Hybrid Precoding for Millimeter Wave MIMO Systems: A Matrix Factorization Approach

Juening Jin, Yahong Rosa Zheng, *Fellow, IEEE*, Wen Chen, *Senior Member, IEEE*, and Chengshan Xiao, *Fellow, IEEE*

***Abstract*— This paper investigates the hybrid precoding design for millimeter wave multiple-input multiple-output systems with finite-alphabet inputs. The precoding problem is a joint optimization of analog and digital precoders, and we treat it as a matrix factorization problem with power and constant modulus constraints. This paper presents three main contributions. First, we present a sufficient condition and a necessary condition for hybrid precoding schemes to realize unconstrained optimal precoders exactly when the number of data streams N_{s} satisfies $N_{\mathrm{s}} = \min\{\mathrm{rank}(\mathrm{H}), N_{\mathrm{rf}}\}$, where H represents the channel matrix and N_{rf} is the number of radio frequency chains. Second, we show that the coupled power constraint in our matrix factorization problem can be removed without loss of optimality. Third, we propose a Broyden-Fletcher-Goldfarb-Shanno-based algorithm to solve our matrix factorization problem using gradient and Hessian information. Several numerical results are provided to show that our proposed algorithm outperforms existing hybrid precoding algorithms.**

***Index Terms*— Hybrid precoding, finite-alphabet inputs, matrix factorization, nonconvex optimization.**

I. Introduction

MILLIMETER wave (mm-wave) multiple-input multiple-output (MIMO) communication is a promising technology for future generation cellular systems to address the wireless spectrum crunch. It makes use of the mm-wave band from 30 GHz to 300 GHz, which implies a much wider bandwidth than current cellular systems operating in microwave bands. Moreover, a short wavelength of radio signals in the mm-wave band enables large number of antennas to be equipped in transceivers, and this allows for applying massive multiple-input multiple-output (MIMO) technique in mm-wave communication systems.

For conventional MIMO systems, linear precoding is utilized to maximize the data rate, and it is implemented in the digital domain by the unconstrained optimal precoder. However, the implementation of unconstrained optimal precoders requires one radio frequency (RF) chain per antenna, which will result in prohibitive cost and power consumption in mm-wave MIMO systems. To address this issue, a hybrid precoding scheme has been proposed for mm-wave MIMO systems to reduce the number of RF chains [1]–[7]. This scheme divides the linear precoder into analog and digital precoders, which are implemented in analog and digital domains, respectively. The digital precoder is realized by a small amount of RF chains, and the analog precoder is realized by phase shifters. Due to the property of phase shifters, each entry of the analog precoder satisfies the constant modulus constraint. These nonconvex constant modulus constraints form a major barrier for hybrid precoding design.

Several hybrid precoding algorithms have been proposed for mm-wave MIMO systems [1]–[7]. The work in [1] first formulated the hybrid precoding problem as a matrix factorization problem, and then applied the orthogonal matching pursuit (OMP) algorithm to find near-optimal analog and digital precoders. Yu *et al.* [3] utilized the formulation proposed in [1], and then employed a manifold based alternating minimization algorithm to design hybrid precoders. References [5] and [7] introduced and analyzed low complexity hybrid precoding algorithms based on the matrix factorization. There were also some studies on how to achieve the performance of unconstrained optimal precoders with hybrid precoding schemes [2], [4], yet requiring the number of RF chains to be twice as much as the number of data streams.

Most existing works on hybrid precoding assume Gaussian inputs, which are rarely realized in practice. It is well known that practical systems utilize finite-alphabet inputs, such as phase-shift keying (PSK) or quadrature amplitude modulation (QAM). Furthermore, it has been shown that precoding designs under Gaussian inputs are quite suboptimal for practical

Manuscript received July 29, 2017; revised November 29, 2017, January 13, 2018, and February 12, 2018; accepted February 12, 2018. Date of publication March 5, 2018; date of current version May 8, 2018. The work of Y. R. Zheng and C. Xiao was supported by the U.S. National Science Foundation under Grant ECCS-1231848, Grant ECCS-1408316, and Grant ECCS-1539316. The work of W. Chen was supported in part by the National Natural Science Foundation of China under Grant 61671294, in part by the STCSM Project under Grant 16JC1402900 and Grant 17510740700, in part by the Natural Science Foundation of Guangxi Province under Grant 2015GXNSFDA139037, and in part by the National Science and Technology Major Project under Grant 2017ZX03001002-005 and Grant 2018ZX03001009-002. This paper was presented in part at the IEEE GLOBECOM, Singapore, 2017. The associate editor coordinating the review of this paper and approving it for publication was J. Zhang. *(Corresponding author: Juening Jin.)*

J. Jin was with the Missouri University of Science and Technology, Rolla, MO 65409 USA. He is now with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: jueningjin@gmail.com).

Y. R. Zheng is with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409 USA (e-mail: zhengyr@mst.edu).

W. Chen is with the Shanghai Key Laboratory of Navigation and Location Based Services, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: wenchen@sjtu.edu.cn).

C. Xiao is with the Department of Electrical and Computer Engineering, Lehigh University, Bethlehem, PA 18015 USA (e-mail: xiaoc@lehigh.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2018.2810072

systems with finite-alphabet inputs [8]–[14]. A unified framework for linear precoding design under finite-alphabet inputs has been proposed in [15]. Recently, the authors in [6] presented an iterative gradient ascent algorithm for mm-wave MIMO systems with finite-alphabet inputs. In each iteration, the gradient ascent algorithm updated the unconstrained precoder using gradient information, and then it employed a heuristic way to partition the unconstrained precoder into analog and digital precoders. Simulation results illustrated that the gradient ascent algorithm can achieve up to 0.4 bps/Hz gains compared to the Gaussian inputs scenario.

A. Contributions

In this paper, we investigate the hybrid precoding design for mm-wave MIMO systems with finite-alphabet inputs. The contributions of this paper are summarized as follows:

- We first provide a sufficient condition under which hybrid precoding schemes can realize any unconstrained optimal precoders exactly. When the sufficient condition does not hold, we also present a necessary condition for hybrid precoding to achieve the performance of unconstrained optimal precoders.
- We prove that the power constraint in the hybrid precoding problem (10) can be removed without loss of local and/or global optimality. This result greatly simplifies the precoding design, and it enable us to design an efficient algorithm for the hybrid precoding problem.
- We present closed form expressions for gradient and Hessian of the hybrid precoding problem. Then we utilize these information to design a BFGS-based algorithm. The proposed algorithm outperforms existing hybrid precoding algorithms.

B. Notations

The following notations are adopted throughout the paper: Boldface lowercase letters, boldface uppercase letters, and calligraphic letters are used to denote vectors, matrices and sets, respectively. The real and complex number fields are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. The superscripts $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathrm{H}}$ stand for transpose, conjugate, and conjugate transpose operations, respectively. $\mathrm{tr}(\cdot)$ is the trace of a matrix; $\|\cdot\|$ denotes the Euclidean norm of a vector; $\|\cdot\|_F$ represents the Frobenius norm of a matrix; $E_{\mathbf{x}}(\cdot)$ represents the statistical expectation with respect to $\mathbf{x}$; $\mathbf{X}_{kl}$ represents the (k,l)-th element of $\mathbf{X}$; $\mathbf{I}$ and 0 denote an identity matrix and a zero matrix, respectively, with appropriate dimensions; $\mathbf{X} \succeq 0$ denotes a positive semidefinite matrix; $\otimes$ and $\circ$ are Kronecker and Hadamard matrix products, respectively; $\mathcal{I}(\cdot)$ represents the mutual information; $\Re$ and $\Im$ are the real and image parts of a complex value; $\log(\cdot)$ is used for the base two logarithm.

II. System Model and Problem Formulation

In this section, we present system and channel models for mm-wave MIMO systems, and then formulate the hybrid precoding design as a matrix factorization problem. Finally, we briefly introduce a few notations on complex matrix derivatives.

A. System Model

Consider a point-to-point mm-wave MIMO system, where a transmitter with N_{t} antennas sends N_{s} data streams to a receiver with N_{r} antennas. The number of RF chains at the transmitter is N_{rf}, which satisfies $N_{\mathrm{s}} \leq N_{\mathrm{rf}} \leq N_{\mathrm{t}}$. We consider the hybrid precoding scheme, where N_{s} data streams are first precoded using a digital precoder, and then shaped by an analog precoder. The received baseband signal $\mathbf{y} \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$ can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{r}} \times N_{\mathrm{t}}}$ is the channel matrix; $\mathbf{F}_{\mathrm{RF}} \in \mathbb{F}^{N_{\mathrm{t}} \times N_{\mathrm{rf}}}$ is the analog precoder with $\mathbb{F} = \left\{f \middle| |f| = \frac{1}{\sqrt{N_{\mathrm{t}}}}\right\}$ being the constant modulus set; $\mathbf{F}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{rf}} \times N_{\mathrm{s}}}$ is the digital precoder; $\mathbf{x} \in \mathbb{C}^{N_{\mathrm{s}} \times 1}$ is the input data vector and $\mathbf{n} \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$ is the independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian noise with zero-mean and covariance $\sigma^2\mathbf{I}$.

Suppose that the channel $\mathbf{H}$ is known at both the transmitter and receiver, and each entry of the input data vector $\mathbf{x}$ is uniformly distributed from a given constellation set with cardinality M. Then the input-output mutual information is given by [12]

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = N_{\mathrm{s}}\log M - \frac{1}{M^{N_{\mathrm{s}}}}\sum_{m=1}^{M^{N_{\mathrm{s}}}} E_{\mathbf{n}}\left\{\log\sum_{k=1}^{M^{N_{\mathrm{s}}}} e^{-d_{mk}}\right\} \tag{2}$$

where $d_{mk} = \sigma^{-2}(\|\mathbf{H}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}(\mathbf{x}_m - \mathbf{x}_k) + \mathbf{n}\|^2 - \|\mathbf{n}\|^2)$, with $\mathbf{x}_m$ and $\mathbf{x}_k$ being two possible input data vectors from $\mathbf{x}$.

B. Channel Model

The mm-wave MIMO channel can be characterized by standard multipath models. Suppose the number of physical propagation paths between the transmitter and the receiver is L. Each path ℓ is described by three parameters: complex gain α_ℓ, angle of arrival $\theta_{\mathrm{r},\ell}$ and angle of departure $\theta_{\mathrm{t},\ell}$. The angles $\{\theta_{\mathrm{r},\ell}\}_{\ell=1}^{L}$ and $\{\theta_{\mathrm{t},\ell}\}_{\ell=1}^{L}$ are i.i.d. uniformly distributed over $[0, 2\pi)$, and the complex gains $\{\alpha_\ell\}_{\ell=1}^{L}$ are i.i.d. complex Gaussian distributed with zero-mean and unit-variance. Under this model, the channel matrix $\mathbf{H}$ is given by [16, Ch. 7.3.2]

$$\mathbf{H} = \sqrt{\frac{N_{\mathrm{r}}N_{\mathrm{t}}}{L}}\sum_{\ell=1}^{L}\alpha_\ell\mathbf{a}(\theta_{\mathrm{r},l})\mathbf{a}(\theta_{\mathrm{t},l})^H \tag{3}$$

where $\mathbf{a}(\theta_{\mathrm{t},l})$ and $\mathbf{a}(\theta_{\mathrm{r},l})$ are array steering vectors of the transmit and receive antenna arrays. In this paper, the transmitter and receiver adopt uniform linear arrays, whose array steering vector $\mathbf{a}(\theta)$ is given by

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{N}}\left[1, e^{-\jmath\frac{2\pi}{\lambda}d\sin\theta}, \ldots, e^{-\jmath\frac{2\pi}{\lambda}d(N-1)\sin\theta}\right]^T \tag{4}$$

where N is the number of antenna element, λ is the wavelength of the carrier frequency and $d = \frac{1}{2}\lambda$ is the antenna spacing.

The channel in (3) can be rewritten in a more compact form as

$$\mathbf{H} = \mathbf{A}_{\mathrm{r}}\mathrm{diag}(\boldsymbol{\alpha})\mathbf{A}_{\mathrm{t}}^H \tag{5}$$

where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_L]^T$, $\mathbf{A}_\mathrm{r} \in \mathbb{C}^{N_\mathrm{r}\times L}$ and $\mathbf{A}_\mathrm{t} \in \mathbb{C}^{N_\mathrm{t}\times L}$ are array steering matrices with constant modulus entries, given by

$$\mathbf{A}_\mathrm{r} = \big[\mathbf{a}(\theta_{\mathrm{r},1}), \ldots, \mathbf{a}(\theta_{\mathrm{r},L})\big] \tag{6}$$

$$\mathbf{A}_\mathrm{t} = \big[\mathbf{a}(\theta_{\mathrm{t},1}), \ldots, \mathbf{a}(\theta_{\mathrm{t},L})\big]. \tag{7}$$

Note that $\mathbf{A}_\mathrm{t}$ is a full rank matrix when the angles $\{\theta_{\mathrm{t},l}\}_{l=1}^{L}$ are distinct [1], and this event occurs with probability one because $\{\theta_{\mathrm{t},l}\}$ are drawn independently from the uniform distribution. Similarly, $\mathbf{A}_\mathrm{r}$ and $\mathrm{diag}(\boldsymbol{\alpha})$ are also full rank matrices with probability one. Therefore, the rank of $\mathbf{H}$ is given by

$$\mathrm{rank}(\mathbf{H}) = \min\{L, N_\mathrm{r}, N_\mathrm{t}\}. \tag{8}$$

C. Problem Formulation

A fundamental approach for hybrid precoding design is to maximize the input-output mutual information under the power and constant modulus constraints. Suppose that the mm-wave receiver can optimally decode data using the received signal $\mathbf{y}$, then the hybrid precoding problem is formulated as

$$\begin{aligned} &\underset{\mathbf{F}_\mathrm{RF}\in\mathcal{U},\mathbf{F}_\mathrm{BB}}{\mathrm{maximize}} && \mathcal{I}(\mathbf{x};\mathbf{y}) \\ &\mathrm{subject\ to} && \mathrm{tr}\big(\mathbf{F}_\mathrm{BB}^H\mathbf{F}_\mathrm{RF}^H\mathbf{F}_\mathrm{RF}\mathbf{F}_\mathrm{BB}\big) \le P \end{aligned} \tag{9}$$

where $\mathcal{I}(\mathbf{x};\mathbf{y})$ is given in (2), P is the transmit power constraint and $\mathcal{U} = \mathbb{F}^{N_\mathrm{t}\times N_\mathrm{rf}}$ is the feasible set of analog precoders. It is challenging to solve problem (9) directly due to two reasons: First, problem (9) is nonconvex because both $\mathcal{I}(\mathbf{x};\mathbf{y})$ and $\mathcal{U}$ are neither convex nor concave with respect to $(\mathbf{F}_\mathrm{RF}, \mathbf{F}_\mathrm{BB})$. Second, iterative algorithms for problem (9) have to evaluate the objective function $\mathcal{I}(\mathbf{x};\mathbf{y})$ in each iteration, which can be very costly because $\mathcal{I}(\mathbf{x};\mathbf{y})$ has no closed form expressions.

To mitigate these difficulties and simplify the precoding design, we adopt the following matrix factorization formulation [1], where hybrid precoders $(\mathbf{F}_\mathrm{RF}, \mathbf{F}_\mathrm{BB})$ are found by approximating the unconstrained optimal precoder $\mathbf{F}_\mathrm{opt}$, i.e.,

$$\begin{aligned} &\underset{\mathbf{F}_\mathrm{RF}\in\mathcal{U},\mathbf{F}_\mathrm{BB}}{\mathrm{minimize}} \ \|\mathbf{F}_\mathrm{opt} - \mathbf{F}_\mathrm{RF}\mathbf{F}_\mathrm{BB}\|_F^2 \\ &\mathrm{subject\ to}\ \mathrm{tr}\big(\mathbf{F}_\mathrm{BB}^H\mathbf{F}_\mathrm{RF}^H\mathbf{F}_\mathrm{RF}\mathbf{F}_\mathrm{BB}\big) \le P. \end{aligned} \tag{10}$$

The unconstrained optimal precoder $\mathbf{F}_\mathrm{opt}$ is given by [12], [13]

$$\mathbf{F}_\mathrm{opt} = \underset{\mathbf{F}\in\mathcal{F}}{\mathrm{maximize}}\ \mathcal{I}(\mathbf{x};\mathbf{y}) \tag{11}$$

where $\mathcal{F} = \big\{\mathbf{F}\big|\mathrm{tr}(\mathbf{F}^H\mathbf{F}) \le P\big\}$.

D. Preliminaries on Complex Matrix Derivatives

The problems investigated in this paper are nonlinear optimization with complex matrix variables, thus we briefly introduce a few definitions on complex matrix derivatives. For a univariate function $f(x) : \mathbb{C} \to \mathbb{R}$, the definition of the complex derivative is given in [17]:

$$\frac{\partial f}{\partial x} \triangleq \frac{1}{2}\left[\frac{\partial f}{\partial \Re(x)} - j\frac{\partial f}{\partial \Im(x)}\right] \tag{12}$$

$$\frac{\partial f}{\partial x^*} \triangleq \frac{1}{2}\left[\frac{\partial f}{\partial \Re(x)} + j\frac{\partial f}{\partial \Im(x)}\right]. \tag{13}$$

For a multivariate function $f(\mathbf{X}) : \mathbb{C}^{n\times r} \to \mathbb{R}$, the partial derivatives with respect to $\mathbf{X}$ and $\mathbf{X}^*$ are matrices

$$\frac{\partial f}{\partial \mathbf{X}} \triangleq \left[\frac{\partial f}{\partial \mathbf{X}_{k\ell}}\right] \ \text{and}\ \frac{\partial f}{\partial \mathbf{X}^*} \triangleq \left[\frac{\partial f}{\partial \mathbf{X}_{k\ell}^*}\right] \tag{14}$$

where $\mathbf{X}_{k\ell}$ denotes the (k,ℓ)-th element of $\mathbf{X}$. In addition, the complex gradient matrix $\nabla_\mathbf{X} f(\mathbf{X})$ is defined as

$$\nabla_\mathbf{X} f(\mathbf{X}) \triangleq \frac{\partial f}{\partial \mathbf{X}^*}. \tag{15}$$

Let $\mathbf{X}_1 \in \{\mathbf{X}, \mathbf{X}^*\}$ and $\mathbf{X}_2 \in \{\mathbf{X}, \mathbf{X}^*\}$, then the complex Hessian of $f(\mathbf{X})$ with respect to $\mathbf{X}_1$ and $\mathbf{X}_2$ is defined in [17]:

$$\mathcal{H}_{\mathbf{X}_1,\mathbf{X}_2} f \triangleq \frac{\partial}{\partial \mathrm{vec}^T(\mathbf{X}_1)}\left[\frac{\partial f}{\partial \mathrm{vec}^T(\mathbf{X}_2)}\right]^T. \tag{16}$$

III. Structures of the Hybrid Precoding Problem

In this section, we first present a sufficient condition and a necessary condition, under which hybrid precoding schemes can realize any unconstrained optimal precoder exactly. Then we prove that the power constraint $\mathrm{tr}\big(\mathbf{F}_\mathrm{BB}^H\mathbf{F}_\mathrm{RF}^H\mathbf{F}_\mathrm{RF}\mathbf{F}_\mathrm{BB}\big) \le P$ in problem (10) can be removed without loss of local and/or global optimality.

A. Optimality of Hybrid Precoding Schemes

The hybrid precoding scheme offers a tradeoff between performance gain and hardware complexity, and its performance is bounded by the unconstrained optimal precoder. When the hybrid precoding scheme can realize any unconstrained optimal precoder exactly, it is an *optimal* scheme. Then a fundamental question arises:

- Question 1: under what conditions can hybrid precoding schemes realize unconstrained optimal precoders exactly?

In other words, we want to find necessary and/or sufficient conditions, under which there exist $(\mathbf{F}_\mathrm{RF}, \mathbf{F}_\mathrm{BB})$ such that $\mathbf{F}_\mathrm{RF} \in \mathcal{U}$ and $\mathbf{F}_\mathrm{opt} = \mathbf{F}_\mathrm{RF}\mathbf{F}_\mathrm{BB}$. The best known result related to this question was shown in [2] and [4]. It states that when the number of data streams N_s satisfies $N_\mathrm{s} \le \frac{1}{2}N_\mathrm{rf}$, we can construct analog and digital precoders to realize any unconstrained optimal precoder with dimensions $N_\mathrm{t} \times N_\mathrm{s}$. However, this result sacrifices the number of data streams to satisfy $\mathbf{F}_\mathrm{opt} = \mathbf{F}_\mathrm{RF}\mathbf{F}_\mathrm{BB}$. In order to achieve the maximum degree of freedom, we should transmit $\min\{\mathrm{rank}(\mathbf{H}), N_\mathrm{rf}\}$ data streams rather than $\frac{1}{2}N_\mathrm{rf}$ data streams. This motivates us to reconsider Question 1 under $N_\mathrm{s} = \min\{\mathrm{rank}(\mathbf{H}), N_\mathrm{rf}\}$.

First, we transform Question 1 into another existence problem through the following proposition.

Proposition 1: Suppose $\mathbf{F}_\mathrm{RF}$ is a full rank matrix, then the following two statements are equivalent:

1) There exists $(\mathbf{F}_\mathrm{RF}, \mathbf{F}_\mathrm{BB})$ such that $\mathbf{F}_\mathrm{RF} \in \mathcal{U}$ and $\mathbf{F}_\mathrm{opt} = \mathbf{F}_\mathrm{RF}\mathbf{F}_\mathrm{BB}$.
2) There exists a full rank square matrix $\mathbf{S} \in \mathbb{C}^{N_\mathrm{rf}\times N_\mathrm{rf}}$ such that $\mathbf{U}_\mathrm{F}\mathbf{S} \in \mathcal{U}$.

Here $\mathbf{U}_\mathrm{F} \in \mathbb{C}^{N_\mathrm{t}\times N_\mathrm{rf}}$ is a semi-unitary matrix whose columns are left singular vectors of $\mathbf{F}_\mathrm{opt}$.

Proof: See Appendix A. ■

Based on Proposition 1, our original problem is equivalent to the existence problem of a full rank square matrix $\mathbf{S}$ satisfying $\mathbf{U}_{\rm F}\mathbf{S} \in \mathcal{U}$. By exploiting the inherent structure of the mm-wave MIMO channel, we provide a sufficient condition to guarantee the existence of such full rank matrix $\mathbf{S}$. The main idea is similar to [18, Th. 1].

Proposition 2: When the number of paths L satisfies $L \leq \min\{N_{\rm r}, N_{\rm t}, N_{\rm rf}\}$, there exists a full rank matrix $\mathbf{S}$ satisfying $\mathbf{A}_{\rm t} = \mathbf{U}_{\rm F}\mathbf{S} \in \mathcal{U}$, where $\mathbf{A}_{\rm t}$ is the array steering matrix given in (7).

Proof: See Appendix A. ■

Combining Propositions 1 and 2, we conclude that when $L \leq \min\{N_{\rm r}, N_{\rm t}, N_{\rm rf}\}$, hybrid precoding schemes can realize any unconstrained optimal precoder $\mathbf{F}_{\rm opt}$ exactly. However, the sufficient condition in Proposition 2 does not always hold in practice because the number of paths may be greater than the number of RF chains. In the rest of this subsection, we propose a necessary condition for the existence of $\mathbf{S}$ satisfying $\mathbf{U}_{\rm F}\mathbf{S} \in \mathcal{U}$, and the proposed necessary condition is independent of L, $N_{\rm rf}$, $N_{\rm r}$ and $N_{\rm t}$.

We first rewrite $\mathbf{U}_{\rm F}\mathbf{S} \in \mathcal{U}$ as

$$\left[\mathbf{U}_{\rm F}\mathbf{s}_\ell\mathbf{s}_\ell^H\mathbf{U}_{\rm F}^H\right]_{kk} = \frac{1}{N_{\rm t}}, \quad k = 1, \ldots, N_{\rm t}, \ \ell = 1, \ldots, N_{\rm rf} \tag{17}$$

where $\mathbf{s}_\ell$ is the ℓth column of $\mathbf{S}$. Combining condition (17) and $\operatorname{rank}(\mathbf{S}) = N_{\rm rf}$, the original problem is equivalent to the existence of $N_{\rm rf}$ linear independent solutions $\{\mathbf{s}_\ell\}_{l=1}^{N_{\rm rf}}$ to the following system of quadratic equations:

$$\left[\mathbf{U}_{\rm F}\mathbf{s}\mathbf{s}^H\mathbf{U}_{\rm F}^H\right]_{kk} = \frac{1}{N_{\rm t}}, \quad k = 1, \ldots, N_{\rm t}. \tag{18}$$

Unfortunately, problem (18) is intractable because checking the existence of solutions to a general quadratic system is NP-hard [19]. Instead, we investigate necessary conditions for the existence of solutions to (18).

The main idea is to transform (18) into a linear system by semidefinite programming. Define $\mathbf{Z} = N_{\rm t}\mathbf{s}\mathbf{s}^H$, the quadratic system (18) can be written as

$$\left[\mathbf{U}_{\rm F}\mathbf{Z}\mathbf{U}_{\rm F}^H\right]_{kk} = 1, \quad \forall k, \ \mathbf{Z} \succeq 0, \ \operatorname{rank}(\mathbf{Z}) = 1. \tag{19}$$

Furthermore, according to

$$\operatorname{vec}\left(\mathbf{U}_{\rm F}\mathbf{Z}\mathbf{U}_{\rm F}^H\right) = \left(\mathbf{U}_{\rm F}^* \otimes \mathbf{U}_{\rm F}\right)\operatorname{vec}(\mathbf{Z}) \tag{20}$$

equations (19) is expressed more compactly as

$$\mathbf{K}_{\rm F}\operatorname{vec}(\mathbf{Z}) = \mathbf{1}, \quad \mathbf{Z} \succeq 0, \ \operatorname{rank}(\mathbf{Z}) = 1 \tag{21}$$

where the kth row of $\mathbf{K}_{\rm F}$ is chosen as the $\left[(k-1)N_{\rm t}+k\right]$th row of $\mathbf{U}_{\rm F}^* \otimes \mathbf{U}_{\rm F}$. Through some standard algebraic manipulations, we can express $\mathbf{K}_{\rm F}$ as

$$\mathbf{K}_{\rm F} = \left[\operatorname{diag}(\mathbf{u}_1^*)\mathbf{U}_{\rm F}, \ldots, \operatorname{diag}\left(\mathbf{u}_{N_{\rm rf}}^*\right)\mathbf{U}_{\rm F}\right] \tag{22}$$

where $\mathbf{u}_\ell$ represents the ℓth column of $\mathbf{U}_{\rm F}$.

The main barrier for solving equations (21) is the nonlinear constraints $\mathbf{Z} \succeq 0$ and $\operatorname{rank}(\mathbf{Z}) = 1$, which restrict solutions of $\mathbf{K}_{\rm F}\operatorname{vec}(\mathbf{Z}) = \mathbf{1}$ with a certain structure. Therefore, we first relax the nonlinear constraints and focus on the linear system $\mathbf{K}_{\rm F}\operatorname{vec}(\mathbf{Z}) = \mathbf{1}$. Clearly, if equations (21) has $N_{\rm rf}$ linear independent solutions, then $\mathbf{K}_{\rm F}\operatorname{vec}(\mathbf{Z}) = \mathbf{1}$ should have at least $N_{\rm rf}$ linear independent solutions. Based on this observation, the following proposition provides a necessary condition for the existence of a full rank $\mathbf{S}$ such that $\mathbf{U}_{\rm F}\mathbf{S} \in \mathcal{U}$.

Proposition 3: If there exist a full rank square matrix $\mathbf{S}$ satisfying $\mathbf{U}_{\rm F}\mathbf{S} \in \mathcal{U}$, then

$$\operatorname{rank}(\mathbf{K}_{\rm F}) \leq N_{\rm rf}^2 - N_{\rm rf} + 1 \tag{23}$$

Proof: See Appendix A. ■

Note that we can compute $\operatorname{rank}(\mathbf{K}_{\rm F})$ without the knowledge of $\mathbf{F}_{\rm opt}$ because its left singular vectors $\mathbf{U}_{\rm F}$ can always be chosen as the first $N_{\rm rf}$ columns of $\mathbf{V}_{\rm H}$, with $\mathbf{V}_{\rm H} \in \mathbb{C}^{N_{\rm t}\times N_{\rm t}}$ being the right singular vectors of $\mathbf{H}$ [12, Proposition 2]. Therefore, when the transmitter has perfect channel state information, it can construct $\mathbf{K}_{\rm F}$ and check whether $\operatorname{rank}(\mathbf{K}_{\rm F}) \leq N_{\rm rf}^2 - N_{\rm rf} + 1$ holds. If the necessary condition does not hold, then hybrid precoding schemes cannot realize unconstrained optimal precoders exactly.

When the sufficient condition in Proposition 2 does not hold, $\mathbf{K}_{\rm F}$ is usually a full rank matrix. In this case, we derive the minimum number of RF chains required for hybrid precoding to achieve the performance of unconstrained optimal precoders.

Corollary 1: When $\mathbf{K}_{\rm F}$ is a full rank matrix, it requires at least $\sqrt{N_{\rm t} - \frac{3}{4}} + \frac{1}{2}$ RF chains for hybrid precoding schemes to realize unconstrained optimal precoders exactly.

Proof: Since $\mathbf{K}_{\rm F}$ is a full rank matrix, $\operatorname{rank}(\mathbf{K}_{\rm F}) = \min\{N_{\rm t}, N_{\rm rf}^2\}$. Inserting $\operatorname{rank}(\mathbf{K}_{\rm F})$ into $\operatorname{rank}(\mathbf{K}_{\rm F}) \leq N_{\rm rf}^2 - N_{\rm rf} + 1$ and using quadratic formula, we obtain

$$N_{\rm rf} \geq \sqrt{N_{\rm t} - \frac{3}{4}} + \frac{1}{2}. \tag{24}$$

This completes the proof. ■

B. Structures of the Matrix Factorization Formulation

Given the unconstrained optimal precoder $\mathbf{F}_{\rm opt}$, the matrix factorization problem (10) belongs to the class of polynomial optimization: The objective function $\|\mathbf{F}_{\rm opt} - \mathbf{F}_{\rm RF}\mathbf{F}_{\rm BB}\|_F^2$ is a convex quartic function with respect to matrix variables $(\mathbf{F}_{\rm RF}, \mathbf{F}_{\rm BB})$, the power constraint $\operatorname{tr}\left(\mathbf{F}_{\rm BB}^H\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF}\mathbf{F}_{\rm BB}\right) \leq P$ is a convex quartic constraint, and the constant modulus constraints $\mathcal{U}$ are nonconvex quadratic equality constraints. Such a problem is nonconvex due to the nonconvexity of $\mathcal{U}$, and theoretical challenges of problem (10) are listed as follows:

1) The optimization variables $\mathbf{F}_{\rm RF}$ and $\mathbf{F}_{\rm BB}$ are coupled through the power constraint. Therefore, we cannot deploy the alternating minimization approach which requires separate variables in constraints. If we jointly optimize $(\mathbf{F}_{\rm RF}, \mathbf{F}_{\rm BB})$, the difficulty also lies in handing the coupled feasible region of problem (10).
2) More importantly, the bilinear mapping $(\mathbf{F}_{\rm RF}, \mathbf{F}_{\rm BB}) \mapsto \mathbf{F}_{\rm RF}\mathbf{F}_{\rm BB}$ is not a one-to-one mapping, thus $(\mathbf{F}_{\rm RF}, \mathbf{F}_{\rm BB})$ and $(\mathbf{F}_{\rm RF}\mathbf{\Sigma}, \mathbf{\Sigma}^{-1}\mathbf{F}_{\rm BB})$ result in the same objective value, where $\mathbf{\Sigma}$ is a diagonal matrix with unit modulus diagonal entries to ensure $\mathbf{F}_{\rm RF}\mathbf{\Sigma} \in \mathcal{U}$. In other words, we should expect problem (10) to have infinite number of local minima and saddle points.

The first issue is fully addressed by the following theorem, which shows the equivalence between problems (10) and the following relaxed problem:

$$\underset{\mathbf{F}_{\mathrm{RF}}\in\mathcal{U},\mathbf{F}_{\mathrm{BB}}}{\text{minimize}} \quad \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2. \tag{25}$$

Theorem 1: If $(\hat{\mathbf{F}}_{\mathrm{RF}}, \hat{\mathbf{F}}_{\mathrm{BB}})$ is a KKT point of problem (25), then it satisfies $\mathrm{tr}(\hat{\mathbf{F}}_{\mathrm{BB}}^H\hat{\mathbf{F}}_{\mathrm{RF}}^H\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}) \leq P$.

Proof: See Appendix A. ∎

According to Theorem 1, any KKT point of problem (25) satisfies $\mathrm{tr}(\mathbf{F}_{\mathrm{BB}}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}) \leq P$, thus the power constraint can be removed without loss of local and global optimality.

The rest of this paper focuses on solving problem (25). Problem (25) is a constant modulus matrix factorization problem where a given matrix $\mathbf{F}_{\mathrm{opt}}$ is factorized into two complex matrices ($\mathbf{F}_{\mathrm{RF}}$, $\mathbf{F}_{\mathrm{BB}}$) under constant modulus constraints on $\mathbf{F}_{\mathrm{RF}}$. Since $(\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}) \mapsto \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}$ is not a one-to-one mapping, problem (25) has infinite number of saddle points, and this issue will be addressed in Section IV.

IV. Constant Modulus Matrix Factorization

A. Problem Reformulation

First, we observe that for any given $\mathbf{F}_{\mathrm{RF}}$, problem (25) is a least square problem

$$\underset{\mathbf{F}_{\mathrm{BB}}}{\text{minimize}} \quad \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2. \tag{26}$$

Suppose that $\mathbf{F}_{\mathrm{RF}}$ has full column rank, then the optimal solution of problem (26) is

$$\mathbf{F}_{\mathrm{BB}} = \mathbf{F}_{\mathrm{RF}}^{+}\mathbf{F}_{\mathrm{opt}} \tag{27}$$

where $\mathbf{F}_{\mathrm{RF}}^{+} = (\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}})^{-1}\mathbf{F}_{\mathrm{RF}}^H$ is the Moore-Penrose pseudoinverse of $\mathbf{F}_{\mathrm{RF}}$. Inserting (27) into problem (25), $\mathbf{F}_{\mathrm{BB}}$ is eliminated and we obtain the modified problem:

$$\underset{\mathbf{F}_{\mathrm{RF}}\in\mathcal{U}}{\text{minimize}} \quad f(\mathbf{F}_{\mathrm{RF}}) = \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{+}\mathbf{F}_{\mathrm{opt}}\|_F^2. \tag{28}$$

The following theorem guarantees that problems (25) and (28) are equivalent.

Theorem 2: If $\hat{\mathbf{F}}_{\mathrm{RF}}$ is a KKT point of problem (28) and $\hat{\mathbf{F}}_{\mathrm{BB}} = \hat{\mathbf{F}}_{\mathrm{RF}}^{+}\mathbf{F}_{\mathrm{opt}}$, then $(\hat{\mathbf{F}}_{\mathrm{RF}}, \hat{\mathbf{F}}_{\mathrm{BB}})$ is a KKT point of problem (25). Furthermore, $\hat{\mathbf{F}}_{\mathrm{RF}}$ is a globally optimal solution of problem (28) if and only if $(\hat{\mathbf{F}}_{\mathrm{RF}}, \hat{\mathbf{F}}_{\mathrm{BB}})$ is a globally optimal solution of problem (25).

Proof: See Appendix B. ∎

The benefit of this reformulation is that problem (28) can be solved more efficiently because its search space is reduced from $(\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}})$ to $\mathbf{F}_{\mathrm{RF}}$.

Problem (28) involves minimizing a polynomial with non-convex constant modulus constraints, which is difficult to handle. Note that the constant modulus constraints imply that only the phase of $\mathbf{F}_{\mathrm{RF}}$ can be changed. Therefore, instead of using $\mathbf{F}_{\mathrm{RF}}$ as the optimization variable, it is more convenient to optimize the phase of $\mathbf{F}_{\mathrm{RF}}$ directly. Let the phase of $\mathbf{F}_{\mathrm{RF}}$ be $\mathbf{\Phi}_{\mathrm{RF}}$, i.e., $\mathbf{F}_{\mathrm{RF}} = \frac{1}{\sqrt{N_{\mathrm{t}}}}e^{j\mathbf{\Phi}_{\mathrm{RF}}}$. Using $\mathbf{\Phi}_{\mathrm{RF}}$ as the optimization variable and rewriting $\mathbf{F}_{\mathrm{RF}}$ as $\mathbf{F}_{\mathrm{RF}}(\mathbf{\Phi}_{\mathrm{RF}})$, we can reformulate problem (28) as the following unconstrained minimization problem

$$\underset{\mathbf{\Phi}_{\mathrm{RF}}}{\text{minimize}} \quad \psi(\mathbf{\Phi}_{\mathrm{RF}}) = \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}(\mathbf{\Phi}_{\mathrm{RF}})\mathbf{F}_{\mathrm{RF}}^{+}(\mathbf{\Phi}_{\mathrm{RF}})\mathbf{F}_{\mathrm{opt}}\|_F^2. \tag{29}$$

Although (29) is a unconstrained problem, it is still not recommended to solve this problem directly because the objective function $\psi(\mathbf{\Phi}_{\mathrm{RF}})$ is ill-behaved: First, $\psi(\mathbf{\Phi}_{\mathrm{RF}}) = \psi(\mathbf{\Phi}_{\mathrm{RF}}+\mathbf{S})$ for any rank one real matrix $\mathbf{S}$. Thus problem (29) has infinite number of local minima and saddle points; Second, the Hessian of $\psi(\mathbf{\Phi}_{\mathrm{RF}})$ at any point $\mathbf{\Phi}_{\mathrm{RF}}$ is a singular matrix. To show this, we expand $\psi(\mathbf{\Phi}_{\mathrm{RF}} + \mathbf{S})$ at $\mathbf{\Phi}_{\mathrm{RF}}$ using Taylor's theorem:

$$\begin{aligned}
&\psi(\mathbf{\Phi}_{\mathrm{RF}} + \mathbf{S}) \\
&\quad = \psi(\mathbf{\Phi}_{\mathrm{RF}}) + \mathrm{vec}\left[\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})\right]^T \mathrm{vec}(\mathbf{S}) \\
&\qquad + \frac{1}{2}\mathrm{vec}(\mathbf{S})^T\left[\nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}})\right]\mathrm{vec}(\mathbf{S}) + o(\|\mathrm{vec}(\mathbf{S})\|^2)
\end{aligned}$$

where $\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})$ and $\nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}})$ are the gradient and Hessian of $\psi(\mathbf{\Phi}_{\mathrm{RF}})$ respectively, and $o(\|\mathrm{vec}(\mathbf{S})\|^2)$ is the Peano's form of the reminder. For any nonzero rank one real matrix $\mathbf{S}$, we have $\psi(\mathbf{\Phi}_{\mathrm{RF}} + \mathbf{S}) = \psi(\mathbf{\Phi}_{\mathrm{RF}})$, which implies

$$\begin{aligned}
\mathrm{vec}(\mathbf{S}) &\neq \mathbf{0} \\
\mathrm{vec}\left[\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})\right]^T\mathrm{vec}(\mathbf{S}) &= 0 \\
\left[\nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}})\right]\mathrm{vec}(\mathbf{S}) &= \mathbf{0}.
\end{aligned} \tag{30}$$

Therefore, $\nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}})$ is a singular matrix.

We address these two issues by restricting the first row of $\mathbf{\Phi}_{\mathrm{RF}}$ being a zero vector. Note that $\mathbf{\Phi}_{\mathrm{RF}}$ can be partitioned into two blocks

$$\mathbf{\Phi}_{\mathrm{RF}} = \begin{bmatrix}\mathbf{r}\\ \mathbf{R}\end{bmatrix} \tag{31}$$

where $\mathbf{r} \in \mathbb{R}^{1\times N_{\mathrm{rf}}}$ is the first row of $\mathbf{\Phi}_{\mathrm{RF}}$, and $\mathbf{R} \in \mathbb{R}^{(N_{\mathrm{t}}-1)\times N_{\mathrm{rf}}}$ is the remaining part of $\mathbf{\Phi}_{\mathrm{RF}}$. If $\mathbf{r}$ is not a zero vector, we can always construct a unique matrix

$$\bar{\mathbf{\Phi}}_{\mathrm{RF}} = \mathbf{\Phi}_{\mathrm{RF}} - \mathbf{1r} = \begin{bmatrix}0\\ \mathbf{\Phi}\end{bmatrix} \tag{32}$$

such that the first row of $\bar{\mathbf{\Phi}}_{\mathrm{RF}}$ is a zero vector, and $\psi(\bar{\mathbf{\Phi}}_{\mathrm{RF}}) = \psi(\mathbf{\Phi}_{\mathrm{RF}})$. Therefore, we can optimize $\psi(\mathbf{\Phi}_{\mathrm{RF}})$ over a special class of $\mathbf{\Phi}_{\mathrm{RF}}$ satisfying

$$\mathbf{\Phi}_{\mathrm{RF}} = \begin{bmatrix}0\\ \mathbf{\Phi}\end{bmatrix} \tag{33}$$

where $\mathbf{\Phi} \in \mathbb{R}^{(N_{\mathrm{t}}-1)\times N_{\mathrm{rf}}}$. Using $\mathbf{\Phi}$ as the optimization variable, problem (29) is further reformulated as

$$\underset{\mathbf{\Phi}}{\text{minimize}}\ \varphi(\mathbf{\Phi}) = \psi\left\{\begin{bmatrix}0\\ \mathbf{\Phi}\end{bmatrix}\right\}. \tag{34}$$

B. Gradient and Hessian

In this subsection, we derive the gradient and Hessian of $\varphi(\mathbf{\Phi})$, which are the foundation for developing numerical algorithms to solve problem (34). Since the gradient and Hessian of $\varphi(\mathbf{\Phi})$ depend on those of $f(\mathbf{F}_{\text{RF}})$, we first provide the gradient and Hessian of $f(\mathbf{F}_{\text{RF}})$ in the following lemma.

Lemma 1: The complex gradient and Hessian matrices of $f(\mathbf{F}_{\text{RF}})$ are given by

$$\nabla_{\mathbf{F}_{\text{RF}}} f(\mathbf{F}_{\text{RF}}) \triangleq \frac{\partial f(\mathbf{F}_{\text{RF}})}{\partial \mathbf{F}_{\text{RF}}^*} = -\mathbf{Z}_1 \mathbf{F}_{\text{opt}} \mathbf{Z}_2^H \tag{35}$$

$$\begin{aligned}\mathcal{CH}_{\mathbf{F}_{\text{RF}}} f(\mathbf{F}_{\text{RF}}) &\triangleq \begin{bmatrix} \mathcal{H}_{\mathbf{F}_{\text{RF}},\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) & \mathcal{H}_{\mathbf{F}_{\text{RF}}^*,\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) \\ \mathcal{H}_{\mathbf{F}_{\text{RF}},\mathbf{F}_{\text{RF}}} f(\mathbf{F}_{\text{RF}}) & \mathcal{H}_{\mathbf{F}_{\text{RF}}^*,\mathbf{F}_{\text{RF}}} f(\mathbf{F}_{\text{RF}}) \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{H}_{\mathbf{F}_{\text{RF}},\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) & \mathcal{H}_{\mathbf{F}_{\text{RF}}^*,\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) \\ \mathcal{H}^*_{\mathbf{F}_{\text{RF}}^*,\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) & \mathcal{H}^*_{\mathbf{F}_{\text{RF}},\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) \end{bmatrix} \end{aligned} \tag{36}$$

where $\mathbf{Z}_1 = \mathbf{I} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{RF}}^+$, $\mathbf{Z}_2 = \mathbf{F}_{\text{RF}}^+\mathbf{F}_{\text{opt}}$, and

$$\begin{aligned}\mathcal{H}_{\mathbf{F}_{\text{RF}},\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) &= (\mathbf{Z}_2\mathbf{Z}_2^H)^T \otimes \mathbf{Z}_1 \\ &\quad - \left[(\mathbf{F}_{\text{RF}}^H\mathbf{F}_{\text{RF}})^{-1}\right]^T \otimes \mathbf{Z}_1\mathbf{F}_{\text{opt}}\mathbf{F}_{\text{opt}}^H\mathbf{Z}_1^H \\ \mathcal{H}_{\mathbf{F}_{\text{RF}}^*,\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}}) &= \left[(\mathbf{Z}_1\mathbf{F}_{\text{opt}}\mathbf{Z}_2^H)^T \otimes (\mathbf{F}_{\text{RF}}^+)^H\right]\mathbf{K}_{N_{\text{t}},N_{\text{rf}}} \\ &\quad + \mathbf{K}_{N_{\text{t}},N_{\text{rf}}}^T \left[(\mathbf{Z}_1\mathbf{F}_{\text{opt}}\mathbf{Z}_2^H)^T \otimes (\mathbf{F}_{\text{RF}}^+)^H\right]^T. \end{aligned}$$

Here $\mathbf{K}_{N_{\text{t}},N_{\text{rf}}}$ is the commutation matrix satisfying $\text{vec}(\text{d}\mathbf{F}_{\text{RF}}^T) = \mathbf{K}_{N_{\text{t}},N_{\text{rf}}}\text{vec}(\text{d}\mathbf{F}_{\text{RF}})$.

Proof: See Appendix B. ∎

With the help of Lemma 1, we can compute the gradient $\nabla\varphi(\mathbf{\Phi})$ and Hessian $\nabla^2\varphi(\mathbf{\Phi})$. For any given $\mathbf{\Phi}$, we construct the corresponding $\mathbf{\Phi}_{\text{RF}}$ in (33). Then $\nabla\varphi(\mathbf{\Phi})$ is obtained by deleting the first row of $\nabla\psi(\mathbf{\Phi}_{\text{RF}})$, and $\nabla^2\varphi(\mathbf{\Phi})$ is obtained by deleting the $(N_{\text{t}}\ell+1)$th rows and columns of $\nabla^2\psi(\mathbf{\Phi}_{\text{RF}})$, with $\ell = 0, 1, \ldots, N_{\text{rf}} - 1$. The gradient and Hessian of $\psi(\mathbf{\Phi}_{\text{RF}})$ are given in the following theorem.

Theorem 3: The gradient and Hessian matrices of $\psi(\mathbf{\Phi}_{\text{RF}})$ are given by

$$\nabla\psi(\mathbf{\Phi}_{\text{RF}}) = 2\Im\big[\mathbf{G}\big] \tag{37}$$

$$\nabla^2\psi(\mathbf{\Phi}_{\text{RF}}) = 2\Re\big[\mathbf{M}\big] - 2\text{diag}\big\{\text{vec}\big(\Re[\mathbf{G}]\big)\big\} \tag{38}$$

where $\mathbf{G} = \nabla_{\mathbf{F}_{\text{RF}}} f(\mathbf{F}_{\text{RF}}) \circ \mathbf{F}_{\text{RF}}^*$ and

$$\begin{aligned}\mathbf{M} &= [\mathcal{H}_{\mathbf{F}_{\text{RF}},\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}})] \circ \text{vec}(\mathbf{F}_{\text{RF}}^*)\text{vec}(\mathbf{F}_{\text{RF}})^T \\ &\quad - [\mathcal{H}_{\mathbf{F}_{\text{RF}}^*,\mathbf{F}_{\text{RF}}^*} f(\mathbf{F}_{\text{RF}})] \circ \text{vec}(\mathbf{F}_{\text{RF}}^*)\text{vec}(\mathbf{F}_{\text{RF}})^H. \end{aligned} \tag{39}$$

Proof: See Appendix B. ∎

C. BFGS-Based Algorithm

In this subsection, we propose a Broyden-Fletcher-Goldfarb-Shanno (BFGS)-based method to solve problem (34). The BFGS method is a well-known quasi-Newton algorithm for unconstrained optimization problems. It updates the current solution $\mathbf{\Phi}_n$ to $\mathbf{\Phi}_{n+1}$ by the following rule:

$$\mathbf{\Phi}_{n+1} = \mathbf{\Phi}_n + \rho_n\mathbf{S}_n \tag{40}$$

where $\mathbf{S}_n$ is the descent direction, and $\rho_n > 0$ is the stepsize. The descent direction $\mathbf{S}_n$ is given by

$$\text{vec}(\mathbf{S}_n) = -\mathbf{B}_n\text{vec}[\nabla\varphi(\mathbf{\Phi}_n)]. \tag{41}$$

Here $\mathbf{B}_n$ is a symmetric positive definite matrix which approximates the inverse of $\nabla^2\varphi(\mathbf{\Phi}_n)$. Note that the positive definiteness of $\mathbf{B}_n$ ensures that $\mathbf{S}_n$ is a descent direction, i.e.,

$$\text{tr}\big[\nabla\varphi(\mathbf{\Phi}_n)^T\mathbf{S}_n\big] = -\text{vec}[\nabla\varphi(\mathbf{\Phi}_n)]^T\mathbf{B}_n\text{vec}[\nabla\varphi(\mathbf{\Phi}_n)] < 0.$$

The matrix $\mathbf{B}_n$ is usually updated by the inverse BFGS formula

$$\mathbf{B}_{n+1} = \left(\mathbf{I} - \frac{\mathbf{s}_n\mathbf{y}_n^T}{\mathbf{y}_n^T\mathbf{s}_n}\right)\mathbf{B}_n\left(\mathbf{I} - \frac{\mathbf{s}_n\mathbf{y}_n^T}{\mathbf{y}_n^T\mathbf{s}_n}\right)^T + \frac{\mathbf{s}_n\mathbf{s}_n^T}{\mathbf{y}_n^T\mathbf{s}_n} \tag{42}$$

where $\mathbf{s}_n = \text{vec}[\mathbf{\Phi}_{n+1} - \mathbf{\Phi}_n]$ and $\mathbf{y}_n = \text{vec}[\nabla\varphi(\mathbf{\Phi}_{n+1}) - \nabla\varphi(\mathbf{\Phi}_n)]$. Clearly, $\mathbf{B}_{n+1}$ will inherit the positive definiteness of $\mathbf{B}_n$ as long as $\mathbf{y}_n^T\mathbf{s}_n > 0$. However, the condition $\mathbf{y}_n^T\mathbf{s}_n > 0$ does not hold for general nonconvex problems. In order to ensure the positive definiteness of $\mathbf{B}_{n+1}$, a cautious update rule for $\mathbf{B}_n$ is proposed [20]

$$\mathbf{B}_{n+1} = \begin{cases} (42) \quad \text{if } \dfrac{\mathbf{y}_n^T\mathbf{s}_n}{\|\mathbf{s}_n\|^2\|\nabla\varphi(\mathbf{\Phi}_n)\|_F} > \eta_{\text{bfgs}} \\ \mathbf{B}_n \quad \text{otherwise} \end{cases} \tag{43}$$

where $\eta_{\text{bfgs}} = 10^{-6}$ is a small constant. The update rule in (43) guarantees that $\mathbf{B}_n$ is a positive definite matrix in each iteration, and thus $\mathbf{S}_n$ should be a descent direction. However, due to the roundoff error, sometimes the direction generated by (41) may be not a descent direction. To address this numerical issue, we choose $\mathbf{S}_n$ as

$$\text{vec}(\mathbf{S}_n) = \begin{cases} -\mathbf{B}_n\text{vec}[\nabla\varphi(\mathbf{\Phi}_n)] & \text{if } \xi_n > \delta_{\text{bfgs}} \\ -\text{vec}[\nabla\varphi(\mathbf{\Phi}_n)] & \text{otherwise} \end{cases} \tag{44}$$

where $\xi_n = \text{vec}[\nabla\varphi(\mathbf{\Phi}_n)]^T\mathbf{B}_n\text{vec}[\nabla\varphi(\mathbf{\Phi}_n)]$ and $\delta_{\text{bfgs}} = 10^{-6}$ is a small constant.

After obtaining the descent direction $\mathbf{S}_n$, we need to determine the stepsize ρ_n such that the objective function is decreasing in each iteration. We propose a modified backtracking line search method, which is usually more efficient than the classic backtracking line search [21]. The main idea is to use ρ_{n-1} as the initial guess of ρ_n, and then either increases or decreases it to find the largest $\rho_n \in \mathcal{G}_n$ such that

$$\mathcal{G}_n = \left\{\rho \geq 0 \;\middle|\; \begin{aligned} &\varphi\big(\mathbf{\Phi}_n + \rho\mathbf{S}_n\big) \leq \varphi(\mathbf{\Phi}_n) + \\ &\rho\cdot\beta_{\text{bfgs}}\text{tr}\big[\nabla\varphi(\mathbf{\Phi}_n)^T\mathbf{S}_n\big] \end{aligned}\right\} \tag{45}$$

where $\beta_{\text{bfgs}} \in [0, 0.5]$ is a constant to control the stepsize. Specifically, the stepsize ρ_n is set as

$$\rho_n = \begin{cases} 2^{K_1-1}\cdot\rho_{n-1} & \text{if } \rho_{n-1} \in \mathcal{G}_n \\ \left(\dfrac{1}{2}\right)^{K_2}\cdot\rho_{n-1} & \text{if } \rho_{n-1} \notin \mathcal{G}_n \end{cases} \tag{46}$$

where $K_1 \geq 0$ is the smallest integer such that $2^{K_1}\cdot\rho_{n-1} \notin \mathcal{G}_n$, and $K_2 \geq 0$ is the smallest integer such that $(\frac{1}{2})^{K_2}\cdot\rho_{n-1} \in \mathcal{G}_n$.The details of our BFGS-based algorithm is summarized in Algorithm 1.

According to [20], the BFGS-based algorithm proposed in Algorithm 1 can converge to a stationary point of problem (29), i.e., the limit of $\nabla\varphi(\mathbf{\Phi}_n)$ satisfies

$$\lim_{n\to\infty}\|\nabla\varphi(\mathbf{\Phi}_n)\|_F = 0. \tag{47}$$

Algorithm 1 BFGS-Based Algorithm

1. Inputs: $\mathbf{F}_{\text{opt}}$, $\mathbf{\Phi}_0$ and $\mathbf{B}_0$. Set $\rho_0 = 1$, $\beta_{\text{bfgs}} = 0.5$, and $\epsilon = 10^{-4}$.
2. For $n = 0, 1, 2, \ldots$ (outer iterations)
   - Determine the descent direction $\mathbf{S}_n$ by (44).
   - Compute the stepsize ρ_n via (46).
   - Update $\mathbf{\Phi}_n$ to $\mathbf{\Phi}_{n+1}$ according to (40).
   - If $\min\left\{\left|\frac{\varphi(\mathbf{\Phi}_{n+1})-\varphi(\mathbf{\Phi}_n)}{\varphi(\mathbf{\Phi}_{n+1})}\right|, \|\nabla\varphi(\mathbf{\Phi}_{n+1})\|_F\right\} < \epsilon$, stop.
   - Update $\mathbf{B}_n$ to $\mathbf{B}_{n+1}$ by (43).
3. Outputs: $\mathbf{F}_{\text{RF}} = \frac{1}{\sqrt{N_{\text{T}}}} e^{j[\mathbf{0};\mathbf{\Phi}_n]}$, $\mathbf{F}_{\text{BB}} = \mathbf{F}_{\text{RF}}^{+}\mathbf{F}_{\text{opt}}$.

The performance and convergence speed of Algorithm 1 depends on $\mathbf{\Phi}_0$ and $\mathbf{B}_0$. Here a good choice for the initial analog precoder phase is $\angle\mathbf{U}_{\text{F}}$, where $\angle\mathbf{U}_{\text{F}}$ is the phase of $\mathbf{U}_{\text{F}}$. Then the corresponding $\mathbf{\Phi}_0$ is set as

$$\mathbf{\Phi}_0 = \left[\angle\mathbf{U}_{\text{F}}\right]_{2:N_{\text{T}},\bullet} - \mathbf{1}\left[\angle\mathbf{U}_{\text{F}}\right]_{1,\bullet}. \tag{48}$$

The initial inverse Hessian approximation $\mathbf{B}_0$ will greatly affect the efficiency of Algorithm 1, thus we need to design it carefully. Let the eigendecomposition of $\nabla^2\varphi(\mathbf{\Phi}_0)$ be

$$\nabla^2\varphi(\mathbf{\Phi}_0) = \mathbf{U}_0\mathbf{\Sigma}_0\mathbf{U}_0^T \tag{49}$$

where $\mathbf{U}_0 \in \mathbb{C}^{(N_{\text{t}}-1)N_{\text{rf}}\times(N_{\text{t}}-1)N_{\text{rf}}}$ is a unitary matrix, and $\mathbf{\Sigma}_0 \in \mathbb{R}^{(N_{\text{t}}-1)N_{\text{rf}}\times(N_{\text{t}}-1)N_{\text{rf}}}$ is a diagonal matrix with eigenvalues arranged in decreasing order. Then $\mathbf{B}_0$ is given by

$$\mathbf{B}_0 = \mathbf{U}_0\hat{\mathbf{\Sigma}}_0^{-1}\mathbf{U}_0^T \tag{50}$$

where $\hat{\mathbf{\Sigma}}_0$ is a diagonal matrix with the k-th diagonal entry being

$$[\hat{\mathbf{\Sigma}}_0]_{k,k} = \begin{cases} \left|[\mathbf{\Sigma}_0]_{k,k}\right| & \text{if } \left|[\mathbf{\Sigma}_0]_{k,k}\right| \geq \delta_{\min} \\ \delta_{\min} & \text{otherwise.} \end{cases} \tag{51}$$

Here the small constant $\delta_{\min}$ is set as $\delta_{\min} = 10^{-4}$. Since $\hat{\mathbf{\Sigma}}_0^{-1}$ is a diagonal matrix with positive diagonal entries, the positive definiteness condition of $\mathbf{B}_0$ is satisfied.

D. Complexity Analysis

In this subsection, we discuss the per-iteration complexity of the proposed BFGS-based algorithm. Typically, the most time consuming operation in Algorithm 1 is evaluating $\varphi(\mathbf{\Phi})$ and $\nabla\varphi(\mathbf{\Phi})$. Therefore, it is important to analyze the complexity for $\varphi(\mathbf{\Phi})$ and $\nabla\varphi(\mathbf{\Phi})$. Given $\mathbf{\Phi}$, we construct the corresponding analog precoder phase $\mathbf{\Phi}_{\text{RF}}$ satisfying (33) and the analog precoder $\mathbf{F}_{\text{RF}} = \frac{1}{\sqrt{N_{\text{t}}}}e^{j\mathbf{\Phi}_{\text{RF}}}$. Then we decompose $\mathbf{F}_{\text{RF}}$ by QR decomposition

$$\mathbf{F}_{\text{RF}} = \mathbf{Q}_{\text{RF}}\mathbf{R}_{\text{RF}} \tag{52}$$

where $\mathbf{Q}_{\text{RF}} \in \mathbb{C}^{N_{\text{t}}\times N_{\text{rf}}}$ is a unitary matrix, and $\mathbf{R}_{\text{RF}} \in \mathbb{C}^{N_{\text{rf}}\times N_{\text{rf}}}$ is an invertible upper triangle matrix. In this way, we can compute $\varphi(\mathbf{\Phi})$ efficiently as

$$\varphi(\mathbf{\Phi}) = \|\mathbf{F}_{\text{opt}}\|_F^2 - \|\mathbf{Q}_{\text{RF}}^H\mathbf{F}_{\text{opt}}\|_F^2. \tag{53}$$

The QR decomposition requires $\mathcal{O}(N_{\text{t}}N_{\text{rf}}^2)$ flops, and computing $\|\mathbf{Q}_{\text{RF}}^H\mathbf{F}_{\text{opt}}\|_F^2$ requires $\mathcal{O}(N_{\text{t}}N_{\text{rf}}N_{\text{s}})$ flops. Therefore, the complexity for computing $\varphi(\mathbf{\Phi})$ is about $\mathcal{O}(N_{\text{t}}N_{\text{rf}}^2 + N_{\text{t}}N_{\text{rf}}N_{\text{s}})$.

The gradient matrix $\nabla\varphi(\mathbf{\Phi})$ can be expressed as

$$\nabla\varphi(\mathbf{\Phi}) = \left[\nabla\psi(\mathbf{\Phi}_{\text{RF}})\right]_{2:N_{\text{t}},\bullet} \tag{54}$$

where $\nabla\psi(\mathbf{\Phi}_{\text{RF}})$ can be expressed using QR decomposition

$$\nabla\psi(\mathbf{\Phi}_{\text{RF}}) = 2\Im\left[(\mathbf{Q}_{\text{RF}}\mathbf{Z}_{\text{RF}} - \mathbf{F}_{\text{opt}})\mathbf{Z}_{\text{RF}}^H(\mathbf{R}_{\text{RF}}^{-1})^H \circ \mathbf{F}_{\text{RF}}^*\right]. \tag{55}$$

Here $\mathbf{Z}_{\text{RF}} = \mathbf{Q}_{\text{RF}}^H\mathbf{F}_{\text{opt}}$. Then the complexity for computing $\nabla\varphi(\mathbf{\Phi})$ is about $\mathcal{O}(N_{\text{t}}N_{\text{rf}}^2 + N_{\text{rf}}^3 + N_{\text{s}}N_{\text{rf}}^2 + N_{\text{t}}N_{\text{rf}}N_{\text{s}})$.

Finally, since $\mathbf{B}_n \in \mathbb{R}^{(N_{\text{t}}-1)N_{\text{rf}}\times(N_{\text{t}}-1)N_{\text{rf}}}$, the updating rule in (43) requires $\mathcal{O}([N_{\text{t}}-1]^2N_{\text{rf}}^2)$ flops. Then the per-iteration complexity of Algorithm 1 is given by

$$\mathcal{O}\big(N_{\text{t}}N_{\text{rf}}^2 + N_{\text{rf}}^3 + N_{\text{s}}N_{\text{rf}}^2 + N_{\text{t}}N_{\text{rf}}N_{\text{s}} + [N_{\text{t}}-1]^2N_{\text{rf}}^2\big). \tag{56}$$

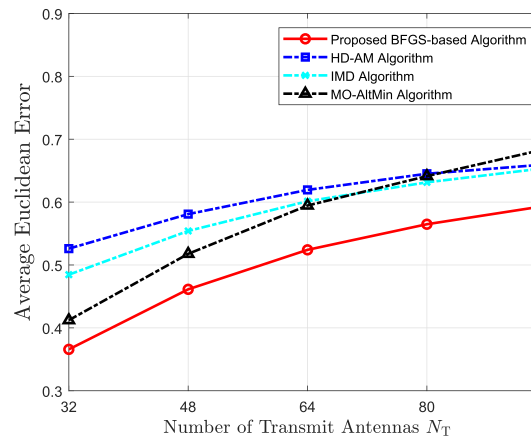


Fig. 1. Average Euclidean error versus N_{t} with 500 randomly generated full rank $\mathbf{F}_{\text{opt}}$.

V. Simulation Results

A. Average Euclidean Error Evaluation

The proposed BFGS-based algorithm solves a general constant modulus matrix factorization problem

$$\underset{\mathbf{F}_{\text{RF}}\in\mathcal{U},\mathbf{F}_{\text{BB}}}{\text{minimize}} \ \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2. \tag{57}$$

Therefore, it is of interest to evaluate the performance of our proposed algorithm for arbitrary given matrix $\mathbf{F}_{\text{opt}}$.

We generate N independent samples $\mathbf{F}_{\text{opt}}^{(i)} \in \mathbb{C}^{N_{\text{t}}\times N_{\text{s}}}$, $i = 1, 2, \ldots, N$ with i.i.d. zero-mean unit-variance complex Gaussian entries. Each sample is then normalized to satisfy

$$\left\|\mathbf{F}_{\text{opt}}^{(i)}\right\|_F^2 = N_{\text{s}}, \quad i = 1, 2, \ldots, N. \tag{58}$$

Subsequently, we evaluate the performance of our proposed algorithm by the average Euclidean error, given by

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{F}_{\text{opt}}^{(i)} - \mathbf{F}_{\text{RF}}^{(i)}\mathbf{F}_{\text{BB}}^{(i)}\right\|_F^2 \tag{59}$$

where $\mathbf{F}_{\text{RF}}^{(i)} \in \mathbb{C}^{N_{\text{t}}\times N_{\text{rf}}}$ and $\mathbf{F}_{\text{BB}}^{(i)} \in \mathbb{C}^{N_{\text{rf}}\times N_{\text{s}}}$ are outputs of Algorithm 1 with the given input $\mathbf{F}_{\text{opt}}^{(i)}$.

TABLE I

AVERAGE RUNNING TIME (IN SECS.) VERSUS N_{T} WITH 500 RANDOMLY GENERATED FULL RANK $\mathbf{F}_{\mathrm{opt}}$

N_{T}	32	48	64	80	96
Proposed BFGS-based algorithm	0.014s	0.021s	0.034s	0.033s	0.149s
HD–AM algorithm	0.008s	0.009s	0.014s	0.017s	0.022s
IMD algorithm	0.012s	0.013s	0.019s	0.020s	0.022s
MO–AltMin algorithm	0.349s	0.696s	1.226s	1.924s	5.429s

TABLE II

AVERAGE MUTUAL INFORMATION WITH GAUSSIAN INPUTS VERSUS SNR FOR VARIOUS ALGORITHMS

SNR(dB)	-35	-30	-25	-20	-15	-10	-5
WF algorithm (benchmark)	0.0767	0.2276	0.6169	1.4544	3.0283	5.5588	9.2619
Proposed BFGS-based algorithm	0.0763	0.2265	0.6141	1.4515	3.0175	5.4993	9.1137
HD–AM algorithm	0.0696	0.2080	0.5685	1.3614	2.8641	5.3070	8.9276
IMD algorithm	0.0698	0.2086	0.5702	1.3649	2.8707	5.3164	8.9382
MO–AltMin algorithm	0.0743	0.2189	0.6033	1.4375	2.9878	5.4584	9.0379

We make head-to-head comparisons between our proposed BFGS-based algorithm and three existing algorithms, namely the manifold optimization based alternating minimization (MO–AltMin) [3], the iterative matrix decomposition (IMD) [7] and the hybrid design by alternating minimization (HD–AM) [5]. To the best of our knowledge, these three algorithms are the best existing algorithms based on the matrix factorization approach. Note that the authors in [3] and [7] claim that their proposed algorithms have significant performance gains over other existing algorithms, and the authors in [5] claim that the HD–AM algorithm provides the best solution among four different hybrid precoding algorithms proposed in [5]. Therefore, if the proposed BFGS-based algorithm can beat these algorithms, we believe it outperforms other existing algorithms based on the matrix factorization approach.

The matrix factorization based algorithms [3], [5], [7] involve a normalization procedure to ensure $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = P$. Since the mutual information is monotonically increasing with respect to $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F$, this procedure will increase the achievable rate. However, when we choose the Euclidean error as the performance metric, the normalization procedure will decrease the overall performance because these algorithms and our proposed BFGS-based algorithm are designed to solve problem (57) without the equality power constraint. Therefore, for the sake of fairness, we do not execute the normalization for all algorithms in this subsection.

We set the number of samples as $N = 500$, and N_{rf} and N_{s} are restricted to be $N_{\mathrm{rf}} = N_{\mathrm{s}} = 4$. The initial analog precoders for these four algorithms are set as $\angle\mathbf{F}_{\mathrm{opt}}^{(i)}$. The average Euclidean error and average running time of four algorithms are presented in Fig. 1 and Table I. From Fig. 1 and Table I, we have the following remarks:

1) The proposed BFGS-based algorithm and the MO-AltMin algorithm are guaranteed to converge to the stationary point of problem (57), while the HD–AM and IMD algorithms may not achieve this goal.
2) The proposed BFGS-based algorithm significantly outperforms the HD–AM, IMD and MO–AltMin algorithms in the whole range of N_{T}. In addition, it consumes much lower computational time than the MO–AltMin algorithm.
3) The phenomenon that the BFGS-based algorithm outperforms the MO-AltMin algorithm can be explained as follows. For nonconvex problem (57), its stationary points can be local minimum (positive definite Hessian), local maximum (negative definite Hessian), or saddle point (indefinite Hessian). Most stationary points are saddle points in high dimensional space, and the objective value at the saddle point is usually worse than that at the local optimum [22]. In order to decrease the possibility for converging to the saddle point, we can 1) decrease the dimensions of the search space; 2) use Hessian information to avoid converging to the indefinite Hessian point [22], [23]. Since the proposed BFGS-based algorithm utilize these two techniques to avoid saddle points, its performance is better than that of the MO-AltMin algorithm.

B. Average Mutual Information Evaluation With Gaussian Inputs

We consider a 4×72 MIMO system with $N_{\mathrm{rf}} = N_{\mathrm{s}} = 4$. The number of physical propagation paths is set as $L = 8$, and the signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} = \frac{P}{\sigma^2}$. We generate $N = 1000$ channel realizations by (3), and evaluate the system performance by the following average mutual information with Gaussian inputs:

$$\frac{1}{N}\sum_{i=1}^{N} \log\det\left[\mathbf{I} + \sigma^{-2}\mathbf{H}_i\mathbf{Q}_i\mathbf{H}_i^H\right] \tag{60}$$

where $\mathbf{H}_i$ is the ith channel realization, and $\mathbf{Q}_i = \mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}(\mathbf{F}_{\mathrm{BB}}^{(i)})^H(\mathbf{F}_{\mathrm{RF}}^{(i)})^H$ with $(\mathbf{F}_{\mathrm{RF}}^{(i)}, \mathbf{F}_{\mathrm{BB}}^{(i)})$ being the analog and digital precoder solution corresponding to $\mathbf{H}_i$.

We set the performance of unconstrained optimal precoder as a benchmark, and then compare our proposed BFGS-based algorithm with the IMD algorithm, the HD–AM algorithm and the MO–AltMin algorithm. The unconstrained optimal precoder $\mathbf{F}_{\mathrm{opt}}$ under Gaussian inputs can be obtained by the waterfilling (WF) algorithm, and all hybrid precoding algorithms in this subsection use the same $\mathbf{F}_{\mathrm{opt}}$ to design analog and digital precoders. Moreover, the initial analog precoders of these algorithms are set as $\mathbf{F}_{\mathrm{RF}} = \frac{1}{\sqrt{N_{\mathrm{t}}}}e^{j[\mathbf{V}_{\mathrm{H}}]_{\bullet,1:N_{\mathrm{rf}}}}$, where $[\mathbf{V}_{\mathrm{H}}]_{\bullet,1:N_{\mathrm{rf}}}$ is the first N_{rf} right singular vectors of $\mathbf{H}$.

Table II demonstrates the average mutual information with Gaussian inputs versus SNR for various algorithms. From Table II, we have the following remarks:

1) The proposed BFGS-based algorithm has about 10% performance gain over HD–AM and IMD algorithms in low SNR regimes because HD–AM and IMD algorithms are designed for full rank $\mathbf{F}_{\rm opt}$. However, the unconstrained optimal precoder $\mathbf{F}_{\rm opt}$ is not a full rank matrix in low SNR regimes. In addition, the HD–AM and IMD algorithms can be applied only when $N_{\rm rf} = N_{\rm s}$, while our proposed BFGS-based algorithm and the MO-AltMin algorithm can work for arbitrary $N_{\rm rf}$ and $N_{\rm s}$.
2) When the unconstrained optimal precoder is obtained by WF algorithm and the performance metric is chosen as the average mutual information, the gain of our proposed BFGS-based algorithm over the MO-AltMin algorithm is not very significant compared with Fig. 1. However, as shown in Table I, our proposed BFGS-based algorithm is much faster than the MO-AltMin algorithm. Therefore, our proposed BFGS-based algorithm also has advantages over the MO-AltMin algorithm.

C. Average Mutual Information Evaluation With Finite-Alphabet Inputs

We first consider a 64×64 MIMO system with $N_{\rm rf} = N_{\rm s} = 4$. The number of physical propagation paths is set as $L = 6$. The input signal is drawn from QPSK modulation, and SNR is defined as $\text{SNR} = \frac{P}{\sigma^2}$. In addition, the system performance is measured by the average mutual information, which is averaged over 1000 channel realizations generated by (3).

We set the unconstrained optimal precoder under finite-alphabet inputs as a benchmark, and then compare our proposed BFGS-based algorithm with the gradient ascent algorithm [6], the classic waterfilling (WF) algorithm, the HD–AM algorithm [5] and the MO–AltMin algorithm [3]. For fair comparisons, the initial analog precoders of these algorithms are set as $\mathbf{F}_{\rm RF} = \frac{1}{\sqrt{N_{\rm t}}} e^{j[\mathbf{V}_{\rm H}]_{\bullet,1:N_{\rm rf}}}$.

Among these algorithms, our proposed BFGS-based algorithm and the gradient ascent algorithm are designed for finite-alphabet inputs, and the remaining three algorithms are designed under Gaussian inputs. Specifically, the HD–AM and MO–AltMin algorithms decompose the WF optimal precoder into digital and analog precoders, and then evaluate the corresponding mutual information under finite-alphabet inputs.

Fig. 2 demonstrates the average mutual information versus SNR for different algorithms. The results in Fig. 2 imply three observations. First, our proposed BFGS-based algorithm has the potential to achieve the performance of unconstrained optimal precoders. Second, our algorithm has about 0.2 bps/Hz improvement compared to the gradient ascent algorithm. Since mm-wave provide very large bandwidths, a gain of 0.2 bps/Hz would translate to a large increase in the effective data rate. Third, the proposed BFGS-based algorithm has about 3dB gain over the HD–AM and MO–AltMin algorithms. This is mainly because the unconstrained optimal precoder designed under Gaussian inputs will lead to significant performance loss when applying to finite-alphabet systems.

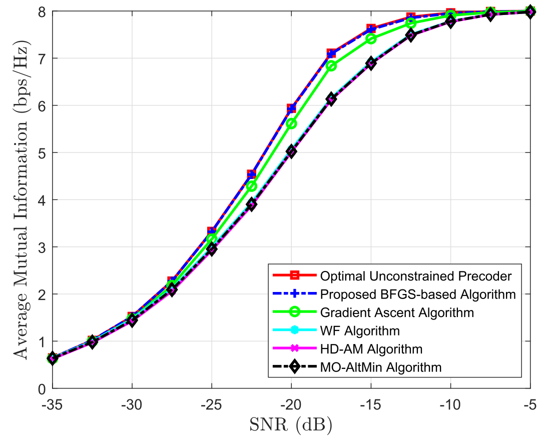


Fig. 2. Average mutual information versus SNR for different algorithms in a 64×64 system with $N_{\rm rf} = N_{\rm s} = 4$.

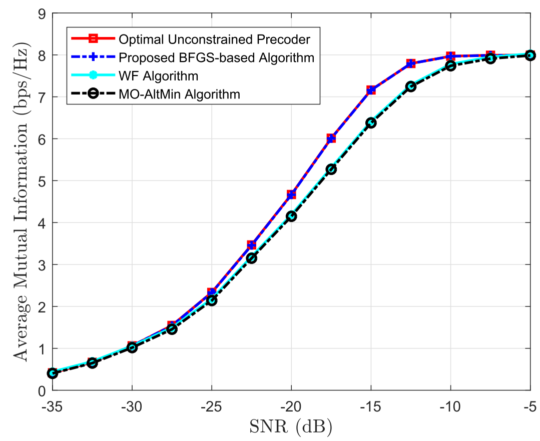


Fig. 3. Average mutual information versus SNR for different methods in a 32×80 system with $N_{\rm rf} = 6$ and $N_{\rm s} = 4$.

Next, we consider a 32×80 MIMO system with $L = 8$, $N_{\rm rf} = 6$ and $N_{\rm s} = 4$. The input signal is drawn from QPSK modulation. In this case, the gradient ascent and HD–AM algorithms cannot work because they assume $N_{\rm s} = N_{\rm rf}$. Therefore, we only compare our proposed BFGS-based algorithm with the MO-AltMin Algorithm. The simulation result is shown in Fig. 3. Based on the results in Fig. 3, we have the following remarks:

- The proposed BFGS-based algorithm and the MO–AltMin Algorithm are more general than the gradient ascent and HD–AM algorithms because they can work when $N_{\rm s} < N_{\rm rf}$.
- Our proposed algorithm can achieve the performance of unconstrained optimal precoder in whole SNR regimes. In addition, the MO–AltMin algorithm with WF optimal precoder has about 2–3dB performance loss compared with the our proposed BFGS-based algorithm.

VI. Conclusion

This paper considers the hybrid precoding design for mm-wave MIMO systems with finite-alphabet inputs. The precoding problem has been formulated as a matrix factorization problem with constant modulus constraints. We first proposed a sufficient and a necessary condition for the hybrid precoding scheme to achieve the performance of unconstrained optimal precoders. Next, we decoupled the constant modulus matrix factorization problem by showing that the power constraint can be removed without loss of local and/or global optimality. Then we proposed a BFGS-based method to solve the constant modulus matrix factorization problem. Numerical results have demonstrated the effectiveness of our proposed algorithm for hybrid precoding designs in mm-wave MIMO systems.

Appendix A
Proofs of Propositions 1–3 and Theorem 1

Proof of Proposition 1: If there exists a full rank square matrix $\mathbf{S}$ such that $\mathbf{U}_{\mathrm{F}}\mathbf{S} \in \mathcal{U}$, we can construct $\mathbf{F}_{\mathrm{RF}}$ and $\mathbf{F}_{\mathrm{BB}}$ as

$$\mathbf{F}_{\mathrm{RF}} = \mathbf{U}_{\mathrm{F}}\mathbf{S}, \quad \mathbf{F}_{\mathrm{BB}} = \mathbf{S}^{-1}\mathbf{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{H} \tag{61}$$

where $\mathbf{\Sigma}_{\mathrm{F}}$ is a diagonal matrix with singular values of $\mathbf{F}_{\mathrm{opt}}$ arranged in decreasing order, and $\mathbf{V}_{\mathrm{F}}$ is a unitary matrix with right singular vectors of $\mathbf{F}_{\mathrm{opt}}$. Then

$$\mathbf{F}_{\mathrm{RF}} \in \mathcal{U}, \quad \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}} = \mathbf{U}_{\mathrm{F}}\mathbf{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{H} = \mathbf{F}_{\mathrm{opt}}. \tag{62}$$

Conversely, if there exists $(\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}})$ such that $\mathbf{F}_{\mathrm{opt}} = \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}$, $\mathcal{C}(\mathbf{F}_{\mathrm{opt}})$ is a subspace of $\mathcal{C}(\mathbf{F}_{\mathrm{RF}})$, where $\mathcal{C}(\cdot)$ represents the space spanned by columns of a matrix. Moreover, according to $\mathbf{F}_{\mathrm{opt}} = \mathbf{U}_{\mathrm{F}}\mathbf{\Sigma}_{\mathrm{F}}\mathbf{V}_{\mathrm{F}}^{H}$, the first $\mathrm{rank}(\mathbf{F}_{\mathrm{opt}})$ columns of $\mathbf{U}_{\mathrm{F}}$ form an orthogonal basis of $\mathcal{C}(\mathbf{F}_{\mathrm{opt}})$. Since $\mathcal{C}(\mathbf{F}_{\mathrm{opt}})$ is a subspace of $\mathcal{C}(\mathbf{F}_{\mathrm{RF}})$, we can use the Gram-Schmidt algorithm to construct the remaining $N_{\mathrm{rf}} - \mathrm{rank}(\mathbf{F}_{\mathrm{opt}})$ columns of $\mathbf{U}_{\mathrm{F}}$ such that the columns of $\mathbf{U}_{\mathrm{F}}$ form an orthogonal basis of $\mathcal{C}(\mathbf{F}_{\mathrm{RF}})$. Then there exists a full rank matrix $\mathbf{S}$ satisfying $\mathbf{F}_{\mathrm{RF}} = \mathbf{U}_{\mathrm{F}}\mathbf{S} \in \mathcal{U}$. This completes the proof. ■

Proof of Proposition 2: Let the SVD of $\mathbf{H}$ be

$$\mathbf{H} = \mathbf{U}_{\mathrm{H}}\mathbf{\Sigma}_{\mathrm{H}}\mathbf{V}_{\mathrm{H}}^{H} \tag{63}$$

where $\mathbf{U}_{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{r}} \times \mathrm{rank}(\mathbf{H})}$ is a unitary matrix with left singular vectors, $\mathbf{\Sigma}_{\mathrm{H}} \in \mathbb{C}^{\mathrm{rank}(\mathbf{H}) \times \mathrm{rank}(\mathbf{H})}$ is a diagonal matrix with singular values arranged in decreasing order, and $\mathbf{V}_{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{t}} \times \mathrm{rank}(\mathbf{H})}$ is a unitary matrix with right singular vectors. Based on equation (8), when $L \le \min(N_{\mathrm{r}}, N_{\mathrm{t}})$, $\mathrm{rank}(\mathbf{H}) = L$. Then the columns of $\mathbf{V}_{\mathrm{H}}$ form an orthogonal basis of $\mathcal{C}(\mathbf{H}^{H})$. Moreover, since $\mathbf{H} = \mathbf{A}_{\mathrm{r}}\mathrm{diag}(\boldsymbol{\alpha})\mathbf{A}_{\mathrm{t}}^{H}$ and $\mathrm{rank}(\mathbf{A}_{\mathrm{t}}) = L$, the columns of $\mathbf{A}_{\mathrm{t}}$ also form a basis of $\mathcal{C}(\mathbf{H}^{H})$. Therefore, there exists a full rank square matrix $\mathbf{S} \in \mathbb{C}^{L \times L}$ such that $\mathbf{A}_{\mathrm{t}} = \mathbf{V}_{\mathrm{H}}\mathbf{S} \in \mathcal{U}$. The semi-unitary matrix $\mathbf{V}_{\mathrm{H}}$ has a close connection with the left singular vectors of $\mathbf{F}_{\mathrm{opt}}$. Specifically, the left singular vectors of $\mathbf{F}_{\mathrm{opt}}$ can always be chosen as the first N_{s} columns of $\mathbf{V}_{\mathrm{H}}$ [12, Proposition 2], i.e.,

$$\mathbf{U}_{\mathrm{F}} = \left[\mathbf{V}_{\mathrm{H}}\right]_{\bullet, 1:N_{\mathrm{s}}}. \tag{64}$$

Therefore, when $L = N_{\mathrm{s}} = \min\{L, N_{\mathrm{rf}}\} \le \min\{N_{\mathrm{r}}, N_{\mathrm{t}}\}$, we have $\mathbf{A}_{\mathrm{t}} = \tilde{\mathbf{V}}_{\mathrm{H}}\mathbf{S} = \mathbf{U}_{\mathrm{F}}\mathbf{S} \in \mathcal{U}$. Finally, $L = \min\{L, N_{\mathrm{rf}}\} \le \min\{N_{\mathrm{r}}, N_{\mathrm{t}}\}$ holds if and only if $L \le \min\{N_{\mathrm{r}}, N_{\mathrm{t}}, N_{\mathrm{rf}}\}$. This completes the proof. ■

Proof of Proposition 3: We first rewrite the solutions of $\mathbf{K}_{\mathrm{F}}\mathrm{vec}(\mathbf{Z}) = \mathbf{1}$ as

$$\mathrm{vec}(\mathbf{Z}) = \boldsymbol{\xi}_0 + \sum_{i=1}^{I} \alpha_i \boldsymbol{\xi}_i. \tag{65}$$

Here $\boldsymbol{\xi}_0$ is a particular solution to $\mathbf{K}_{\mathrm{F}}\mathrm{vec}(\mathbf{Z}) = \mathbf{1}$, $\{\alpha_i\}_{i=1}^{I}$ are complex numbers, and $\{\boldsymbol{\xi}_i\}_{i=1}^{I}$ is a basis of $\mathcal{N}(\mathbf{K}_{\mathrm{F}})$, where $\mathcal{N}(\cdot)$ represents the null space of a matrix. Since the nonliear equations

$$\mathbf{K}_{\mathrm{F}}\mathrm{vec}(\mathbf{Z}) = \mathbf{1}, \quad \mathbf{Z} \succeq 0, \ \mathrm{rank}(\mathbf{Z}) = 1 \tag{66}$$

have N_{RF} linear independent solutions, the dimension of $\mathcal{N}(\mathbf{K}_{\mathrm{F}})$ should be at least $N_{\mathrm{RF}} - 1$, which implies

$$\dim\left[\mathcal{N}(\mathbf{K}_{\mathrm{F}})\right] = N_{\mathrm{RF}}^2 - \mathrm{rank}(\mathbf{K}_{\mathrm{F}}) \ge N_{\mathrm{RF}} - 1. \tag{67}$$

This completes the proof. ■

Proof of Theorem 1: If $(\hat{\mathbf{F}}_{\mathrm{RF}}, \hat{\mathbf{F}}_{\mathrm{BB}})$ is a KKT point of problem (25), then it satisfies the following KKT conditions:

$$-(\mathbf{F}_{\mathrm{opt}} - \hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}})\hat{\mathbf{F}}_{\mathrm{BB}}^{H} + \mathbf{\Upsilon} \circ \hat{\mathbf{F}}_{\mathrm{RF}} = \mathbf{0} \tag{68}$$

$$\hat{\mathbf{F}}_{\mathrm{RF}}^{H}(\mathbf{F}_{\mathrm{opt}} - \hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}) = \mathbf{0} \tag{69}$$

$$\hat{\mathbf{F}}_{\mathrm{RF}}^{*} \circ \hat{\mathbf{F}}_{\mathrm{RF}} = \frac{1}{N_{\mathrm{t}}}\mathbf{1} \tag{70}$$

where $\mathbf{\Upsilon}_{ij}$ is the lagrangian multiplier associated with the equality constraint $[\mathbf{F}_{\mathrm{RF}}]_{ij}^{*}[\mathbf{F}_{\mathrm{RF}}]_{ij} = 1/N_{\mathrm{t}}$. Suppose that $\hat{\mathbf{F}}_{\mathrm{RF}}$ has full column rank, then equation (69) becomes

$$\hat{\mathbf{F}}_{\mathrm{BB}} = \hat{\mathbf{F}}_{\mathrm{RF}}^{+}\mathbf{F}_{\mathrm{opt}}. \tag{71}$$

where $\hat{\mathbf{F}}_{\mathrm{RF}}^{+} = (\hat{\mathbf{F}}_{\mathrm{RF}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}})^{-1}\hat{\mathbf{F}}_{\mathrm{RF}}^{H}$ is the Moore-Penrose pseudoinverse of $\mathbf{F}_{\mathrm{RF}}$. Inserting equation (71) into $\mathrm{tr}\big(\hat{\mathbf{F}}_{\mathrm{BB}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\big)$, we obtain

$$\begin{aligned}
&\mathrm{tr}\big(\hat{\mathbf{F}}_{\mathrm{BB}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\big) \\
&\quad = \mathrm{tr}\big(\mathbf{F}_{\mathrm{opt}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{RF}}^{+}\mathbf{F}_{\mathrm{opt}}\big) = \mathrm{tr}\big(\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{RF}}^{+}\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{opt}}^{H}\big) \\
&\quad \le \sum_{i=1}^{N_{\mathrm{t}}} \lambda_i\big(\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{RF}}^{+}\big)\lambda_i\big(\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{opt}}^{H}\big)
\end{aligned} \tag{72}$$

where $\lambda_i(\cdot)$ represents the eigenvalue of a Hermitian matrix in decreasing order. The inequality in (72) follows from [24, Lemma II.1]:

$$\sum_{i=1}^{n} \lambda_i(\mathbf{A})\lambda_{n-i+1}(\mathbf{B}) \le \mathrm{tr}(\mathbf{AB}) \le \sum_{i=1}^{n} \lambda_i(\mathbf{A})\lambda_i(\mathbf{B}) \tag{73}$$

where $\mathbf{A} \in \mathbb{C}^{n \times n}$ and $\mathbf{B} \in \mathbb{C}^{n \times n}$ are Hermitian matrices. Since $\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{RF}}^{+}$ is a projection matrix, its eigenvalues are given by

$$\lambda_i(\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{RF}}^{+}) = \begin{cases} 1, & i = 1, 2, \ldots, N_{\mathrm{rf}} \\ 0, & \text{otherwise} \end{cases} \tag{74}$$

Then $\mathrm{tr}\big(\hat{\mathbf{F}}_{\mathrm{BB}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\big)$ can be further upper bounded by

$$\mathrm{tr}\big(\hat{\mathbf{F}}_{\mathrm{BB}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}^{H}\hat{\mathbf{F}}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\big) \le \textstyle\sum_{i=1}^{N_{\mathrm{rf}}} \lambda_i\big(\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{opt}}^{H}\big) \le \mathrm{tr}(\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{opt}}^{H}) = P.$$

This completes the proof. ■

APPENDIX B
PROOFS OF THEOREM 2–3 AND LEMMA 1

Proof of Theorem 2: The KKT conditions of problem (25) are given by

$$-(\mathbf{F}_{\rm opt}-\mathbf{F}_{\rm RF}\mathbf{F}_{\rm BB})\mathbf{F}_{\rm BB}^H+\boldsymbol{\Upsilon}\circ\mathbf{F}_{\rm RF}=\mathbf{0} \quad (75)$$
$$\mathbf{F}_{\rm RF}^H(\mathbf{F}_{\rm opt}-\mathbf{F}_{\rm RF}\mathbf{F}_{\rm BB})=\mathbf{0} \quad (76)$$
$$\mathbf{F}_{\rm RF}^*\circ\mathbf{F}_{\rm RF}=\frac{1}{N_{\rm t}}\mathbf{1} \quad (77)$$

where $\boldsymbol{\Upsilon}_{ij}$ is the lagrangian multiplier associated with the equality constraint $[\mathbf{F}_{\rm RF}]_{kl}^*[\mathbf{F}_{\rm RF}]_{kl}=1/N_{\rm t}$. Suppose $\hat{\mathbf{F}}_{\rm RF}$ is a KKT point of problem (28) and $\hat{\mathbf{F}}_{\rm BB}=\hat{\mathbf{F}}_{\rm RF}^+\mathbf{F}_{\rm opt}$, $(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})$ satisfies equations (76) and (77). Moreover, $\hat{\mathbf{F}}_{\rm RF}$ satisfies the following stationarity condition of problem (28):

$$-\big(\mathbf{I}-\hat{\mathbf{F}}_{\rm RF}\hat{\mathbf{F}}_{\rm RF}^+\big)\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H(\hat{\mathbf{F}}_{\rm RF}^+)^H+\boldsymbol{\Upsilon}\circ\hat{\mathbf{F}}_{\rm RF}=\mathbf{0} \quad (78)$$

where $-\big(\mathbf{I}-\hat{\mathbf{F}}_{\rm RF}\hat{\mathbf{F}}_{\rm RF}^+\big)\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H(\hat{\mathbf{F}}_{\rm RF}^+)^H$ is the complex gradient of $f(\mathbf{F}_{\rm RF})$, and $\boldsymbol{\Upsilon}$ is the lagrangian multiplier. Inserting $\hat{\mathbf{F}}_{\rm BB}=\hat{\mathbf{F}}_{\rm RF}^+\mathbf{F}_{\rm opt}$ into equation (78), it becomes

$$-(\hat{\mathbf{F}}_{\rm opt}-\hat{\mathbf{F}}_{\rm RF}\hat{\mathbf{F}}_{\rm BB})\hat{\mathbf{F}}_{\rm BB}^H+\boldsymbol{\Upsilon}\circ\hat{\mathbf{F}}_{\rm RF}=\mathbf{0} \quad (79)$$

which is exactly the stationarity condition of problem (25) given in equation (75). Therefore, the KKT point of problem (28) satisfies equations (75)–(77) and it is a KKT point of problem (25).

Suppose that $\hat{\mathbf{F}}_{\rm RF}$ is a globally optimal solution of problem (28) and $\hat{\mathbf{F}}_{\rm BB}=\hat{\mathbf{F}}_{\rm RF}^+\mathbf{F}_{\rm opt}$, then

$$r(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})=f(\hat{\mathbf{F}}_{\rm RF}) \quad (80)$$

where $r(\mathbf{F}_{\rm RF},\mathbf{F}_{\rm BB})=\|\mathbf{F}_{\rm opt}-\mathbf{F}_{\rm RF}\mathbf{F}_{\rm BB}\|_F^2$. We further assume $(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})$ is not a globally optimal solution of problem (25), i.e., there exists a feasible solution $(\tilde{\mathbf{F}}_{\rm RF},\tilde{\mathbf{F}}_{\rm BB})$ such that $r(\tilde{\mathbf{F}}_{\rm RF},\tilde{\mathbf{F}}_{\rm BB})<r(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})$. Since for any given $\mathbf{F}_{\rm BB}$, $f(\mathbf{F}_{\rm RF})\le r(\mathbf{F}_{\rm RF},\mathbf{F}_{\rm BB})$, we have

$$f(\tilde{\mathbf{F}}_{\rm RF})\le r(\tilde{\mathbf{F}}_{\rm RF},\tilde{\mathbf{F}}_{\rm BB})<r(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})=f(\hat{\mathbf{F}}_{\rm RF}) \quad (81)$$

which is a contradiction to the fact that $\hat{\mathbf{F}}_{\rm RF}$ is a globally optimal solution of problem (28). Therefore, $(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})$ is a globally optimal solution of problem (25).

Conversely, suppose that $(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})$ is a globally optimal solution of problem (25), then

$$r(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})=f(\hat{\mathbf{F}}_{\rm RF}). \quad (82)$$

Similarly, we assume $\hat{\mathbf{F}}_{\rm RF}$ is not a globally optimal solution of problem (28), i.e., there exists a feasible $\tilde{\mathbf{F}}_{\rm RF}$ such that $f(\tilde{\mathbf{F}}_{\rm RF})<f(\hat{\mathbf{F}}_{\rm RF})$. Let $\tilde{\mathbf{F}}_{\rm BB}=\tilde{\mathbf{F}}_{\rm RF}^+\mathbf{F}_{\rm opt}$, then

$$f(\tilde{\mathbf{F}}_{\rm RF})=r(\tilde{\mathbf{F}}_{\rm RF},\tilde{\mathbf{F}}_{\rm BB})<f(\hat{\mathbf{F}}_{\rm RF})=r(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB}) \quad (83)$$

which is a contradiction to the fact that $(\hat{\mathbf{F}}_{\rm RF},\hat{\mathbf{F}}_{\rm BB})$ is a globally optimal solution of problem (25). Therefore, $\hat{\mathbf{F}}_{\rm RF}$ is a globally optimal solution of problem (28). This completes the proof. ■

Proof of Lemma 1: We first compute the complex gradient matrix $\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})$. Note that $f(\mathbf{F}_{\rm RF})$ can be rewritten as

$$f(\mathbf{F}_{\rm RF})=\|\mathbf{F}_{\rm opt}\|_F^2-{\rm tr}\big(\mathbf{F}_{\rm RF}^+\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H\mathbf{F}_{\rm RF}\big). \quad (84)$$

Then the differential of $f(\mathbf{F}_{\rm RF})$ is given by

$$\mathrm{d}f(\mathbf{F}_{\rm RF})=-{\rm tr}\big(\mathrm{d}\mathbf{F}_{\rm RF}^+\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H\mathbf{F}_{\rm RF}\big)-{\rm tr}\big(\mathbf{F}_{\rm RF}^+\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H\mathrm{d}\mathbf{F}_{\rm RF}\big). \quad (85)$$

The differential of $\mathbf{F}_{\rm RF}^+=(\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF})^{-1}\mathbf{F}_{\rm RF}^H$ in equation (85) can be computed as follows:

$$\begin{aligned}\mathrm{d}\mathbf{F}_{\rm RF}^+&=\mathrm{d}\big[(\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF})^{-1}\big]\mathbf{F}_{\rm RF}^H+(\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF})^{-1}\mathrm{d}\mathbf{F}_{\rm RF}^H\\&=(\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF})^{-1}\mathrm{d}\mathbf{F}_{\rm RF}^H(\mathbf{I}-\mathbf{F}_{\rm RF}\mathbf{F}_{\rm RF}^+)-\mathbf{F}_{\rm RF}^+\mathrm{d}\mathbf{F}_{\rm RF}\mathbf{F}_{\rm RF}^+\end{aligned} \quad (86)$$

where the second equality in (86) holds due to the following equation

$$\mathrm{d}\big(\mathbf{A}^{-1}\big)=-\mathbf{A}^{-1}\mathrm{d}\mathbf{A}\mathbf{A}^{-1}. \quad (87)$$

Inserting (86) into (85), we have

$$\begin{aligned}\mathrm{d}f(\mathbf{F}_{\rm RF})&\triangleq{\rm tr}\big(\mathrm{d}\mathbf{F}_{\rm RF}^H\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})+\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})^H\mathrm{d}\mathbf{F}_{\rm RF}\big) && (88)\\&=-{\rm tr}\big(\mathrm{d}\mathbf{F}_{\rm RF}^H\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H\big)-{\rm tr}\big(\mathbf{Z}_2\mathbf{F}_{\rm opt}^H\mathbf{Z}_1^H\mathrm{d}\mathbf{F}_{\rm RF}\big) && (89)\end{aligned}$$

where $\mathbf{Z}_1=\mathbf{I}-\mathbf{F}_{\rm RF}\mathbf{F}_{\rm RF}^+$ and $\mathbf{Z}_2=\mathbf{F}_{\rm RF}^+\mathbf{F}_{\rm opt}$. Thus the complex gradient matrix of $f(\mathbf{F}_{\rm RF})$ is $\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})=-\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H$.

Next, we compute the Hessian matrix $\mathcal{CH}_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})$. Since $\mathcal{CH}_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})$ contains four blocks, we first determine $\mathcal{H}_{\mathbf{F}_{\rm RF},\mathbf{F}_{\rm RF}^*}f(\mathbf{F}_{\rm RF})$ and $\mathcal{H}_{\mathbf{F}_{\rm RF}^*,\mathbf{F}_{\rm RF}^*}f(\mathbf{F}_{\rm RF})$. According to the definition

$$\begin{aligned}{\rm vec}\big[\mathrm{d}\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})\big]&\triangleq\mathcal{H}_{\mathbf{F}_{\rm RF},\mathbf{F}_{\rm RF}^*}f(\mathbf{F}_{\rm RF}){\rm vec}(\mathrm{d}\mathbf{F}_{\rm RF})\\&\quad+\mathcal{H}_{\mathbf{F}_{\rm RF}^*,\mathbf{F}_{\rm RF}^*}f(\mathbf{F}_{\rm RF}){\rm vec}(\mathrm{d}\mathbf{F}_{\rm RF}^*)\end{aligned} \quad (90)$$

we obtain $\mathcal{H}_{\mathbf{F}_{\rm RF},\mathbf{F}_{\rm RF}^*}f(\mathbf{F}_{\rm RF})$ and $\mathcal{H}_{\mathbf{F}_{\rm RF}^*,\mathbf{F}_{\rm RF}^*}f(\mathbf{F}_{\rm RF})$ through computing the differential of $\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})$:

$$\mathrm{d}\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})=-\mathrm{d}\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H-\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathrm{d}\mathbf{Z}_2^H. \quad (91)$$

where

$$\mathrm{d}\mathbf{Z}_1=-\mathrm{d}\mathbf{F}_{\rm RF}\mathbf{F}_{\rm RF}^+-\mathbf{F}_{\rm RF}\mathrm{d}\mathbf{F}_{\rm RF}^+,\quad \mathrm{d}\mathbf{Z}_2^H=\mathbf{F}_{\rm opt}^H(\mathrm{d}\mathbf{F}_{\rm RF}^+)^H. \quad (92)$$

Inserting $\mathrm{d}\mathbf{F}_{\rm RF}^+$ in (86) into (92), $\mathrm{d}\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})$ can be expressed as

$$\begin{aligned}\mathrm{d}\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})=\;&\mathbf{Z}_1\mathrm{d}\mathbf{F}_{\rm RF}\mathbf{Z}_2\mathbf{Z}_2^H+(\mathbf{F}_{\rm RF}^+)^H\mathrm{d}\mathbf{F}_{\rm RF}^H\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H\\&-\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H\mathbf{Z}_1^H\mathrm{d}\mathbf{F}_{\rm RF}(\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF})^{-1}\\&+\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H\mathrm{d}\mathbf{F}_{\rm RF}^H(\mathbf{F}_{\rm RF}^+)^H.\end{aligned} \quad (93)$$

Then we vectorize $\mathrm{d}\nabla_{\mathbf{F}_{\rm RF}}f(\mathbf{F}_{\rm RF})$ using the formula ${\rm vec}(\mathbf{AXB})=\big(\mathbf{B}^T\otimes\mathbf{A}\big){\rm vec}(\mathbf{X})$:

$$\begin{aligned}&{\rm vec}\big[\mathbf{Z}_1\mathrm{d}\mathbf{F}_{\rm RF}\mathbf{Z}_2\mathbf{Z}_2^H\big]\\&\quad={\rm vec}\big[(\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H)^T\otimes(\mathbf{F}_{\rm RF}^+)^H\big]{\rm vec}(\mathrm{d}\mathbf{F}_{\rm RF}) && (94)\\&{\rm vec}\big[\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H\mathbf{Z}_1^H\mathrm{d}\mathbf{F}_{\rm RF}(\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF})^{-1}\big]\\&\quad={\rm vec}\Big[\big[(\mathbf{F}_{\rm RF}^H\mathbf{F}_{\rm RF})^{-1}\big]^T\otimes\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{F}_{\rm opt}^H\mathbf{Z}_1^H\Big]{\rm vec}(\mathrm{d}\mathbf{F}_{\rm RF}) && (95)\\&{\rm vec}\big[(\mathbf{F}_{\rm RF}^+)^H\mathrm{d}\mathbf{F}_{\rm RF}^H\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H\big]\\&\quad=\Big[(\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H)^T\otimes(\mathbf{F}_{\rm RF}^+)^H\Big]\mathbf{K}_{N_{\rm t},N_{\rm rf}}{\rm vec}(\mathrm{d}\mathbf{F}_{\rm RF}^*) && (96)\\&{\rm vec}\big[\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H\mathrm{d}\mathbf{F}_{\rm RF}^H(\mathbf{F}_{\rm RF}^+)^H\big]\\&\quad=\Big[(\mathbf{F}_{\rm RF}^+)^*\otimes\mathbf{Z}_1\mathbf{F}_{\rm opt}\mathbf{Z}_2^H\Big]\mathbf{K}_{N_{\rm t},N_{\rm rf}}{\rm vec}(\mathrm{d}\mathbf{F}_{\rm RF}^*) && (97)\end{aligned}$$

where $\mathbf{K}_{N_t,N_{rf}}$ is the commutation matrix such that $\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}^H) = \mathbf{K}_{N_t,N_{rf}}\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}^*)$. Then we can obtain

$$\begin{aligned}\mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}}) &= (\mathbf{Z}_2\mathbf{Z}_2^H)^T \otimes \mathbf{Z}_1 \\ &\quad - \left[(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}})^{-1}\right]^T \otimes \mathbf{Z}_1\mathbf{F}_{\mathrm{opt}}\mathbf{F}_{\mathrm{opt}}^H\mathbf{Z}_1^H \quad (98)\\ \mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}}) &= (\mathbf{Z}_1\mathbf{F}_{\mathrm{opt}}\mathbf{Z}_2^H)^T \otimes (\mathbf{F}_{\mathrm{RF}}^+)^H\mathbf{K}_{N_t,N_{rf}} \\ &\quad + (\mathbf{F}_{\mathrm{RF}}^+)^* \otimes \mathbf{Z}_1\mathbf{F}_{\mathrm{opt}}\mathbf{Z}_2^H\mathbf{K}_{N_t,N_{rf}}. \quad (99)\end{aligned}$$

The remaining two blocks $\mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})$ and $\mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})$ can be obtained via $\mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})$ and $\mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})$. Since $\frac{\partial f(\mathbf{F}_{\mathrm{RF}})}{\partial \mathbf{F}_{\mathrm{RF}}} = [\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})]^*$, $\mathrm{vec}\left[\mathrm{d}\frac{\partial f(\mathbf{F}_{\mathrm{RF}})}{\partial \mathbf{F}_{\mathrm{RF}}}\right]$ can be expressed as

$$\begin{aligned}\mathrm{vec}\left[\mathrm{d}\frac{\partial f(\mathbf{F}_{\mathrm{RF}})}{\partial \mathbf{F}_{\mathrm{RF}}}\right] &\triangleq \mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}^*) \\ &\quad + \mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}) \quad (100)\\ &= \left[\mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})\right]^*\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}^*) \\ &\quad + \left[\mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})\right]^*\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}). \quad (101)\end{aligned}$$

As a consequence, one can obtain

$$\begin{aligned}\mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) &= \left[\mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})\right]^* \quad (102)\\ \mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) &= \left[\mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})\right]^*. \quad (103)\end{aligned}$$

This completes the proof. ■

Proof of Theorem 3: We first rewrite $\psi(\mathbf{\Phi}_{\mathrm{RF}})$ as the composition of $f(\mathbf{F}_{\mathrm{RF}})$ and $\mathbf{F}_{\mathrm{RF}}(\mathbf{\Phi}_{\mathrm{RF}})$, i.e.,

$$\psi(\mathbf{\Phi}_{\mathrm{RF}}) = f\left[\mathbf{F}_{\mathrm{RF}}(\mathbf{\Phi}_{\mathrm{RF}})\right]. \quad (104)$$

Using the chain rule in differentiation, the differential of $\psi(\bar{\mathbf{\Phi}})$ is

$$\begin{aligned}\mathrm{d}[\psi(\mathbf{\Phi}_{\mathrm{RF}})] &= \mathrm{tr}\left[\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})^H\mathrm{d}\mathbf{F}_{\mathrm{RF}}(\mathbf{\Phi}_{\mathrm{RF}}) \right.\\ &\quad \left. + \mathrm{d}\mathbf{F}_{\mathrm{RF}}(\mathbf{\Phi}_{\mathrm{RF}})^H\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})\right]. \quad (105)\end{aligned}$$

Inserting $\mathrm{d}\mathbf{F}_{\mathrm{RF}}(\mathbf{\Phi}_{\mathrm{RF}}) = j\mathbf{F}_{\mathrm{RF}} \circ \mathrm{d}\mathbf{\Phi}_{\mathrm{RF}}$ into (105), $\mathrm{d}[\psi(\mathbf{\Phi}_{\mathrm{RF}})]$ is expressed as

$$\begin{aligned}\mathrm{d}[\psi(\mathbf{\Phi}_{\mathrm{RF}})] &= j\mathrm{tr}\left[\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})^H(\mathbf{F}_{\mathrm{RF}} \circ \mathrm{d}\mathbf{\Phi}_{\mathrm{RF}}) \right.\\ &\quad \left. - (\mathbf{F}_{\mathrm{RF}} \circ \mathrm{d}\mathbf{\Phi}_{\mathrm{RF}})^H\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})\right] \quad (106)\\ &= j\mathrm{tr}\left[(\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})^* \circ \mathbf{F}_{\mathrm{RF}})^T\mathrm{d}\mathbf{\Phi}_{\mathrm{RF}} \right.\\ &\quad \left. - (\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) \circ \mathbf{F}_{\mathrm{RF}}^*)^T\mathrm{d}\mathbf{\Phi}_{\mathrm{RF}}\right] \quad (107)\end{aligned}$$

where (107) holds due to the following equality

$$\mathrm{tr}\left[\mathbf{A}^T(\mathbf{B} \circ \mathbf{C})\right] = \mathrm{tr}\left[(\mathbf{A} \circ \mathbf{B})^T\mathbf{C}\right]. \quad (108)$$

Then the gradient of $\psi(\mathbf{\Phi}_{\mathrm{RF}})$ can be obtained from (107):

$$\begin{aligned}\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}}) &= j\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})^* \circ \mathbf{F}_{\mathrm{RF}} - j\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) \circ \mathbf{F}_{\mathrm{RF}}^* \quad (109)\\ &= 2\Im\left[\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) \circ \mathbf{F}_{\mathrm{RF}}^*\right]. \quad (110)\end{aligned}$$

Next, we compute the Hessian of $\psi(\mathbf{\Phi}_{\mathrm{RF}})$. According to the definition

$$\mathrm{vec}\left[\mathrm{d}\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})\right] \triangleq \nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}})\mathrm{vec}(\mathrm{d}\mathbf{\Phi}_{\mathrm{RF}}) \quad (111)$$

we can obtain $\nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}})$ by computing the differential of $\mathrm{vec}[\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})]$:

$$\mathrm{vec}\left[\mathrm{d}\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})\right] = 2\Im\left\{\mathrm{d}\left(\mathrm{vec}\left[\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) \circ \mathbf{F}_{\mathrm{RF}}^*\right]\right)\right\}. \quad (112)$$

Using the product rule in differentiation, $\mathrm{d}\left(\mathrm{vec}\left[\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) \circ \mathbf{F}_{\mathrm{RF}}^*\right]\right)$ is given by

$$\begin{aligned}&\mathrm{d}\left(\mathrm{vec}\left[\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) \circ \mathbf{F}_{\mathrm{RF}}^*\right]\right)\\ &\quad = \mathrm{vec}\left[\mathrm{d}\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})\right] \circ \mathrm{vec}(\mathbf{F}_{\mathrm{RF}}^*)\\ &\quad\quad + \mathrm{vec}\left[\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})\right] \circ \mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}^*) \quad (113)\end{aligned}$$

where

$$\begin{aligned}\mathrm{vec}\left[\mathrm{d}\nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}})\right] &= \mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}) \\ &\quad + \mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})\mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}^*) \quad (114)\\ \mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}) &= j\mathrm{vec}(\mathbf{F}_{\mathrm{RF}}) \circ \mathrm{vec}(\mathrm{d}\mathbf{\Phi}_{\mathrm{RF}}) \quad (115)\\ \mathrm{vec}(\mathrm{d}\mathbf{F}_{\mathrm{RF}}^*) &= -j\mathrm{vec}(\mathbf{F}_{\mathrm{RF}}^*) \circ \mathrm{vec}(\mathrm{d}\mathbf{\Phi}_{\mathrm{RF}}) \quad (116)\end{aligned}$$

Inserting the equations in (114) into (113), $\mathrm{vec}\left[\mathrm{d}\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})\right]$ can be rewritten as

$$\mathrm{vec}\left[\mathrm{d}\nabla\psi(\mathbf{\Phi}_{\mathrm{RF}})\right] = \left\{2\Re(\mathbf{M}) - 2\mathrm{diag}\left(\mathrm{vec}\left[\Re(\mathbf{G})\right]\right)\right\}\mathrm{vec}(\mathrm{d}\mathbf{\Phi}_{\mathrm{RF}}) \quad (117)$$

where $\mathbf{G} = \nabla_{\mathbf{F}_{\mathrm{RF}}}f(\mathbf{F}_{\mathrm{RF}}) \circ \mathbf{F}_{\mathrm{RF}}^*$, and $\mathbf{M} = [\mathcal{H}_{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})] \circ \mathrm{vec}(\mathbf{F}_{\mathrm{RF}}^*)\mathrm{vec}(\mathbf{F}_{\mathrm{RF}})^T - [\mathcal{H}_{\mathbf{F}_{\mathrm{RF}}^*,\mathbf{F}_{\mathrm{RF}}^*}f(\mathbf{F}_{\mathrm{RF}})] \circ \mathrm{vec}(\mathbf{F}_{\mathrm{RF}}^*)\mathrm{vec}(\mathbf{F}_{\mathrm{RF}})^H$. Therefore, the Hessian matrix $\nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}})$ is given by

$$\nabla^2\psi(\mathbf{\Phi}_{\mathrm{RF}}) = 2\Re(\mathbf{M}) - 2\mathrm{diag}\left(\mathrm{vec}\left[\Re(\mathbf{G})\right]\right). \quad (118)$$

This completes the proof. ■

References

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

[2] E. Zhang and C. Huang, "On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems," in *Proc. IEEE 80th Veh. Technol. Conf. (VTC Fall)*, Sep. 2014, pp. 1–5.

[3] 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.

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

[5] C. Rusu, R. Mèndez-Rial, N. González-Prelcic, and R. W. Heath, Jr., "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8380–8393, Dec. 2016.

[6] R. Rajashekar and L. Hanzo, "Hybrid beamforming in mm-wave MIMO systems having a finite input alphabet," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3337–3349, Aug. 2016.

[7] R. Rajashekar and L. Hanzo, "Iterative matrix decomposition aided block diagonalization for mm-wave multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1372–1384, Mar. 2017.

[8] D. Guo, S. Shamai (Shitz), and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[9] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, Jul. 2006.

[10] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector Gaussian channels with finite discrete inputs," in *Proc. IEEE Globecom*, Nov. 2008, pp. 1–5.

[11] F. Pérez-Cruz, M. R. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1070–1084, Mar. 2010.

[12] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3301–3314, Jul. 2011.

[13] W. Zeng, C. Xiao, and J. Lu, "A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 38–41, Feb. 2012.

[14] J. Jin, C. Xiao, M. Tao, and W. Chen, "Linear precoding for fading cognitive multiple-access wiretap channel with finite-alphabet inputs," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3059–3070, Apr. 2017.

[15] J. Jin, Y. R. Zheng, W. Chen, and C. Xiao, "Generalized quadratic matrix programming: A unified framework for linear precoding with arbitrary input distributions," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4887–4901, Sep. 2017.

[16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[17] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.

[18] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.

[19] Y. Chen and E. Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 739–747.

[20] D.-H. Li and M. Fukushima, "On the global convergence of the BFGS method for nonconvex unconstrained optimization problems," *SIAM J. Optim.*, vol. 11, no. 4, pp. 1054–1064, May 2001.

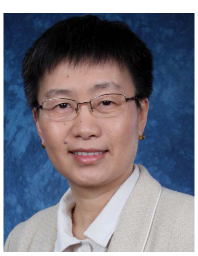
[21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[22] Y. N. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, "Identifying and attacking the saddle point problem in high-dimensional non-convex optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2933–2941.

[23] S. J. Reddi *et al.* (Sep. 2017). "A generic approach for escaping saddle points." [Online]. Available: https://arxiv.org/abs/1709.01434

[24] J. B. Lasserre, "A trace inequality for matrix product," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1500–1501, Aug. 1995.

Juening Jin received the B.S. degree in electronic engineering from Southeast University, Nanjing, China, in 2013. He is currently pursuing the Ph.D. degree with the Network Coding and Transmission Laboratory, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China. During his Ph.D. degree, he conducted cooperative research at the Missouri University of Science and Technology, Rolla, MO, USA. His research interests include convex and non-convex optimization, and physical-layer security.



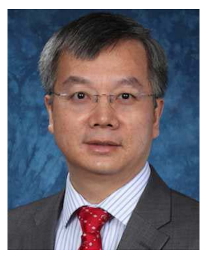
Yahong Rosa Zheng (M'03–SM'07–F'15) received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 1987, and the M.S. degree from Tsinghua University, Beijing, China, in 1989, and the Ph.D. degree from Carleton University, Ottawa, ON, Canada, in 2002. From 1989 to 1996, she was an Engineer and a member of the Technical Staff at several companies including the Beijing Peony Electronics Group, Beijing, China, Sagem Australasia, Sydney, Australia, and Polytronics, Markham, ON, Canada. From 2003 to 2005, she was an NSERC Post-Doctoral Fellow with the University of Missouri, Columbia, MO, USA. Since 2005, he has been a Faculty Member with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO, where she is currently the Wilkens Missouri Telecommunications Chair Professor.

Her research interests include array signal processing, wireless communications, and wireless sensor networks. She has published over 70 journal papers and over 120 conference papers in these areas. She was a recipient of the NSF CAREER Award in 2009. She has been a Distinguished Lecturer of the IEEE Vehicular Technology Society since 2015. She was an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She serves as an Associate Editor for the IEEE JOURNAL OF OCEANIC ENGINEERING.

Wen Chen (M'03–SM'11) received the B.S. and M.S. degrees from Wuhan University, China, in 1990 and 1993, respectively, and the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1999. In 2001, he joined the University of Alberta, Canada, starting as a Post-Doctoral Fellow and then becoming a Research Associate. Since 2006, he has been a Full Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China, where he is also the Director of the Institute for Signal Processing and Systems. His research interests include multiple access, coded cooperation, and green heterogenous networks. From 2014 to 2015, he was the Dean of the School of Electronics Engineering and Automations, Guilin University of Electronic Technology. Since 2016, he has been the Chairman of the SJTU Intellectual Property Management Corporation. He was a Research Fellow of the Japan Society for the Promotion of Sciences from 1999 to 2001.

Dr. Chen received the Ariyama Memorial Research Prize in 1997 and the PIMS Post-Doctoral Fellowship in 2001. He was a recipient of New Century Excellent Scholar in China in 2006 and Pujiang Excellent Scholar in Shanghai in 2007. He was elected the Vice General Secretary of the Shanghai Institute of Electronics in 2008. He received the Best Service Award of the China Institute of Electronics in 2013, the Best Paper Award of Chinese Information Theory Society in 2014, and the Innovate 5G Competition Award in 2015. He was invited to deliver a keynote speech at the IEEE APCC 2016 and tutorials at the IEEE ICCC 2016 and the IEEE VTC 2017.

He has published 88 papers in IEEE journals and over 110 papers in IEEE conferences. He is the Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS, and the IEEE ACCESS. His research interests include multiple access, coded cooperation, and green heterogenous networks.



Chengshan Xiao (M'99–SM'02–F'10) received the B.Sc. degree in electronic engineering from the University of Electronic Science and Technology of China in 1987, the M.Sc. degree in electronic engineering from Tsinghua University in 1989, and the Ph.D. degree in electrical engineering from The University of Sydney in 1997.

He served as the Program Director with the Division of Electrical, Communications and Cyber Systems, USA National Science Foundation. He was a Senior Member of Scientific Staff with Nortel Networks, Ottawa, ON, Canada, and a Faculty Member at Tsinghua University, Beijing, China, the University of Alberta, Edmonton, AB, Canada, the University of Missouri, Columbia, MO, USA, and the Missouri University of Science and Technology, Rolla, MO, USA. He also held visiting professor positions in Germany and Hong Kong. He is currently the Chandler Weaver Professor and the Chair of the Department of Electrical and Computer Engineering, Lehigh University. He holds several patents granted in the USA, Canada, China, and Europe. His invented algorithms have been implemented into Nortel's base station radio products after successful technical field trials and network integration. His research interests include wireless communications, signal processing, and underwater acoustic communications. He is a fellow of the Canadian Academy of Engineering.

Dr. Xiao served as an elected member of the Board of Governors, a member of the Fellow Evaluation Committee, the Director of Conference Publications, a Distinguished Lecturer of the IEEE Communications Society, and a Distinguished Lecturer of the IEEE Vehicular Technology. He received several distinguished awards, including the 2014 Humboldt Research Award, the 2014 IEEE Communications Society Joseph LoCicero Award, the 2015 IEEE Wireless Communications Technical Committee Recognition Award, and the 2017 IEEE Communications Society Harold Sobol Award. He is the Awards Committee Chair of the IEEE Communications Society. He was the Technical Program Chair of the 2010 IEEE International Conference on Communications, Cape Town, South Africa, and a Technical Program Co-Chair of the 2017 IEEE Global Communications Conference, Singapore. He served as the Founding Chair of the IEEE Wireless Communications Technical Committee. He served as an Editor, an Area Editor, and the Editor-in-Chief for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I.