of the proposed AO algorithm, we depict Fig. 5(c) to intuitively compare its average CPU runtime with that of benchmark schemes for solving the capacity maximization problem, where SNR$=$ 5 dB and $N_r = 8$. We firstly observe that the proposed novel AO algorithm has the lowest average CPU runtime, since it avoids the complicated matrix operation. Moreover, it shows a sharp decrease of the average CPU runtime relative to the element-wise BCD algorithm. For example, at $N_t = 96$, there is 98% CPU runtime decrement achieved by the proposed algorithm implying its low-complexity advantage. Whereas, the element-wise BCD algorithm suffers from the highest time overhead which also increases as the number of BS antennas increases, since the dimension of the involved matrix inversion is the same as the number of BS antennas.

V. Conclusions

In this paper, we investigated complex matrix derivatives for two special matrices, i.e., diagonal structured matrices and constant modulus structured matrices. Under the diagonal structure constraints, the optimal closed-form solutions of the capacity maximization problem, the MSE minimization problem and their variants can be obtained using complex matrix derivatives. Whereas for constant modulus constraints, the optimal solutions of these classical optimization problems are derived utilizing element-wise phase derivatives. Further, in

$$\begin{aligned}
\tan(\angle g_{i,j}(\widehat{\boldsymbol{X}}_m)) =& \frac{\Im\{g_{i,j}(\widehat{\boldsymbol{X}}_m)\}}{\Re\{g_{i,j}(\widehat{\boldsymbol{X}}_m)\}} \\
=& \frac{\Re\{A^{\mathrm{CL}}_{i,j}\}\Im\{\widehat{\boldsymbol{X}}_m\} + \Im\{A^{\mathrm{CL}}_{i,j}\}\Re\{\widehat{\boldsymbol{X}}_m\} + \Im\{B^{\mathrm{CL}}_{i,j}\}\Re\{\widehat{\boldsymbol{X}}_m\} - \Re\{B^{\mathrm{CL}}_{i,j}\}\Im\{\widehat{\boldsymbol{X}}_m\} + \Im\{C^{\mathrm{CL}}_{i,j}\}}{\Re\{A^{\mathrm{CL}}_{i,j}\}\Re\{\widehat{\boldsymbol{X}}_m\} - \Im\{A^{\mathrm{CL}}_{i,j}\}\Im\{\widehat{\boldsymbol{X}}_m\} + \Re\{B^{\mathrm{CL}}_{i,j}\}\Re\{\widehat{\boldsymbol{X}}_m\} + \Im\{B^{\mathrm{CL}}_{i,j}\}\Im\{\widehat{\boldsymbol{X}}_m\} + \Re\{C^{\mathrm{CL}}_{i,j}\}} \\
=& \frac{[\boldsymbol{w}_{i,j}]_1\Im\{\widehat{\boldsymbol{X}}_m\} + [\boldsymbol{w}_{i,j}]_2\Re\{\widehat{\boldsymbol{X}}_m\} + [\boldsymbol{w}_{i,j}]_4\Re\{\widehat{\boldsymbol{X}}_m\} - [\boldsymbol{w}_{i,j}]_3\Im\{\widehat{\boldsymbol{X}}_m\} + [\boldsymbol{w}_{i,j}]_5}{[\boldsymbol{w}_{i,j}]_1\Re\{\widehat{\boldsymbol{X}}_m\} - [\boldsymbol{w}_{i,j}]_2\Im\{\widehat{\boldsymbol{X}}_m\} + [\boldsymbol{w}_{i,j}]_3\Re\{\widehat{\boldsymbol{X}}_m\} + [\boldsymbol{w}_{i,j}]_4\Im\{\widehat{\boldsymbol{X}}_m\} + 1},\ m = 1, \cdots, 5. \qquad (90)
\end{aligned}$$

order to avoid the complicated matrix operations, we explore the inherent structure of the element-wise phase derivatives, and develop a novel AO algorithm with the aid of several arbitrary feasible solutions. Finally, numerical simulations demonstrate the global optimality and low complexity of the proposed novel AO algorithm.

Appendix

A. Proof of Proposition 1

Based on Sec. III-B, we can easily conclude that for both capacity maximization problems (i.e. **Prob.8**, **Prob.11**) with log-determinant functions and WMMSE minimization problems (i.e. **Prob.10**, **Prob.13**) with trace-linear and trace-quadratic functions, the corresponding element-wise phase derivatives have the same linear forms as in (85a). Whereas, for the MSE minimization problems (i.e. **Prob.9**, **Prob.12**) with trace-inverse functions, the element-wise phase derivatives satisfy the conjugate linear forms in (85b).

Furthermore, in order to determine the optimal solution from two zero-derivative points $\left\{[\boldsymbol{X}_1]_{i,j}, [\boldsymbol{X}_2]_{i,j}\right\}$ satisfying $g_{i,j}(\boldsymbol{X}) = 0, \forall i, j$, we resort to the second-order derivative of the corresponding objective function $f(\boldsymbol{X})$ w.r.t. $[\boldsymbol{\Theta}]_{i,j}$'s. Specifically, by taking the conjugate linear element-wise phase derivatives as an example, we have

$$\frac{\partial^2 f(\boldsymbol{X})}{\partial\,[\boldsymbol{\Theta}]^2_{i,j}} = \frac{\partial g_{i,j}(\boldsymbol{X})}{\partial\,[\boldsymbol{\Theta}]_{i,j}} = 2j\Re\{A^{\mathrm{CL}}_{i,j}[\boldsymbol{X}]_{i,j} - B^{\mathrm{CL}}_{i,j}[\boldsymbol{X}]^*_{i,j}\}. \quad (91)$$

According to the optimization theory, it is readily inferred that these two zero-derivative points are local minimum when $\frac{\partial^2 f(\boldsymbol{X})}{\partial[\boldsymbol{\Theta}]^2_{i,j}} \geq 0$ holds. In contrast, they become local maximum when $\frac{\partial^2 f(\boldsymbol{X})}{\partial[\boldsymbol{\Theta}]^2_{i,j}} < 0$ holds [25]. This completes the proof.

B. Proof of Proposition 2

It follows from **Proposition 1** that the derivation of the optimal $\boldsymbol{X}_{i,j}$'s is based on $A^{\mathrm{CL}}_{i,j}$'s, $B^{\mathrm{CL}}_{i,j}$'s and $C^{\mathrm{CL}}_{i,j}$'s. Thus, we firstly define a 5-dimensional vector as follows.

$$\begin{aligned}
\boldsymbol{w}_{i,j} =& [[\boldsymbol{w}_{i,j}]_1, [\boldsymbol{w}_{i,j}]_2, \cdots, [\boldsymbol{w}_{i,j}]_5]^{\mathrm{T}} \qquad (92)\\
=& \left[\frac{\Re\{A^{\mathrm{CL}}_{i,j}\}}{\Re\{C^{\mathrm{CL}}_{i,j}\}}, \frac{\Im\{A^{\mathrm{CL}}_{i,j}\}}{\Re\{C^{\mathrm{CL}}_{i,j}\}}, \frac{\Re\{B^{\mathrm{CL}}_{i,j}\}}{\Re\{C^{\mathrm{CL}}_{i,j}\}}, \frac{\Im\{B^{\mathrm{CL}}_{i,j}\}}{\Re\{C^{\mathrm{CL}}_{i,j}\}}, \frac{\Im\{C^{\mathrm{CL}}_{i,j}\}}{\Re\{C^{\mathrm{CL}}_{i,j}\}}\right]^{\mathrm{T}},\ \forall i, j.
\end{aligned}$$

Moreover, by recalling **Prob.12**, $A^{\mathrm{CL}}_{i,j}$'s, $B^{\mathrm{CL}}_{i,j}$'s and $C^{\mathrm{CL}}_{i,j}$'s are closely related to the functions $g_{i,j}(\boldsymbol{X})$'s associated with the original element-wise phase derivatives as (90), as shown at the top of this page. Note that the equations in (90) can be further rewritten as the following homogeneous linear equations, i.e.,

$$\begin{aligned}
&[\boldsymbol{w}_{i,j}]_1\left(\Im\{\widehat{\boldsymbol{X}}_m\} - \tan(\angle g_{i,j}(\widehat{\boldsymbol{X}}_m))\Re\{\widehat{\boldsymbol{X}}_m\}\right)\\
&+[\boldsymbol{w}_{i,j}]_2\left(\Re\{\widehat{\boldsymbol{X}}_m\} + \tan(\angle g_{i,j}(\widehat{\boldsymbol{X}}_m))\Im\{\widehat{\boldsymbol{X}}_m\}\right)\\
&-[\boldsymbol{w}_{i,j}]_3\left(\Im\{\widehat{\boldsymbol{X}}_m\} + \tan(\angle g_{i,j}(\widehat{\boldsymbol{X}}_m))\Re\{\widehat{\boldsymbol{X}}_m\}\right)\\
&+[\boldsymbol{w}_{i,j}]_4\left(\Re\{\widehat{\boldsymbol{X}}_m\} - \tan(\angle g_{i,j}(\widehat{\boldsymbol{X}}_m))\Im\{\widehat{\boldsymbol{X}}_m\}\right)\\
&+[\boldsymbol{w}_{i,j}]_5 - \tan(\angle g_{i,j}(\widehat{\boldsymbol{X}}_m)) = 0,\ m = 1, \cdots, 5. \qquad (93)
\end{aligned}$$

It follows from (93) that the optimal $\boldsymbol{w}_{i,j}$'s can be obtained by jointly solving its involved five homogeneous linear equations, whose closed-form structures are further shown in (87) and the proof of **Proposition 2** is completed.

References

[1] J. Yang, and S. Roy, "On joint transmitter and receiver optimization or multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[2] S. Sugiura, S. Chen, and L. Hanzo, "A universal space-time architecture for multiple-antenna aided systems," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 401–420, Apr.–Jun. 2012.

[3] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Commun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[4] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[5] E. Vlachos, G. C. Alexandropoulos, and J. Thompson, "Massive MIMO channel estimation for millimeter wave systems via matrix completion," *IEEE Signal Process. Lett.*, vol. 25, no. 11, pp. 1675–1679, Nov. 2018.

[6] C. Xing, S. Wang, S. Chen, S. Ma, H. V. Poor, and L. Hanzo, "Matrix-monotonic optimization - Part I: Single-variable optimization," *IEEE Trans. Signal Process.*, vol. 69, pp. 738–754, 2021.

[7] Z. Chen, N. Zhao, D. K. C. So, J. Tang, X. Y. Zhang, and K. -K. Wong, "Joint altitude and hybrid beamspace precoding optimization for UAV-enabled multiuser mmWave MIMO system," *IEEE Trans. Veh. Tech.*, vol. 71, no. 2, pp. 1713–1725, Feb. 2022.

[8] H. Vaezy, M. J. Omidi, M. M. Naghsh, and H. Yanikomeroglu, "Energy efficient transceiver design in MIMO interference channels: The selfish, unselfish, worst-case, and robust methods," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5377–5389, Aug. 2019.

[9] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[10] W. M. Jang, and W. Wu, "Distributed and centralized multiuser detection with antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 855–860, May 2005.

[11] S. Han, C. -l. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

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

[13] T. Qiao, Y. Cao, J. Tang, N. Zhao, and K. -K. Wong, "IRS-aided uplink security enhancement via energy-harvesting jammer," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8286–8297, Dec. 2022.

[14] J. Xu, and Y. Liu, "A novel physics-based channel model for reconfigurable intelligent surface-assisted multi-user communication systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1183–1196, Feb. 2022.

[15] X. Jin, P. Zhang, C. Wan, D. Ma, and Y. Yao, "RIS assisted dual-function radar and secure communications based on frequency-shifted chirp spread spectrum index modulation," *China Commun.*, vol. 20, no. 10, pp. 85–99, 2023.

[16] S. Gong, C. Xing, Y. Jing, S. Wang, J. Wang, S. Chen, and L. Hanzo, "A unified MIMO optimization framework relaying on the KKT conditions," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7251–7268, Aug. 2021.

[17] C. Xing, S. Xie, S. Gong, X. Yang, S. Chen, and L. Hanzo, "A KKT conditions based transceiver optimization framework for RIS-aided multi-user MIMO networks," *IEEE Trans. Commun.*, vol. 71, no. 5, pp. 2602–2617, May 2023.

[18] S. Zhang, and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.

[19] X. Zhao, K. Xu, S. Ma, S. Gong, G. Yang, and C. Xing, "Joint transceiver optimization for IRS-aided MIMO communications," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3467–3482, Mar. 2022.

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

[21] S. Gong, C. Xing, V. K. N. Lau, S. Chen, and L. Hanzo, "Majorization-minimization aided hybrid transceivers for MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 4903–4918, 2020.

[22] C. Xing, Y. Li, S. Gong, J. An, S. Chen, and L. Hanzo, "A general matrix variable optimization framework for MIMO assisted wireless communications," *IEEE Trans. Veh. Tech.*, 2023

[23] H. -T. Wai, W. -K. Ma, and A. M. -C. So, "Cheap semidefinite relaxation MIMO detection using row-by-row block coordinate descent," *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Prague, Czech Republic, 2011, pp. 3256–3259.

[24] A. Hjørungnes, *Complex-valued matrix derivatives: With applications in signal processing and communications*. Cambridge University Press: Cambridge, UK, 2011.

[25] S. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[26] Y. Yu, and Y. Hua, "Power allocation for a MIMO relay system with multiple-antenna users," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2823–2835, May 2010

[27] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5849–5863, Sept. 2020.

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

[29] S. S. Christensen, R. Agarwal, E. d. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[30] S. Serbetli, and A. Yener, "Transceiver optimization for multisuer MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 214–226, Jan. 2004.

[31] S.-R. Lee, J.-S. Kim, S.-H. Moon, H.-B. Kong, and I. Lee, "Zero-forcing beamforming in multiuser MISO downlink systems under per-antenna power constraint and equal-rate metric," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 228–236, Jan. 2013

[32] D. P. Palomar, and Y. Jiang, "MIMO transceiver design via majorization theory," *Found. Trends Commun. Inf. Theory*, vol. 3, no. 4, pp. 331–551, 2006.

[33] C. Xing, X. Zhao, W. Xu, X. Dong, and G. Y. Li, "A framework on hybrid MIMO transceiver design based on matrix-monotonic optimization," *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3531–3546, Jul. 2019.

[34] Q. Shi, M. Razaviyayn, M. Hong, and Z. -Q. Luo, "SINR constrained beamforming for a MIMO multi-user downlink system: Algorithms and convergence analysis," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2920–2933, Jun. 2016.

[35] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Trans. Veh. Tech.*, vol. 63, no. 4, pp. 1678–1690, May 2014.

[36] M. V. Solodov, "On the convergence of constrained parallel variable distribution algorithms," *SIAM J. Optim.*, vol. 8, no. 1, pp. 187–196, Feb. 1998.

[37] M. Grant, and S. Boyd. (Sep. 2013). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.0 Beta.* [Online]. Available: http://cvxr.com/cvx

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