

Article

Joint Hybrid Beamforming Design for Millimeter Wave Amplify-and-Forward Relay Communication Systems

Jinxian Zhao ¹, Dongfang Jiang ¹, Heng Wei ^{1,*}, Bingjie Liu ¹, Yifeng Zhao ², Yi Zhang ^{2,3}, Haoyuan Yu ¹ and Xuewei Liu ¹

¹ Key Laboratory of Smart Earth, Beijing 100094, China; jinxian_zhao@sohu.com (J.Z.); jiang_dongfang@126.com (D.J.); lbj012@126.com (B.L.); haoyuan_yu@foxmail.com (H.Y.); liuxuewei@foxmail.com (X.L.)

² Department of Information and Communication Engineering, Xiamen University, Xiamen 361102, China; zhaoyf@xmu.edu.cn (Y.Z.); yizhang@xmu.edu.cn (Y.Z.)

³ Perception and Efficient Computing, Ministry of Education of China, Xiamen 361102, China

* Correspondence: wh_killer@foxmail.com

Abstract: Hybrid beamforming (HBF) has been regarded as one of the most promising technologies in millimeter Wave (mmWave) communication systems. In order to guarantee the communication quality in non-line-of-sight (NLOS) scenarios, joint HBF design for the mmWave amplify-and-forward (AF) relay communication system is studied in this paper. The ideal case is first considered where the mmWave half-duplex (HD) AF relay system operates with channel state information (CSI) accurately known. In order to tackle the non-convex problem, a manifold optimization (MO)-based alternating optimization algorithm is proposed, where an optimization problem containing only constant modulus constraints in Euclidean space can be converted to an unconstrained optimization problem in a Riemann manifold. Furthermore, considering more practical cases with estimation errors of CSI, we investigate the robust joint HBF design with the system operating in full-duplex (FD) mode to obtain higher spectral efficiency (SE). A null-space projection (NP) based self-interference cancellation (SIC) algorithm is developed to attenuate the self-interference (SI). Different from the traditional SI suppression algorithm, there's no limit on the number of RF chains. Numerical results reveal that our proposed algorithms has a good convergence and can effectively deal with the influence of different CSI estimation errors. A significant performance improvement can be achieved in contrast with other approaches.

Keywords: mmWave communication system; AF relay; hybrid beamforming; manifold optimization; channel state information



Citation: Zhao, J.; Jiang, D.; Wei, H.; Liu, B.; Zhao, Y.; Zhang, Y.; Yu, H.; Liu, X. Joint Hybrid Beamforming Design for Millimeter Wave Amplify-and-Forward Relay Communication Systems. *Appl. Sci.* **2024**, *14*, 3713. <https://doi.org/10.3390/app14093713>

Academic Editor: Christos Bouras

Received: 26 March 2024

Revised: 22 April 2024

Accepted: 24 April 2024

Published: 26 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Compared with microwave communication systems, millimeter wave (mmWave) communication systems have larger bandwidth and higher data rate [1], receiving great attention from both academic and industry. However, high propagation loss is a major characteristic of mmWave signals [2], which should be compensated for by deploying large antenna arrays to guarantee the communication quality. Fortunately, due to the shorter wavelength of mmWave, more antennas can be integrated into the same space, enabling Massive MIMO technology in mmWave communication systems [3]. Nevertheless, the drastic increase in the number of antennas leads to a high cost, based on which researchers proposed the hybrid beamforming (HBF) technology.

mmWave communication is mostly used in line-of-sight (LOS) dominant scenarios owing to their high vulnerability to blockages. However, the source-destination link is probably non-line-of-sight (NLOS). Practically, relay nodes are commonly deployed to solve the problem [3,4]. Combined with the relay technology, hybrid transceivers can

combat severe attenuation of mmWave signals and extend network convergence. In addition, an mmWave relay system operating in full-duplex (FD) mode can achieve higher spectral efficiency (SE) than the one based on half-duplex (HD) by enabling simultaneous reception and transmission of relay nodes [5]. However, the self-interference (SI) imposed by simultaneous transmission should be mitigated to ensure the communication quality [6]. With effective suppression of SI, the FD technology can reduce end-to-end communication delay and improve spectral efficiency significantly.

Unfortunately, the HBF design problem in mmWave relay communication systems has not been extensively investigated. In [7], the authors exploited a transceiver HBF design on the basis of orthogonal matching pursuit (OMP), whose complexity is relatively low while the performance significantly relates to the orthogonality of the pre-defined analog beamforming (ABF) matrices. In [8], a wideband HBF algorithm was adopted in the multiple-relays scenario. The HBF design techniques of a relay node has been explored in [9,10]. In particular, the alternating direction method of multipliers (ADMM) algorithm was proposed to reduce the minimum square error (MSE) between received signals and transmit signals in [9]. To overcome high complexity, the optimization problem was decomposed into three sub-problems in [10], which is solved by an iterative successive approximation (ISA) algorithm. However, the above two algorithms only optimize the relay node. The authors in [11] developed an HBF algorithm to jointly design all nodes in the mmWave amplify-and-forward (AF) relay system, including the source node, the relay node and the destination node. In [12], the HBF design in mmWave multi-user MIMO relay system was investigated. The ABF of the base station, relay and users could be iteratively designed by solving the weighted minimum MSE (WMMSE) problem.

It is noteworthy that all of these works were under accurate channel state information (CSI). Due to channel estimation errors and user mobility, CSI is not available perfectly in practice. The works of [13,14] studied the robust HBF design for the mmWave FD relay system with the premise of imperfect CSI. To be specific, the work in [13] introduced a robust OMP HBF scheme for mmWave channels with Gaussian-distributed errors to maximize the average receive signal-to-noise ratio (SNR). To reduce complexity, the authors in [14] designed the radio frequency (RF) and baseband separately and showed a robust WMMSE-based algorithm with respect to the sum of rate for the fully connected AF relay networks. On the other hand, the robust design for mmWave HD relay systems were investigated in [15,16]. Given the impact of SI, a minimum MSE (MMSE) optimization problem for the relay node was deduced in [15]. In [16], the SI power constraint was taken into account in the worst-case sum rate maximization problem. To make it tractable [16], reformulated the problem by using the penalty dual decomposition technique. Currently, there are still few researchers focusing on HBF design for mmWave relay communication systems on the basis of channel estimation errors.

Against this background, the above problem of HBF design for the mmWave relay communication system is addressed in this paper. Instead of considering only relay nodes, all nodes are jointly optimized in this paper. Additionally, we take both perfect and imperfect CSI into account. It is worth noting that our proposed algorithms can be applied to both HD and FD mode. For clarity, the main contributions of our work are summarized as follows:

- We first consider the HD relay system under the assumption of perfect CSI. Different from conventional MSE minimization problems, where the power constraints introduce additional complexity, we model the improved MSE (IMSE) as the optimization objective by designing a power scaling factor.
- To deal with the complicated non-convex problem, a manifold optimization (MO) based alternating optimization algorithm is proposed, which decomposes the problem into three sub-problems. Simulation results demonstrate that our proposed method is superior to its counterparts by approaching the full digital solution.
- In terms of the FD relay communication system in the practical scenario, we develop a robust HBF algorithm. To mitigate SI introduced by the FD model, we further develop

a null-space-projection (NP)-based SI cancellation (SIC) method, which has no limit on the number of RF chains in contrast with traditional methods. Simulation shows that the proposed approach can achieve sufficient suppression of SI, providing significant beamforming gain.

The remaining sections are organized as follows. Section 2 presents the system model and the mmWave channel model. Section 3 introduces the HBF design for the HD AF relay communication system. The SIC-NP and HBF-MO-FD-R algorithms for the FD AF relay communication system are proposed in Section 4. Section 5 provides our simulation results. Finally, Section 6 draws conclusions.

Notations: In this paper, vectors and matrixes are denoted by bold lowercase letters and uppercase letters, respectively. $\|\cdot\|$ and $\|\cdot\|_F$ stand for the Euclidean norm of a vector and the Frobenius norm of a matrix. $\text{tr}(\cdot)$ and $(\cdot)^H$ represent the trace and conjugate transposition operations, respectively. $|\cdot|$ represents the magnitude of a complex number. $E\{\cdot\}$ denotes the expectation operator. $d(\cdot)$ and $\nabla(\cdot)$ denote the differential and the gradient of a function. \mathbb{R} and $\mathbb{C}^{m \times n}$ denote the set of all real matrixes and all complex $m \times n$ matrixes. \mathbf{I}_m denotes the $m \times m$ identity matrix. Some important symbols and their meanings are listed in Table 1.

Table 1. Notations.

Symbol	Meaning
N_T	The number of antennas of the source node
N_R	The number of antennas of the relay node
N_D	The number of antennas of the destination node
N_{RF}	The number of antennas of RF chains at each node
N_S	The number of data streams
\mathbf{s}	The original signal of the source node
\mathbf{F}	The HBF matrix of the source node
\mathbf{F}_B	The DBF matrix of the source node
\mathbf{F}_{RF}	The ABF matrix of the source node
P_B	The power threshold of the source node
\mathbf{H}_1	The mmWave channel matrix of the source-to-relay link
\mathbf{G}	The HBF matrix of the HD relay node
\mathbf{G}_{RF1}	The receive ABF matrix of the HD relay node
\mathbf{G}_{RF2}	The transmit ABF matrix of the HD relay node
\mathbf{G}_B	The DBF matrix of the HD relay node
σ_R^2	The noise variance of the relay node
P_R	The power threshold of the relay node
\mathbf{H}_2	The mmWave channel matrix of the relay-to-destination link
\mathbf{W}	The HBF matrix of the destination node
\mathbf{W}_{RF}	The ABF matrix of the destination node
\mathbf{W}_B	The DBF matrix of the destination node
σ_D^2	The noise variance of the destination node
$\gamma, \gamma_1, \gamma_2$	The scaling factors
\mathbf{F}_U	The unnormalized DBF matrix of the source node
\mathbf{G}_U	The unnormalized DBF matrix of the relay node
$\bar{\mathbf{F}}$	The unnormalized HBF matrix of the source node
$\bar{\mathbf{G}}$	The unnormalized HBF matrix of the relay node
\mathbf{G}_{B1}	The receive DBF matrix of the FD relay node
\mathbf{G}_{B2}	The transmit DBF matrix of the FD relay node
\mathbf{G}_1	The receive HBF matrix of the FD relay node
\mathbf{G}_2	The transmit HBF matrix of the FD relay node
\mathbf{H}_{SI}	The SI channel matrix of the FD relay node
$\overline{\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_{SI}}$	The estimated channel matrixes
Θ	The covariance matrix of estimation error at the receiver side
Φ	The covariance matrix of estimation error at the transmit side
Δ	The unknown part of CSI

2. System Model

As shown in Figure 1, an AF mmWave relay communication system is adopted in this paper, including a source node with N_T antennas, an HD AF relay node with N_R antennas and a destination node with N_D antennas. And all nodes are equipped with N_{RF} RF chains. In our scenario, the blockage between the source–destination link prevents direct communication. Consequently, it is necessary for the two nodes to communicate through the AF relay node.

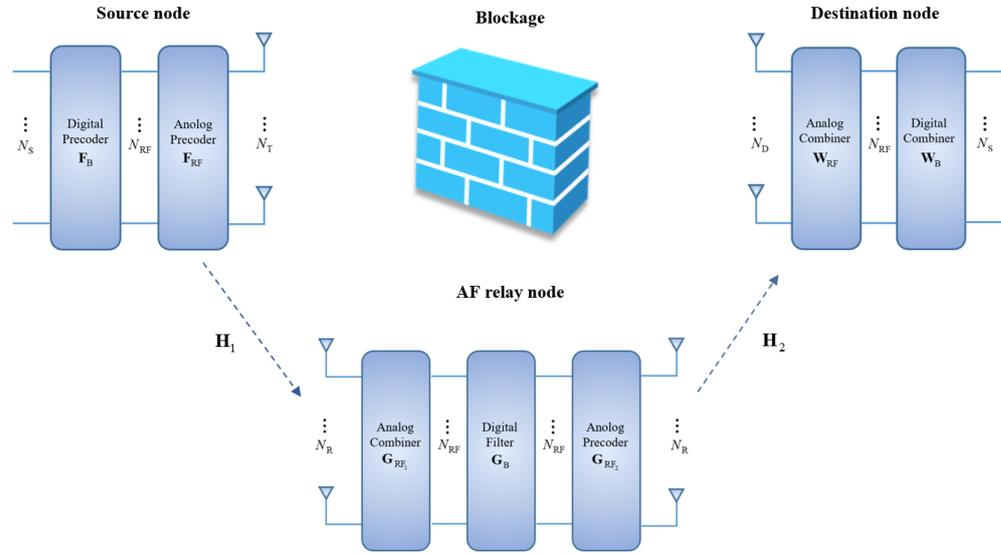


Figure 1. HD AF relay communication system.

We assume that $\mathbf{s} \in \mathbb{C}^{N_S \times 1}$ denotes the original signal of the source node characterized by $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_{N_S}$, with N_S representing the number of data streams. The signal is processed by the digital beamformer $\mathbf{F}_B \in \mathbb{C}^{N_{RF} \times N_S}$ and then the analog beamformer $\mathbf{F}_{RF} \in \mathbb{C}^{N_T \times N_{RF}}$, which subjects to the constant-modulus constraint, i.e., $|\mathbf{F}_{RF}|_{a,b}|^2 = 1$. Therefore, the source transmit signal can be given by $\mathbf{x}_B = \mathbf{F}\mathbf{s}$, where $\mathbf{F} = \mathbf{F}_{RF}\mathbf{F}_B$ is the HBF matrix of the source node. The source transmit power is stated as $E\{\|\mathbf{x}_B\|_F^2\} = P_B$.

Through the mmWave channel $\mathbf{H}_1 \in \mathbb{C}^{N_R \times N_T}$, the signal is received by the AF relay node, the transmit signal of which can be given by $\mathbf{x}_R = \mathbf{G}(\mathbf{H}_1\mathbf{x}_B + \mathbf{n}_R)$, where $\mathbf{G} = \mathbf{G}_{RF2}\mathbf{G}_B\mathbf{G}_{RF1}^H$ is the relay node HBF matrix, with $\mathbf{G}_{RF1} \in \mathbb{C}^{N_R \times N_{RF}}$, $\mathbf{G}_{RF2} \in \mathbb{C}^{N_R \times N_{RF}}$ and $\mathbf{G}_B \in \mathbb{C}^{N_{RF} \times N_{RF}}$ being the receive ABF matrix, transmit ABF matrix and digital beamforming (DBF) matrix, respectively. $\mathbf{n}_R \in \mathbb{C}^{N_R \times 1}$ denotes the additive complex Gaussian noise with zero mean and covariance matrix $E\{\mathbf{n}_R\mathbf{n}_R^H\} = \sigma_R^2\mathbf{I}_{N_R}$ at the relay node. Similarly, the analog beamformers should subject to the constant-modulus constraints, while the transmit signal \mathbf{x}_R should satisfy the power constraint $E\{\|\mathbf{x}_R\|_F^2\} = P_R$.

With an analog beamformer $\mathbf{W}_{RF} \in \mathbb{C}^{N_D \times N_{RF}}$ followed by a digital beamformer $\mathbf{W}_B \in \mathbb{C}^{N_{RF} \times N_S}$, the received signal at the destination node takes the form of

$$\mathbf{y}_D = \mathbf{W}^H\mathbf{H}_2\mathbf{G}\mathbf{H}_1\mathbf{F}\mathbf{s} + \mathbf{W}^H\mathbf{H}_2\mathbf{G}\mathbf{n}_R + \mathbf{W}^H\mathbf{n}_D, \quad (1)$$

where $\mathbf{H}_2 \in \mathbb{C}^{N_D \times N_R}$ denotes the mmWave channel of the relay-to-destination link, $\mathbf{n}_D \in \mathbb{C}^{N_D \times 1}$ is the zero-mean additive complex Gaussian noise vector with covariance matrix $E\{\mathbf{n}_D\mathbf{n}_D^H\} = \sigma_D^2\mathbf{I}_{N_D}$ at the destination node. For simplicity, we define $\mathbf{W} = \mathbf{W}_{RF}\mathbf{W}_B$ as the HBF matrix of the destination node.

Channel Model

The mmWave channel model in [17–20] is considered in this paper, which takes the following form:

$$\mathbf{H} = \sqrt{\frac{N_T N_R}{L}} \sum_{l=1}^L \alpha_l \mathbf{a}_r(\varphi_l^r) \mathbf{a}_t^H(\varphi_l^t), \tag{2}$$

where L denotes the number of propagation paths, α_l denotes the complex channel gain of the l^{th} path. $\mathbf{a}_t(\varphi_l^t)$ and $\mathbf{a}_r(\varphi_l^r)$ are the transmitting and receiving array response vectors. φ_l^t and φ_l^r stand for the elevation angles of departure (AOD) and azimuth angles of arrival (AOA) for the l^{th} path, respectively. It is worth noting that our proposed algorithms can be applied to any antenna array. For simplicity, the commonly applied uniform linear array (ULA) with N elements are considered. The array response vector can be given by

$$\mathbf{a}(\varphi) = \frac{1}{\sqrt{N}} [1, e^{j\frac{2\pi}{\lambda}d \sin(\varphi)}, \dots, e^{j(N-1)\frac{2\pi}{\lambda}d \sin(\varphi)}]^T, \tag{3}$$

where λ represents the wavelength and d stand for the distance between antenna elements, which is assumed to satisfy $d = \lambda/2$. In this section, both \mathbf{H}_1 and \mathbf{H}_2 are modelled as (2).

3. Hybrid Beamforming Design for HD AF Relay Systems

3.1. Problem Formulation

A mathematical model of the joint HBF design problem is formulated in this subsection. According to [21,22], the IMSE is introduced as the metric to optimize the HBF matrix, which can be defined as

$$\text{IMSE} \triangleq \mathbb{E} \left\{ \left\| \gamma^{-1} \mathbf{y}_D - \mathbf{s} \right\|^2 \right\}, \tag{4}$$

where $\gamma = \gamma_1 \gamma_2$, with γ_1 and γ_2 being the scaling factors, which are jointly optimized with the HBF matrixes at the source and relay nodes, contributing to simplify and solve the problem. By substituting (1) into (4), we have

$$\begin{aligned} \text{IMSE} = & \text{tr}(\gamma_2^{-2} \gamma_1^{-2} \mathbf{W}^H \mathbf{H}_2 \mathbf{G} \mathbf{H}_1 \mathbf{F} \mathbf{F}^H \mathbf{H}_1^H \mathbf{G}^H \mathbf{H}_2^H \mathbf{W} - \gamma_2^{-1} \gamma_1^{-1} \mathbf{W}^H \mathbf{H}_2 \mathbf{G} \mathbf{H}_1 \mathbf{F} \\ & - \gamma_2^{-1} \gamma_1^{-1} \mathbf{F}^H \mathbf{H}_1^H \mathbf{G}^H \mathbf{H}_2^H \mathbf{W} + \gamma_2^{-2} \gamma_1^{-2} \sigma_R^2 \mathbf{W}^H \mathbf{H}_2 \mathbf{G} \mathbf{G}^H \mathbf{H}_2^H \mathbf{W} \\ & + \gamma_2^{-2} \gamma_1^{-2} \sigma_D^2 \mathbf{W}^H \mathbf{W} + \mathbf{I}_{N_S}). \end{aligned} \tag{5}$$

Similar to that in [22–24], the source and relay DBF matrixes can be decomposed as $\mathbf{F}_B = \gamma_1 \mathbf{F}_U$ and $\mathbf{G}_B = \gamma_2 \mathbf{G}_U$, with \mathbf{F}_U and \mathbf{G}_U being the unnormalized DBF matrix of the source and relay nodes, respectively. So we can rewrite the modified MSE as

$$\begin{aligned} \text{IMSE} = & \text{tr}(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \overline{\mathbf{F}}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} - \mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} - \overline{\mathbf{F}}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} \\ & + \gamma_1^{-2} \sigma_R^2 \mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} + \gamma_1^{-2} \gamma_2^{-2} \sigma_D^2 \mathbf{W}^H \mathbf{W} + \mathbf{I}_{N_S}) \end{aligned} \tag{6}$$

where we define $\overline{\mathbf{F}} = \mathbf{F}_{RF} \mathbf{F}_U$ and $\overline{\mathbf{G}} = \mathbf{G}_{RF2} \mathbf{G}_U \mathbf{G}_{RF1}^H$ for convenience. Hence, we have $\mathbf{F} = \gamma_1 \overline{\mathbf{F}}$ and $\mathbf{G} = \gamma_2 \overline{\mathbf{G}}$. As a further step, the optimization problem can be formulated as

$$\max_{\gamma_1, \gamma_2, \mathbf{F}_{RF}, \mathbf{F}_U, \mathbf{G}_{RF1}, \mathbf{G}_U, \mathbf{G}_{RF2}, \mathbf{W}_{RF}, \mathbf{W}_B} \text{IMSE} \tag{7a}$$

$$\text{s.t.} \quad \left| [\mathbf{F}_{RF}]_{a,b} \right|^2 = 1, \forall a, b, \tag{7b}$$

$$\left| [\mathbf{G}_{RF1}]_{c,d} \right|^2 = 1, \forall c, d, \tag{7c}$$

$$\left| [\mathbf{G}_{RF2}]_{e,f} \right|^2 = 1, \forall e, f, \tag{7d}$$

$$\left| [\mathbf{W}_{RF}]_{g,h} \right|^2 = 1, \forall g, h, \tag{7e}$$

$$\mathbb{E} \left\{ \|\mathbf{x}_B\|_F^2 \right\} = P_B, \tag{7f}$$

$$\mathbb{E} \left\{ \|\mathbf{x}_R\|_F^2 \right\} = P_R, \tag{7g}$$

where (7b)–(7e) are the constant-modulus constraints of ABF matrixes, (7f) and (7g) are the source and relay transmit power constraints.

3.2. MO-Based HBF Design

It can be observed from the optimization problem (7) that there exist nine variables and four nonconvex constraints, which is tough to get the global optimal solution. In this subsection, an alternating optimization algorithm on the basis of the MO method is proposed. To overcome the difficulty of solving the original problem, it is decomposed into three HBF design sub-problems [25,26].

(1) *The Source Node Design:* When \mathbf{G} and \mathbf{W} are fixed, the sub-problem can be further expressed as

$$\min_{\gamma_1, \mathbf{F}_{RF}, \mathbf{F}_U} \text{IMSE} \tag{8a}$$

$$\text{s.t.} \quad \left| [\mathbf{F}_{RF}]_{a,b} \right|^2 = 1, \forall a, b, \tag{8b}$$

$$E \left\{ \|\mathbf{x}_B\|_F^2 \right\} = P_B. \tag{8c}$$

According to $\mathbf{F} = \gamma_1 \mathbf{F}_{RF} \mathbf{F}_U$, (8c) can be rewritten as $\gamma_1^2 \text{tr}(\overline{\mathbf{F}} \mathbf{F}^H) = P_B$, and we can obtain

$$\gamma_1 = \left[\frac{\text{tr}(\mathbf{F}_{RF} \mathbf{F}_U \mathbf{F}_U^H \mathbf{F}_{RF}^H)}{P_B} \right]^{-\frac{1}{2}}. \tag{9}$$

To this end, the Lagrange function $L1(\lambda_1, \mathbf{F}_U, \gamma_1)$ of problem (8) can be constructed. The minimum of $L1(\lambda_1, \mathbf{F}_U, \gamma_1)$ is achieved when its partial derivatives equal to zero, which gives the optimal \mathbf{F}_U as

$$\mathbf{F}_U = \left(\mathbf{F}_{RF}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} \mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} + \mathbf{B}_{11} \right)^{-1} \mathbf{F}_{RF}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W}, \tag{10}$$

with $\mathbf{B}_{11} = b_{11} (\mathbf{F}_{RF}^H \mathbf{F}_{RF})$ and $b_{11} = \frac{\sigma_R^2}{P_B} \text{tr}(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{G}^H \mathbf{H}_2^H \mathbf{W}) + \frac{\sigma_D^2}{P_B} \gamma_2^{-2} \text{tr}(\mathbf{W}^H \mathbf{W})$.

One more step, substituting (8a) with (10), the optimization sub-problem can be concisely re-formulated as

$$\min_{\mathbf{F}_{RF}} T1(\mathbf{F}_{RF}) \tag{11a}$$

$$\text{s.t.} \quad \left| [\mathbf{F}_{RF}]_{a,b} \right|^2 = 1, \forall a, b, \tag{11b}$$

where we denote

$$T1(\mathbf{F}_{RF}) = \text{tr} \left\{ \left[\mathbf{I}_{N_S} + \left(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \right) \mathbf{B}_{11}^{-1} \left(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \right)^H \right]^{-1} \right\} \tag{12}$$

We can see that (11) is an optimization problem with only a constant-modulus constraint. On the grounds of [21,27], we can tackle it by applying the MO method. However, it is difficult to use MO in the Euclidean space directly. Fortunately, if the gradient of objective function in Euclidean space is derived, the above problem can be easily solved [27]. The gradient of $T1(\mathbf{F}_{RF})$ with respect to \mathbf{F}_{RF} is defined as $\nabla T1(\mathbf{F}_{RF}) = \frac{\partial T1(\mathbf{F}_{RF})}{\partial \mathbf{F}_{RF}^*}$ [28], which is given by

$$\nabla T1(\mathbf{F}_{RF}) = \left(b_{11} \mathbf{F}_{RF} \mathbf{B}_{11}^{-1} \mathbf{F}_{RF}^H - \mathbf{I}_{N_T} \right) \left(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \right)^H \mathbf{B}_{12}^{-2} \mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \mathbf{B}_{11}^{-1}, \tag{13}$$

where $\mathbf{B}_{12} \triangleq \mathbf{I}_{N_S} + \left(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \right) \mathbf{B}_{11}^{-1} \left(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \right)^H$ for brevity. The detailed derivation is provided in Appendix A.

(2) *The Relay Node Design:* In a similar manner, it is assumed that \mathbf{F} and \mathbf{W} are fixed. The optimization sub-problem of the relay node can be formulated as

$$\min_{\mathbf{G}_{\text{RF}_1}, \mathbf{G}_B, \mathbf{G}_{\text{RF}_2}} \text{IMSE} \tag{14a}$$

$$\text{s.t.} \quad \left| [\mathbf{G}_{\text{RF}_1}]_{c,d} \right|^2 = 1, \forall c, d, \tag{14b}$$

$$\left| [\mathbf{G}_{\text{RF}_2}]_{e,f} \right|^2 = 1, \forall e, f, \tag{14c}$$

$$\mathbb{E} \left\{ \|\mathbf{x}_R\|_F^2 \right\} = P_R. \tag{14d}$$

For any given \mathbf{G}_{RF_1} and \mathbf{G}_{RF_2} , we can obtain

$$\gamma_2 = \left[\frac{\text{tr} \left(\gamma_1^2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \overline{\mathbf{F}}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H + \sigma_R^2 \overline{\mathbf{G}} \overline{\mathbf{G}}^H \right)}{P_R} \right]^{-\frac{1}{2}}. \tag{15}$$

This leads to the optimization problem with only power constraint. Based on the Lagrange function, we have

$$\begin{aligned} \mathbf{G}_U &= \left(\mathbf{G}_{\text{RF}_2}^H \mathbf{H}_2^H \mathbf{W} \mathbf{W}^H \mathbf{H}_2 \mathbf{G}_{\text{RF}_2} + \frac{\sigma_D^2}{P_R} \text{tr}(\mathbf{W}^H \mathbf{W}) \mathbf{G}_{\text{RF}_2}^H \mathbf{G}_{\text{RF}_2} \right)^{-1} \\ &\times \left(\mathbf{G}_{\text{RF}_2}^H \mathbf{H}_2^H \mathbf{W} \overline{\mathbf{F}}^H \mathbf{H}_1^H \mathbf{G}_{\text{RF}_1} \mathbf{G}_{\text{RF}_1}^H \mathbf{H}_1 \overline{\mathbf{F}} \overline{\mathbf{F}}^H \mathbf{H}_1^H \mathbf{G}_{\text{RF}_1} + \gamma_1^{-2} \sigma_R^2 \mathbf{G}_{\text{RF}_1}^H \mathbf{G}_{\text{RF}_1} \right)^{-1}. \end{aligned} \tag{16}$$

By substituting (15) into (14a), the objective function is reduced to

$$\text{T2}(\mathbf{G}_{\text{RF}_1}, \mathbf{G}_{\text{RF}_2}) = \text{tr} \left(\mathbf{I}_{N_{\text{RF}}} - \mathbf{B}_{22}^H \mathbf{B}_{21}^{-1} \mathbf{B}_{22} \mathbf{B}_{23}^{-1} \right), \tag{17}$$

where $\mathbf{B}_{21} = \mathbf{G}_{\text{RF}_2}^H \mathbf{H}_2^H \mathbf{W} \mathbf{W}^H \mathbf{H}_2 \mathbf{G}_{\text{RF}_2} + \frac{\sigma_D^2}{P_R} \text{tr}(\mathbf{W}^H \mathbf{W}) \mathbf{G}_{\text{RF}_2}^H \mathbf{G}_{\text{RF}_2}$, $\mathbf{B}_{22} = \mathbf{G}_{\text{RF}_2}^H \mathbf{H}_2^H \mathbf{W} \overline{\mathbf{F}}^H \mathbf{H}_1^H \mathbf{G}_{\text{RF}_1}$ and $\mathbf{B}_{23} = \mathbf{G}_{\text{RF}_1}^H \mathbf{H}_1 \overline{\mathbf{F}} \overline{\mathbf{F}}^H \mathbf{H}_1^H \mathbf{G}_{\text{RF}_1} + \gamma_1^{-2} \sigma_R^2 \mathbf{G}_{\text{RF}_1}^H \mathbf{G}_{\text{RF}_1}$. The partial derivative of $\text{T2}(\mathbf{G}_{\text{RF}_1}, \mathbf{G}_{\text{RF}_2})$ for $\mathbf{G}_{\text{RF}_1}^*$ and $\mathbf{G}_{\text{RF}_2}^*$ can be obtained as

$$\begin{aligned} \nabla_{\mathbf{G}_{\text{RF}_1}} \text{T2}(\mathbf{G}_{\text{RF}_1}, \mathbf{G}_{\text{RF}_2}) &= \left(\mathbf{H}_1 \overline{\mathbf{F}} \overline{\mathbf{F}}^H \mathbf{H}_1^H \mathbf{G}_{\text{RF}_1} \mathbf{B}_{23}^{-1} \mathbf{B}_{22}^H \right. \\ &\left. + \gamma_1^{-2} \sigma_R^2 \mathbf{G}_{\text{RF}_1} \mathbf{B}_{23}^{-1} \mathbf{B}_{22}^H - \mathbf{H}_1 \overline{\mathbf{F}} \mathbf{W}^H \mathbf{H}_2 \mathbf{G}_{\text{RF}_2} \right) \mathbf{B}_{21}^{-1} \mathbf{B}_{22} \mathbf{B}_{23}^{-1}, \end{aligned} \tag{18}$$

$$\begin{aligned} \nabla_{\mathbf{G}_{\text{RF}_2}} \text{T2}(\mathbf{G}_{\text{RF}_1}, \mathbf{G}_{\text{RF}_2}) &= \left(\mathbf{H}_2^H \mathbf{W} \mathbf{W}^H \mathbf{H}_2 \mathbf{G}_{\text{RF}_2} \mathbf{B}_{21}^{-1} \mathbf{B}_{22} \right. \\ &\left. + \frac{\sigma_D^2}{P_R} \text{tr}(\mathbf{W}^H \mathbf{W}) \mathbf{G}_{\text{RF}_2} \mathbf{B}_{21}^{-1} \mathbf{B}_{22} - \mathbf{H}_2^H \mathbf{W} \overline{\mathbf{F}}^H \mathbf{H}_1^H \mathbf{G}_{\text{RF}_1} \right) \mathbf{B}_{23}^{-1} \mathbf{B}_{22} \mathbf{B}_{21}^{-1}. \end{aligned} \tag{19}$$

(3) *The Destination Node Design:* With \mathbf{F} and \mathbf{G} fixed, the optimization sub-problem of the destination node can be expressed as

$$\min_{\mathbf{W}_{\text{RF}}, \mathbf{W}_B} \text{IMSE} \tag{20a}$$

$$\text{s.t.} \quad \left| [\mathbf{W}_{\text{RF}}]_{g,h} \right|^2 = 1, \forall g, h. \tag{20b}$$

Compared with the source and relay nodes, the destination node only transmits signals without receiving signals. Hence, there is no power constraint. When \mathbf{W}_{RF} is constant, (20) can be converted to an unconstrained optimization problem. By taking the derivative, the optimal \mathbf{W}_B is given by

$$\mathbf{W}_B = \left(\mathbf{W}_{RF}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \mathbf{F}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W}_{RF} + \gamma_1^{-2} \sigma_R^2 \mathbf{W}_{RF}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{G}^H \mathbf{H}_2^H \mathbf{W}_{RF} + \gamma_1^{-2} \gamma_2^{-2} \sigma_D^2 \mathbf{W}_{RF}^H \mathbf{W}_{RF} \right)^{-1} \mathbf{W}_{RF}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}}. \quad (21)$$

Substituting (21) into (20a), let $\mathbf{B}_{31} = \gamma_1^{-2} \sigma_R^2 \mathbf{W}_{RF}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{G}^H \mathbf{H}_2^H \mathbf{W}_{RF} + \gamma_1^{-2} \gamma_2^{-2} \sigma_D^2 \mathbf{W}_{RF}^H \mathbf{W}_{RF}$, we can recast the objective function as

$$T3(\mathbf{W}_{RF}) = \text{tr} \left[\left(\mathbf{I}_{N_S} + \overline{\mathbf{F}}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W}_{RF} \mathbf{B}_{31}^{-1} \mathbf{W}_{RF}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \right)^{-1} \right]. \quad (22)$$

Likewise, the gradient is given by

$$\begin{aligned} \nabla T3(\mathbf{W}_{RF}) = & \left(\gamma_1^{-2} \sigma_R^2 \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{G}^H \mathbf{H}_2^H \mathbf{W}_{RF} \mathbf{B}_{31}^{-1} \mathbf{W}_{RF}^H \right. \\ & \left. + \gamma_1^{-2} \gamma_2^{-2} \sigma_D^2 \mathbf{W}_{RF} \mathbf{B}_{31}^{-1} \mathbf{W}_{RF}^H - \mathbf{I}_{N_D} \right) \mathbf{B}_{33} \mathbf{B}_{32}^{-2} \mathbf{B}_{33}^H \mathbf{W}_{RF} \mathbf{B}_{31}^{-1}, \end{aligned} \quad (23)$$

where we define $\mathbf{B}_{32} = \mathbf{I}_{N_S} + \overline{\mathbf{F}}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W}_{RF} \mathbf{B}_{31}^{-1} \mathbf{W}_{RF}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}}$ and $\mathbf{B}_{33} = \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}}$.

(4) *Overall Algorithm:* On the basis of the above description, the proposed HBF-MO algorithm for the HD relay system is summarized in Algorithm 1. The original joint IMSE problem can be decomposed into three sub-problems, which are optimized alternately until the stopping criterion is met, as shown in Step 3 to Step 12. Moreover, the MO-based approach generates a local optimal solution instead of a global one. As a consequence, the performance of our proposed algorithm depends on the initialization of system parameters. In practice, the HBF matrix keeps approaching to the full DBF matrix with the number of iterations increasing. In this paper, we start with the full DBF matrix calculated by the singular value decomposition (SVD).

Algorithm 1 HBF-MO Algorithm.

Input: $\mathbf{H}_1, \mathbf{H}_2, \sigma_R^2, \sigma_D^2$.

Output: $\gamma_1, \gamma_2, \mathbf{F}_{RF}, \mathbf{F}_U, \mathbf{G}_{RF1}, \mathbf{G}_U, \mathbf{G}_{RF2}, \mathbf{W}_{RF}, \mathbf{W}_B$.

- 1: **Initialize** $\mathbf{F}_{RF}, \mathbf{W}_{RF}, \mathbf{G}_{RF1}$ and \mathbf{G}_{RF2} randomly.
 - 2: **Initialize** $\mathbf{F}_U, \mathbf{G}_U$ and \mathbf{W}_B with full DBF matrixes, and set $\gamma_1 = \gamma_2 = 1, k = 1$.
 - 3: **while** $\text{IMSE}_k - \text{IMSE}_{k-1} > \varepsilon$ **do**
 - 4: Calculate $\nabla T1(\mathbf{F}_{RF,k})$ according to (13) and compute $\mathbf{F}_{RF,k+1}$ by using the MO method.
 - 5: Update $\gamma_{1,k+1}, \mathbf{F}_{U,k+1}$ and \mathbf{F}_{k+1} according to (9) and (10).
 - 6: Calculate $\nabla_{\mathbf{G}_{RF1}} T2(\mathbf{G}_{RF1,k}, \mathbf{G}_{RF2,k})$ according to (18) and compute $\mathbf{G}_{RF1,k+1}$ by using the MO method.
 - 7: Calculate $\nabla_{\mathbf{G}_{RF2}} T2(\mathbf{G}_{RF1,k+1}, \mathbf{G}_{RF2,k})$ according to (19) and compute $\mathbf{G}_{RF2,k+1}$ by using the MO method.
 - 8: Update $\gamma_{2,k+1}, \mathbf{G}_{U,k+1}$ and \mathbf{G}_{k+1} according to (15) and (16).
 - 9: Calculate $\nabla T3(\mathbf{W}_{RF,k})$ according to (23) and compute $\mathbf{W}_{RF,k+1}$ by using the MO method.
 - 10: Update $\mathbf{W}_{B,k+1}$ and \mathbf{W}_{k+1} according to (21).
 - 11: $k \leftarrow k + 1$.
 - 12: **end while**
-

3.3. Algorithm Evaluation

(1) *Convergence Analysis:* The HBF-MO algorithm mainly includes two loops, which are the outer loop of Step 3 to Step 12 and the inner loop of Step 4, 6, 7, and 9 in Algorithm 1, respectively. With respect to the inner loop, by the definition of Theorem 4.3.1 in [29], the MO-based algorithm guarantees to converge to a point at which the first-order optimality condition is met [21,27]. As a result, the objective function is ensured to decrease

in each iteration of the alternating algorithm. For the outer loop, the DBF matrixes are obtained by taking the derivative, guaranteeing the convergence.

(2) *Complexity Analysis*: The complexity consists of three parts: (a) calculating the gradient, with the complexity of $\mathcal{O}(N_R^2 N_{MAX2})$, (b) orthogonal projection and retraction, whose complexity is $\mathcal{O}(N_{MAX1} N_{RF})$, and (c) line search, whose complexity is $\mathcal{O}(N_R^2 N_{MAX2})$, where $N_{MAX1} = \max(N_T, N_R, N_D)$, $N_{MAX2} = \max(N_T, N_D)$ and $\min(N_T, N_R, N_D) > N_{RF}$. Therefore, the total complexity is $N_{out} N_{in} \mathcal{O}(N_R^2 N_{MAX2})$, with N_{in} and N_{out} denoting the number of iterations of the inner and outer loops.

4. Hybrid Beamforming Design for FD AF Relay Systems

In this section, a more common case in practical applications is further considered, where the FD AF relay communication system operates with CSI unavailable.

4.1. System Model

(1) *System Model*: It is inevitable that the AF relay system will lead to severe SI in FD mode. As shown in Figure 2, in order to mitigate the SI, we decompose the DBF matrix \mathbf{G}_B of the relay node into $\mathbf{G}_{B1} \in \mathbb{C}^{N_{RF} \times N_S}$ and $\mathbf{G}_{B2} \in \mathbb{C}^{N_{RF} \times N_S}$ to design individually, which are the relay digital receive and transmit beamforming matrixes, respectively.

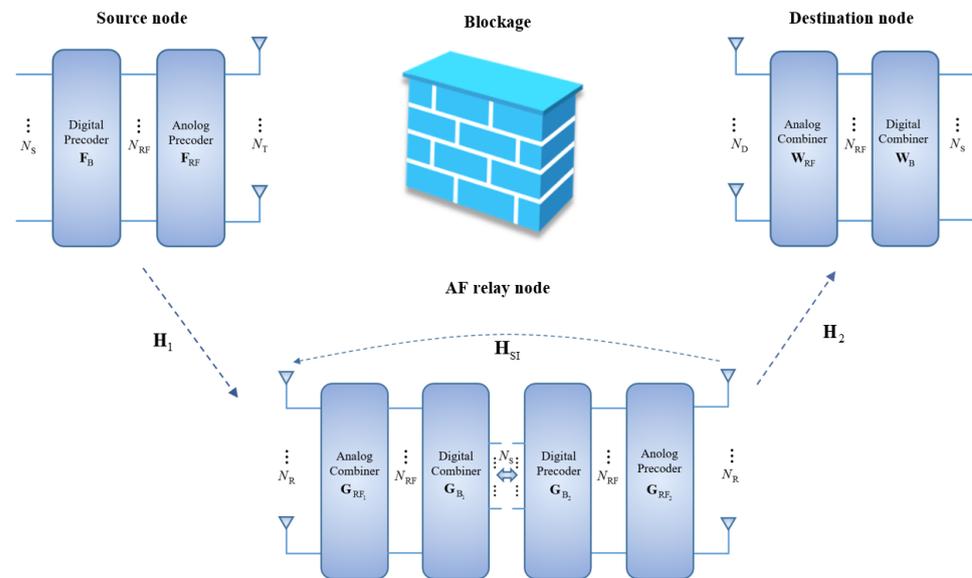


Figure 2. FD AF relay communication system.

The FD relay node can receive its own transmit signal and that of the source node simultaneously. Hence, the received signal of the relay node takes the form of [30,31]

$$\mathbf{y}_R = \mathbf{G}_1^H \mathbf{H}_1 \mathbf{F}_S \mathbf{s}_B + \mathbf{G}_1^H \mathbf{H}_{SI} \mathbf{G}_2 \mathbf{s}_R + \mathbf{G}_1^H \mathbf{n}_R, \tag{24}$$

where $\mathbf{G}_1 = \mathbf{G}_{RF1} \mathbf{G}_{B1} \in \mathbb{C}^{N_R \times N_S}$ and $\mathbf{G}_2 = \mathbf{G}_{RF2} \mathbf{G}_{B2} \in \mathbb{C}^{N_R \times N_S}$ are the relay hybrid receive and transmit beamforming matrixes. \mathbf{s}_B and \mathbf{s}_R denote the transmit signals of the source and relay nodes, which are subject to $E\{\mathbf{s}_B \mathbf{s}_B^H\} = \mathbf{I}_{N_S}$ and $E\{\mathbf{s}_R \mathbf{s}_R^H\} = P_{th} \mathbf{I}_{N_S}$, with P_{th} denoting the power of interference signals. And $\mathbf{H}_{SI} \in \mathbb{C}^{N_R \times N_R}$ stands for the SI channel matrix. The received signal at the destination node can be expressed as

$$\begin{aligned} \mathbf{y}_D = & \underbrace{\mathbf{W}^H \mathbf{H}_2 \mathbf{G}_2 \mathbf{G}_1^H \mathbf{H}_1 \mathbf{F}_S \mathbf{s}_B}_{\text{desired signal}} + \underbrace{\mathbf{W}^H \mathbf{H}_2 \mathbf{G}_2 \mathbf{G}_1^H \mathbf{H}_{SI} \mathbf{G}_2 \mathbf{s}_R}_{\text{SI}} \\ & + \underbrace{\mathbf{W}^H \mathbf{H}_2 \mathbf{G}_2 \mathbf{G}_1^H \mathbf{n}_R}_{\text{noise at the relay node}} + \underbrace{\mathbf{W}^H \mathbf{n}_D}_{\text{noise at the destination node}}. \end{aligned} \tag{25}$$

With the SI constraint $\mathbf{G}_1^H \mathbf{H}_{\text{SI}} \mathbf{G}_2 = \mathbf{0}$ set in the HBF design, \mathbf{y}_D can be further expressed as

$$\mathbf{y}_D = \mathbf{W}^H \mathbf{H}_2 \mathbf{G}_2 \mathbf{G}_1^H \mathbf{H}_1 \mathbf{F}_{\text{SB}} + \mathbf{W}^H \mathbf{H}_2 \mathbf{G}_2 \mathbf{G}_1^H \mathbf{n}_R + \mathbf{W}^H \mathbf{n}_D. \quad (26)$$

(2) *Imperfect CSI Model*: In practice, the CSI is usually not available accurately at every node because of user mobility and algorithm processing delay. The channel model with estimation errors is modelled as

$$\mathbf{H} = \bar{\mathbf{H}} + \Delta \mathbf{H}, \quad (27)$$

where \mathbf{H} and $\bar{\mathbf{H}}$ are the true and estimated channel matrixes, respectively. $\Delta \mathbf{H}$ denotes the channel estimation error, which is given by [14,15]

$$\Delta \mathbf{H} = \Phi^{\frac{1}{2}} \Delta \Theta^{\frac{1}{2}}, \quad (28)$$

where Θ and Φ are the covariance matrixes of estimation error at the receiver and transmitter sides. Δ denotes the unknown part of CSI, which is a complex Gaussian matrix with identically distributed zero mean and unit variance entries. Therefore, the true channel matrix satisfies $\mathbf{H} \sim \mathcal{CN}(\bar{\mathbf{H}}, \Theta \otimes \Phi)$.

In this correspondence, the exponential model is adopted to design the covariance matrixes of CSI error [14,15,32,33], i.e., $\Phi(i, j) = \sigma_e^2 \beta^{|i-j|}$, $\Theta(i, j) = \alpha^{|i-j|}$, with $\alpha, \beta \in \mathbb{R}$ being correlation coefficients and $\sigma_e^2 \in \mathbb{R}$ being estimation error variance. (i, j) represents the element at the intersection of row i and column j .

(3) *SI Channel Model*: In mmWave FD relay communication systems, SI consists of the LOS part and the NLOS part, where the former is much bigger than the latter on account of the path loss of mmWave signals. On one hand, NLOS SI is caused by the reflection of transmit signals on adjacent obstacles. Typically, the reflection channel $\mathbf{H}_{\text{NLOS}} \in \mathbb{C}^{N_R \times N_R}$ can be modelled as (2), i.e., the mmWave channel [15]. On the other hand, the direct path SI is caused by transmit signals of the FD relay node through the LOS channel $\mathbf{H}_{\text{LOS}} \in \mathbb{C}^{N_R \times N_R}$ between the transmitter–receiver link. According to [34], it can be formulated as the near-field model, whose $(m, n)_{th}$ element can be expressed as

$$\mathbf{H}_{\text{LOS}}^{(m,n)} = \mu_{mn} e^{-j \frac{2\pi}{\lambda} d_{mn}}, \quad (29)$$

where μ_{mn} is the normalization factor, d_{mn} denotes the distance between the m_{th} receive antenna and the n_{th} transmit antenna. Consequently, the SI channel can be modelled as

$$\mathbf{H}_{\text{SI}} = \sqrt{\frac{\kappa}{1+\kappa}} \mathbf{H}_{\text{LOS}} + \sqrt{\frac{1}{1+\kappa}} \mathbf{H}_{\text{NLOS}}, \quad (30)$$

with the scalar κ being the Rician factor.

In practice, the transmit and receive arrays of the FD AF relay node are close and fixed. In this paper, it is assumed that there is an accurate estimation for \mathbf{H}_{LOS} while uncertainty for \mathbf{H}_{NLOS} [35]. As a result, the estimated SI channel is given by [30]

$$\bar{\mathbf{H}}_{\text{SI}} = \sqrt{\frac{\kappa}{1+\kappa}} \mathbf{H}_{\text{LOS}} + \sqrt{\frac{1}{1+\kappa}} \bar{\mathbf{H}}_{\text{NLOS}}, \quad (31)$$

where $\bar{\mathbf{H}}_{\text{NLOS}}$ denotes the estimated reflection channel.

Due to the channel estimation errors, the SI constraint should be transformed into $\mathbf{G}_1^H \bar{\mathbf{H}}_{\text{SI}} \mathbf{G}_2 = \mathbf{0}$, which means the LOS part can be completely eliminated, while the NOLS part cannot. Considering the NLOS SI is much weaker than the LOS SI, we ignore the influence of residual SI in the subsequent problem formulation and algorithm design [30,35], but its impact on spectral efficiency will be shown in the simulation results.

4.2. Proposed HBF-MO-FD-R Algorithm

In this subsection, the IMSE criterion is still adopted in the HBF design for each node in the FD AF relay communication system, where the optimization objective can be written as

$$\text{IMSE} = E_{\Delta\mathbf{H}_1, \Delta\mathbf{H}_2} \left\{ \text{tr} \left(\mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \mathbf{F}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} - \mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} - \overline{\mathbf{F}}^H \mathbf{H}_1^H \overline{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} + \gamma_1^{-2} \sigma_R^2 \mathbf{W}^H \mathbf{H}_2 \overline{\mathbf{G}} \mathbf{G}^H \mathbf{H}_2^H \mathbf{W} + \gamma_1^{-2} \gamma_2^{-2} \sigma_D^2 \mathbf{W}^H \mathbf{W} + \mathbf{I}_{N_S} \right) \right\}. \quad (32)$$

When a matrix \mathbf{X} satisfies $\mathbf{X} \sim \mathcal{CN}(\overline{\mathbf{X}}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$, we have [36]

$$E \{ \mathbf{X} \mathbf{C} \mathbf{X}^H \} = \overline{\mathbf{X}} \mathbf{C} \overline{\mathbf{X}}^H + \text{tr}(\mathbf{C} \mathbf{\Sigma}^T) \mathbf{\Psi}, \quad (33)$$

$$E \{ \mathbf{X}^H \mathbf{C} \mathbf{X} \} = \overline{\mathbf{X}}^H \mathbf{C} \overline{\mathbf{X}} + \text{tr}(\mathbf{\Psi} \mathbf{C}) \mathbf{\Sigma}^T. \quad (34)$$

Since the error matrixes $\Delta\mathbf{H}_1$ and $\Delta\mathbf{H}_2$ are independent from each other, we can rewrite IMSE as

$$\begin{aligned} \text{IMSE} = & \text{tr} \left[\mathbf{W}^H \overline{\mathbf{H}}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \mathbf{F}^H \overline{\mathbf{H}}_1^H \overline{\mathbf{G}}^H \overline{\mathbf{H}}_2^H \mathbf{W} + \text{tr} \left(\overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \mathbf{F}^H \overline{\mathbf{H}}_1^H \overline{\mathbf{G}}^H \mathbf{\Theta}_2 \right) \mathbf{W}^H \mathbf{\Phi}_2 \mathbf{W} \right. \\ & + \text{tr} \left(\overline{\mathbf{F}} \mathbf{F}^H \mathbf{\Theta}_1 \right) \mathbf{W}^H \overline{\mathbf{H}}_2 \overline{\mathbf{G}} \mathbf{\Phi}_1 \overline{\mathbf{G}}^H \overline{\mathbf{H}}_2^H \mathbf{W} + \text{tr} \left(\overline{\mathbf{F}} \mathbf{F}^H \mathbf{\Theta}_1 \right) \text{tr} \left(\overline{\mathbf{G}} \mathbf{\Phi}_1 \overline{\mathbf{G}}^H \mathbf{\Theta}_2 \right) \mathbf{W}^H \mathbf{\Phi}_2 \mathbf{W} \\ & - \mathbf{W}^H \overline{\mathbf{H}}_2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} - \overline{\mathbf{F}}^H \overline{\mathbf{H}}_1^H \overline{\mathbf{G}}^H \overline{\mathbf{H}}_2^H \mathbf{W} + \gamma_1^{-2} \sigma_R^2 \mathbf{W}^H \overline{\mathbf{H}}_2 \overline{\mathbf{G}} \mathbf{G}^H \overline{\mathbf{H}}_2^H \mathbf{W} \\ & \left. + \gamma_1^{-2} \sigma_R^2 \text{tr} \left(\overline{\mathbf{G}} \mathbf{G}^H \mathbf{\Theta}_2 \right) \mathbf{W}^H \mathbf{\Phi}_2 \mathbf{W} + \gamma_1^{-2} \gamma_2^{-2} \sigma_D^2 \mathbf{W}^H \mathbf{W} + \mathbf{I}_{N_S} \right] \end{aligned} \quad (35)$$

in view of (33) and (34). Therefore, the IMSE problem can be formulated as

$$\min_{\gamma_1, \gamma_2, \mathbf{F}_{\text{RF}}, \mathbf{F}_U, \mathbf{G}_{\text{RF}_1}, \mathbf{G}_{\text{B}_1}, \mathbf{G}_{\text{RF}_2}, \mathbf{G}_{\text{U}_2}, \mathbf{W}_{\text{RF}}, \mathbf{W}_B} \text{IMSE} \quad (36a)$$

$$\text{s.t.} \quad \left| [\mathbf{F}_{\text{RF}}]_{a,b} \right|^2 = 1, \forall a, b, \quad (36b)$$

$$\left| [\mathbf{G}_{\text{RF}_1}]_{c,d} \right|^2 = 1, \forall c, d, \quad (36c)$$

$$\left| [\mathbf{G}_{\text{RF}_2}]_{e,f} \right|^2 = 1, \forall e, f, \quad (36d)$$

$$\left| [\mathbf{W}_{\text{RF}}]_{g,h} \right|^2 = 1, \forall g, h, \quad (36e)$$

$$E \left\{ \|\mathbf{x}_B\|_F^2 \right\} = P_B, \quad (36f)$$

$$E_{\Delta\mathbf{H}_1} \left\{ \|\mathbf{x}_R\|_F^2 \right\} = P_R, \quad (36g)$$

$$\mathbf{G}_1^H \overline{\mathbf{H}}_{\text{SI}} \mathbf{G}_2 = \mathbf{0}, \quad (36h)$$

where (36f) and (36g) denote the source and AF relay power constraints. Specifically, the former has the same form as that in Section III, while the latter can be given by $E_{\Delta\mathbf{H}_1} \left\{ \|\mathbf{x}_R\|_F^2 \right\} = \gamma_2^2 \text{tr} \left[\gamma_1^2 \overline{\mathbf{G}} \mathbf{H}_1 \overline{\mathbf{F}} \mathbf{F}^H \overline{\mathbf{H}}_1^H \overline{\mathbf{G}}^H + \gamma_1^2 \text{tr} \left(\overline{\mathbf{F}} \mathbf{F}^H \mathbf{\Theta}_1 \right) \overline{\mathbf{G}} \mathbf{\Phi}_1 \overline{\mathbf{G}}^H + \sigma_R^2 \overline{\mathbf{G}} \mathbf{G}^H \right]$.

(1) *The Source Node Design:* While fixing \mathbf{G} and \mathbf{W} , the HBF design sub-problem of the source node can be formulated as

$$\min_{\gamma_1, \mathbf{F}_{\text{RF}}, \mathbf{F}_U} \text{IMSE} \quad (37a)$$

$$\text{s.t.} \quad \left| [\mathbf{F}_{\text{RF}}]_{a,b} \right|^2 = 1, \forall a, b, \quad (37b)$$

$$E \left\{ \|\mathbf{x}_B\|_F^2 \right\} = P_B. \quad (37c)$$

with $\mathbf{F} = \gamma_1 \mathbf{F}_{\text{RF}} \mathbf{F}_U$, we have the power constraint as $\gamma_1^2 \text{tr} \left(\overline{\mathbf{F}} \mathbf{F}^H \right) = P_B$. Therefore, γ_1 can be expressed as

$$\gamma_1 = \left[\frac{\text{tr}(\mathbf{F}_{\text{RF}} \mathbf{F}_U \mathbf{F}_U^H \mathbf{F}_{\text{RF}}^H)}{P_B} \right]^{-\frac{1}{2}}. \quad (38)$$

Upon having fixed \mathbf{F}_{RF} , we can obtain \mathbf{F}_U by setting the partial derivatives of the Lagrange function $L1(\lambda_1, \mathbf{F}_U, \gamma_1)$ to zero:

$$\mathbf{F}_U = \left(\mathbf{F}_{\text{RF}}^H \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}} + \mathbf{C}_{11} \right)^{-1} \mathbf{F}_{\text{RF}}^H \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}, \quad (39)$$

where $\mathbf{C}_{11} = c_{11} \mathbf{F}_{\text{RF}}^H \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}} + c_{12} \mathbf{F}_{\text{RF}}^H \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W} + c_{13} \mathbf{F}_{\text{RF}}^H \mathbf{F}_{\text{RF}}$, $c_{11} = \text{tr}(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}})$, $c_{12} = \text{tr}(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}) + \text{tr}(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}}) \text{tr}(\bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H)$ and $c_{13} = \gamma_2^{-2} \frac{\sigma_D^2}{P_B} \text{tr}(\mathbf{W}^H \mathbf{W}) + \frac{\sigma_R^2}{P_B} \text{tr}(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}) + \frac{\sigma_D^2}{P_B} \text{tr}(\bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H) \text{tr}(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}})$. Substituting (38) and (39) into (37a), the optimization objective can be further expressed as

$$J1(\mathbf{F}_{\text{RF}}) = \text{tr} \left\{ \left[\mathbf{I}_{N_S} + \left(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}} \right) \mathbf{C}_{11}^{-1} \left(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}} \right)^H \right]^{-1} \right\}. \quad (40)$$

The sub-problem with only constant-modulus constraint can be re-formulated as

$$\min_{\mathbf{F}_{\text{RF}}} J1(\mathbf{F}_{\text{RF}}) \quad (41a)$$

$$\text{s.t.} \quad \left| [\mathbf{F}_{\text{RF}}]_{a,b} \right|^2 = 1, \forall a, b. \quad (41b)$$

By mathematical derivation, we can obtain the gradient of $J1(\mathbf{F}_{\text{RF}})$ with respect to \mathbf{F}_{RF} :

$$\begin{aligned} \nabla J1(\mathbf{F}_{\text{RF}}) &= \left(c_{11} \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{C}_{11}^{-1} \mathbf{F}_{\text{RF}}^H \right. \\ &\quad \left. + c_{12} \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{C}_{11}^{-1} \mathbf{F}_{\text{RF}}^H - \mathbf{I}_{N_T} + c_{13} \mathbf{F}_{\text{RF}} \mathbf{C}_{11}^{-1} \mathbf{F}_{\text{RF}}^H \right) \\ &\quad \times \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{C}_{11}^{-2} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \mathbf{F}_{\text{RF}} \mathbf{C}_{11}^{-1}. \end{aligned} \quad (42)$$

(2) *The Relay Node Design:* First, we consider the HBF design for the receiver of the relay node. It is assumed that \mathbf{F} , \mathbf{G}_2 and \mathbf{W} are fixed. Taking SI constraint out of consideration, the optimization problem can be formulated as

$$\min_{\mathbf{G}_{\text{RF}_1}, \mathbf{G}_{\text{B}_1}} \text{IMSE} \quad (43a)$$

$$\text{s.t.} \quad \left| [\mathbf{G}_{\text{RF}_1}]_{c,d} \right|^2 = 1, \forall c, d. \quad (43b)$$

with \mathbf{G}_{RF_1} fixed, the optimization problem (43) can be recast as an unconstrained optimization problem. By taking derivative, we have the closed-form solution of \mathbf{G}_{B_1} as

$$\begin{aligned} \mathbf{G}_{\text{B}_1} &= \left[\mathbf{G}_{\text{RF}_1}^H \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \mathbf{G}_{\text{RF}_1} + \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1) \mathbf{G}_{\text{RF}_1}^H \bar{\mathbf{H}}_1 \mathbf{G}_{\text{RF}_1} + \gamma_1^{-2} \sigma_R^2 \mathbf{G}_{\text{RF}_1}^H \mathbf{G}_{\text{RF}_1} \right]^{-1} \\ &\quad \times \left(\mathbf{G}_{\text{RF}_1}^H \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}}_2 \right) \left[\bar{\mathbf{G}}_2^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}}_2 + \text{tr}(\mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}}_2) \bar{\mathbf{G}}_2^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}}_2 \right]^{-1}. \end{aligned} \quad (44)$$

Substituting the optimization problem (43) with Equation (44), we have

$$\min_{\mathbf{G}_{\text{RF}_1}} Q1(\mathbf{G}_{\text{RF}_1}) \quad (45a)$$

$$\text{s.t.} \quad \left| [\mathbf{G}_{\text{RF}_1}]_{c,d} \right|^2 = 1, \forall c, d, \quad (45b)$$

$$Q1(\mathbf{G}_{\text{RF}_1}) = \text{tr} \left(\mathbf{I}_{N_{\text{RF}}} - \mathbf{D}_{11}^{-1} \mathbf{D}_{12} \mathbf{D}_{13}^{-1} \mathbf{D}_{12}^H \right), \quad (46)$$

with $\mathbf{D}_{11} = \mathbf{G}_{\text{RF}_1}^H \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \mathbf{G}_{\text{RF}_1} + \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\Theta}_1) \mathbf{G}_{\text{RF}_1}^H \boldsymbol{\Phi}_1 \mathbf{G}_{\text{RF}_1} + \gamma_1^{-2} \sigma_{\text{R}}^2 \mathbf{G}_{\text{RF}_1}^H \mathbf{G}_{\text{RF}_1}$, $\mathbf{D}_{12} = \mathbf{G}_{\text{RF}_1}^H \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}}_2$ and $\mathbf{D}_{13} = \bar{\mathbf{G}}_2^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}}_2 + \text{tr}(\mathbf{W}^H \boldsymbol{\Phi}_2 \mathbf{W}) \bar{\mathbf{G}}_2^H \boldsymbol{\Theta}_2 \bar{\mathbf{G}}_2$. The Euclidean gradient of $\text{Q1}(\mathbf{G}_{\text{RF}_1})$ is given as

$$\begin{aligned} \nabla \text{Q1}(\mathbf{G}_{\text{RF}_1}) &= \left[\bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \mathbf{G}_{\text{RF}_1} \mathbf{D}_{11}^{-1} \mathbf{D}_{12} \right. \\ &\quad + \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\Theta}_1) \boldsymbol{\Phi}_1 \mathbf{G}_{\text{RF}_1} \mathbf{D}_{11}^{-1} \mathbf{D}_{12} \\ &\quad + \gamma_1^{-2} \sigma_{\text{R}}^2 \mathbf{G}_{\text{RF}_1} \mathbf{D}_{11}^{-1} \mathbf{D}_{12} \\ &\quad \left. - \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \mathbf{W}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}}_2 \right] \mathbf{D}_{13}^{-1} \mathbf{D}_{12}^H \mathbf{D}_{11}^{-1}. \end{aligned} \tag{47}$$

After finishing the design of the receiver, we turn to the transmitter. For any given \mathbf{F} , \mathbf{G}_1 and \mathbf{W} , we have the optimization sub-problem as follows by ignoring the SI constraint:

$$\min_{\gamma_2, \mathbf{G}_{\text{RF}_2}, \mathbf{G}_{\text{U}_2}} \text{IMSE} \tag{48a}$$

$$\text{s.t.} \quad \left| [\mathbf{G}_{\text{RF}_2}]_{c,d} \right|^2 = 1, \forall c, d, \tag{48b}$$

$$\text{E}_{\Delta \mathbf{H}_1} \left\{ \|\mathbf{x}_{\text{R}}\|_{\text{F}}^2 \right\} = P_{\text{R}}. \tag{48c}$$

Likewise, we can obtain γ_2 by taking transposition of (48b), which is given as

$$\begin{aligned} \gamma_2 &= \left[\text{tr} \left(\gamma_1^2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \bar{\mathbf{G}}^H \right. \right. \\ &\quad \left. \left. + \gamma_1^2 \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\Theta}_1) \bar{\mathbf{G}} \boldsymbol{\Phi}_1 \bar{\mathbf{G}}^H + \sigma_{\text{R}}^2 \bar{\mathbf{G}} \bar{\mathbf{G}}^H \right) / P_{\text{R}} \right]^{-\frac{1}{2}} \end{aligned} \tag{49}$$

To this end, the optimization problem (48) can be equivalently recast as an optimization problem (50) with only power constraint by fixing \mathbf{G}_{RF_2} . Based on the Lagrange function, the closed form of \mathbf{G}_{U_2} is expressed as (51).

$$\min_{\gamma_2, \mathbf{G}_{\text{U}_2}} \text{IMSE} \tag{50a}$$

$$\text{s.t.} \quad \text{s.t.} \text{E}_{\Delta \mathbf{H}_1} \left\{ \|\mathbf{x}_{\text{R}}\|_{\text{F}}^2 \right\} = P_{\text{R}}, \tag{50b}$$

$$\begin{aligned} \mathbf{G}_{\text{U}_2} &= \left[\mathbf{G}_{\text{RF}_2}^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{W}^H \bar{\mathbf{H}}_2 \mathbf{G}_{\text{RF}_2} + \text{tr}(\mathbf{W}^H \boldsymbol{\Phi}_2 \mathbf{W}) \mathbf{G}_{\text{RF}_2}^H \boldsymbol{\Theta}_2 \mathbf{G}_{\text{RF}_2} \right. \\ &\quad \left. + \frac{\sigma_{\text{D}}^2}{P_{\text{R}}} \text{tr}(\mathbf{W}^H \mathbf{W}) \mathbf{G}_{\text{RF}_2}^H \mathbf{G}_{\text{RF}_2} \right]^{-1} \left(\mathbf{G}_{\text{RF}_2}^H \bar{\mathbf{H}}_2^H \mathbf{W} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \mathbf{G}_1 \right) \\ &\quad \times \left[\mathbf{G}_1^H \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \mathbf{G}_1 + \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\Theta}_1) \mathbf{G}_1^H \boldsymbol{\Phi}_1 \mathbf{G}_1 + \gamma_1^{-2} \sigma_{\text{R}}^2 \mathbf{G}_1^H \mathbf{G}_1 \right]^{-1}, \end{aligned} \tag{51}$$

where we define $\mathbf{D}_{23} = \mathbf{G}_1^H \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \mathbf{G}_1 + \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\Theta}_1) \mathbf{G}_1^H \boldsymbol{\Phi}_1 \mathbf{G}_1 + \gamma_1^{-2} \sigma_{\text{R}}^2 \mathbf{G}_1^H \mathbf{G}_1$, $\mathbf{D}_{21} = \mathbf{G}_{\text{RF}_2}^H \bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{W}^H \bar{\mathbf{H}}_2 \mathbf{G}_{\text{RF}_2} + \text{tr}(\mathbf{W}^H \boldsymbol{\Phi}_2 \mathbf{W}) \mathbf{G}_{\text{RF}_2}^H \boldsymbol{\Theta}_2 \mathbf{G}_{\text{RF}_2} + \frac{\sigma_{\text{D}}^2}{P_{\text{R}}} \text{tr}(\mathbf{W}^H \mathbf{W}) \mathbf{G}_{\text{RF}_2}^H \mathbf{G}_{\text{RF}_2}$ and $\mathbf{D}_{22} = \mathbf{G}_{\text{RF}_2}^H \bar{\mathbf{H}}_2^H \mathbf{W} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1^H \mathbf{G}_1$. One further step, we can get the simplified optimization objective (52), the optimization problem (53) and its gradient (54) in a similar way.

$$\text{Q2}(\mathbf{G}_{\text{RF}_2}) = \text{tr} \left(\mathbf{I}_{N_{\text{RF}}} - \mathbf{D}_{21}^{-1} \mathbf{D}_{22} \mathbf{D}_{23}^{-1} \mathbf{D}_{22}^H \right), \tag{52}$$

$$\min_{\mathbf{G}_{\text{RF}_2}} \text{Q2}(\mathbf{G}_{\text{RF}_2}) \tag{53a}$$

$$\text{s.t.} \quad \left| [\mathbf{G}_{\text{RF}_2}]_{e,f} \right|^2 = 1, \forall e, f, \tag{53b}$$

$$\begin{aligned} \nabla Q2(\mathbf{G}_{\text{RF}_2}) = & \left[\bar{\mathbf{H}}_2^H \mathbf{W} \mathbf{W}^H \bar{\mathbf{H}}_2 \mathbf{G}_{\text{RF}_2} \mathbf{D}_{21}^{-1} \mathbf{D}_{22} + \text{tr}(\mathbf{W}^H \Phi_2 \mathbf{W}) \Theta_2 \mathbf{G}_{\text{RF}_2} \mathbf{D}_{21}^{-1} \mathbf{D}_{22} \right. \\ & \left. + \frac{\sigma_D^2}{P_R} \text{tr}(\mathbf{W}^H \mathbf{W}) \mathbf{G}_{\text{RF}_2} \mathbf{D}_{21}^{-1} \mathbf{D}_{22} - \bar{\mathbf{H}}_2^H \mathbf{W} \bar{\mathbf{F}}^H \mathbf{H}_1^H \mathbf{G}_1 \right] \mathbf{D}_{23}^{-1} \mathbf{D}_{22}^H \mathbf{D}_{21}^{-1}. \end{aligned} \quad (54)$$

It's worth noting that we temporarily neglect the SI constraint in the HBF design of the relay node. Consequently, we need to design the SI suppression matrix to satisfy $\mathbf{G}_1^H \bar{\mathbf{H}}_{\text{SI}} \mathbf{G}_2 = \mathbf{0}$. Severe SI acts as a key factor to affect the performance of FD communication systems. There are strict restrictions on the number of RF chains in traditional SI suppression algorithm based on NP. In [31], it is required to meet $N_{\text{RF},T} \geq N_S + N_{\text{RF},R}$, with $N_{\text{RF},T}$ and $N_{\text{RF},R}$ being the number of the transmit and receive RF chains at the AF relay node, while [35] needs to satisfy $N_{\text{RF},R} \geq 2N_S$. The authors of [30] proposed an SI suppression algorithm suitable for the system with $N_{\text{RF},T} = N_{\text{RF},R} = N_S$. Nevertheless, its performance remains to be improved. On the basis of the above literature, we further develop the SIC-NP algorithm.

We first construct an equivalent SI channel matrix $\bar{\mathbf{H}}_{\text{SI-EQ}} \in \mathbb{C}^{N_S \times N_R}$, which can be written as

$$\bar{\mathbf{H}}_{\text{SI-EQ}} = \mathbf{G}_1^H \bar{\mathbf{H}}_{\text{SI}}. \quad (55)$$

It is obvious that the rank of $\bar{\mathbf{H}}_{\text{SI-EQ}}$ is $N_S (N_R \gg N_S)$, so its null-space definitely exists. There are a number of ways to obtain the null-space of a matrix. We adopt the singular value decomposition (SVD) method in this paper, which yields

$$\bar{\mathbf{H}}_{\text{SI-EQ}} = \bar{\mathbf{U}}_{\text{SI-EQ}} \bar{\Sigma}_{\text{SI-EQ}} \bar{\mathbf{V}}_{\text{SI-EQ}}^H, \quad (56)$$

where $\bar{\mathbf{U}}_{\text{SI-EQ}} \in \mathbb{C}^{N_S \times N_S}$ and $\bar{\mathbf{V}}_{\text{SI-EQ}} \in \mathbb{C}^{N_R \times N_R}$ are unitary matrixes, $\bar{\Sigma}_{\text{SI-EQ}} \in \mathbb{C}^{N_S \times N_R}$ is a diagonal matrix. $\bar{\mathbf{V}}_{\text{SI-EQ}}$ can be further decomposed as $\bar{\mathbf{V}}_{\text{SI-EQ}} = [\bar{\mathbf{V}}_{\text{SI-EQ}}^1 \bar{\mathbf{V}}_{\text{SI-EQ}}^0]$, where $\bar{\mathbf{V}}_{\text{SI-EQ}}^1 \in \mathbb{C}^{N_R \times N_S}$ and $\bar{\mathbf{V}}_{\text{SI-EQ}}^0 \in \mathbb{C}^{N_R \times (N_R - N_S)}$ are the right singular matrix for non-zero singular value and zero singular value. And the latter satisfies

$$\bar{\mathbf{H}}_{\text{SI-EQ}} \bar{\mathbf{V}}_{\text{SI-EQ}}^0 = \mathbf{0}. \quad (57)$$

In this correspondence, we can construct the null-space matrix $\mathbf{P} \in \mathbb{C}^{N_R \times N_R}$ of $\bar{\mathbf{H}}_{\text{SI-EQ}}$ using the column vectors of $\bar{\mathbf{V}}_{\text{SI-EQ}}^0$. Subsequently, \mathbf{G}_2 is projected to \mathbf{P} to mitigate the SI. The projection matrix \mathbf{P}_1 and SI suppression matrix $\mathbf{G}_2^{\text{SIC}}$ can be written as

$$\mathbf{P}_1 = \mathbf{P} (\mathbf{P}^H \mathbf{P})^{-1} \mathbf{P}^H, \quad (58)$$

$$\mathbf{G}_2^{\text{SIC}} = \mathbf{P}_1 \mathbf{G}_2. \quad (59)$$

Hence, the HBF matrix of FD AF relay node takes the form of

$$\mathbf{G} = \mathbf{G}_2^{\text{SIC}} \mathbf{G}_1^H. \quad (60)$$

For clarity, the above SI suppression procedure is summarized in Algorithm 2.

Algorithm 2 SIC-NP Algorithm.

Input: $\mathbf{G}_1, \mathbf{G}_2, \bar{\mathbf{H}}_{\text{SI}}$.

Output: $\mathbf{G}_2^{\text{SIC}}$.

- 1: Calculate the equivalent SI channel matrix $\bar{\mathbf{H}}_{\text{SI-EQ}}$ according to (55).
 - 2: Calculate $\bar{\mathbf{V}}_{\text{SI-EQ}}^0$ by taking SVD of $\bar{\mathbf{H}}_{\text{SI-EQ}}$.
 - 3: Construct the null-space matrix \mathbf{P} .
 - 4: Calculate \mathbf{P}_1 according to (58).
 - 5: Compute $\mathbf{G}_2^{\text{SIC}}$ according to (59).
-

(3) *The Destination Node Design:* Similar to that of Section II, the optimization sub-problem of the destination node has only constant-modulus constraint. For any given source and relay HBF matrixes, the destination sub-problem can be formulated as

$$\min_{\mathbf{W}_{\text{RF}}, \mathbf{W}_{\text{B}}} \text{IMSE} \tag{61a}$$

$$\text{s.t.} \quad \left| [\mathbf{W}_{\text{RF}}]_{g,h} \right|^2 = 1, \forall g, h. \tag{61b}$$

with \mathbf{W}_{RF} fixed, the problem (61) can be transformed into an unconstrained optimization problem. By setting the gradient of the objective function to zero, we can obtain

$$\mathbf{W}_{\text{B}} = \left(\mathbf{W}_{\text{RF}}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}_{\text{RF}} + \mathbf{C}_{31} \right)^{-1} \mathbf{W}_{\text{RF}}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{F}}, \tag{62}$$

where $c_{31} = \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\Theta}_1)$, $c_{32} = \gamma_1^{-2} \sigma_{\text{R}}^2$, $c_{33} = \text{tr}(\bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \boldsymbol{\Theta}_2) + \text{tr}(\bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\Theta}_1) \text{tr}(\bar{\mathbf{G}} \boldsymbol{\Phi}_1 \bar{\mathbf{G}}^H \boldsymbol{\Theta}_2) + \gamma_1^{-2} \sigma_{\text{R}}^2 \text{tr}(\bar{\mathbf{G}} \bar{\mathbf{G}}^H \boldsymbol{\Theta}_2)$, $c_{34} = \gamma_1^{-2} \gamma_2^{-2} \sigma_{\text{D}}^2$, and we define

$$\begin{aligned} \mathbf{C}_{31} &= c_{31} \mathbf{W}_{\text{RF}}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \boldsymbol{\Phi}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}_{\text{RF}} + c_{32} \mathbf{W}_{\text{RF}}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}_{\text{RF}} \\ &+ c_{33} \mathbf{W}_{\text{RF}}^H \boldsymbol{\Phi}_2 \mathbf{W}_{\text{RF}} + c_{34} \mathbf{W}_{\text{RF}}^H \mathbf{W}_{\text{RF}}. \end{aligned} \tag{63}$$

Substituting (62) into the equation of IMSE, the objective function is equivalent to

$$\text{J3}(\mathbf{W}_{\text{RF}}) = \text{tr} \left[\left(\mathbf{I}_{N_{\text{S}}} + \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}_{\text{RF}} \mathbf{C}_{31}^{-1} \mathbf{W}_{\text{RF}}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \right)^{-1} \right]. \tag{64}$$

In the meantime, the optimization sub-problem is reduced to

$$\min_{\mathbf{W}_{\text{RF}}} \text{J3}(\mathbf{W}_{\text{RF}}) \tag{65a}$$

$$\text{s.t.} \quad \left| [\mathbf{W}_{\text{RF}}]_{g,h} \right|^2 = 1, \forall g, h. \tag{65b}$$

Finally, the gradient $\nabla \text{J3}(\mathbf{W}_{\text{RF}})$ can be written as

$$\begin{aligned} \nabla \text{J3}(\mathbf{W}_{\text{RF}}) &= \left[\left(c_{33} \boldsymbol{\Phi}_2 + c_{31} \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \boldsymbol{\Phi}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H + c_{32} \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H + c_{34} \mathbf{I}_{N_{\text{D}}} \right) \right. \\ &\quad \left. \times \mathbf{W}_{\text{RF}} \mathbf{C}_{31}^{-1} \mathbf{W}_{\text{RF}}^H - \mathbf{I}_{N_{\text{D}}} \right] \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{F}} \mathbf{C}_{32}^{-2} \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}_{\text{RF}} \mathbf{C}_{31}^{-1}, \end{aligned} \tag{66}$$

with $\mathbf{C}_{32} = \mathbf{I}_{N_{\text{S}}} + \bar{\mathbf{F}}^H \bar{\mathbf{H}}_1 \bar{\mathbf{G}}^H \bar{\mathbf{H}}_2^H \mathbf{W}_{\text{RF}} \mathbf{C}_{31}^{-1} \mathbf{W}_{\text{RF}}^H \bar{\mathbf{H}}_2 \bar{\mathbf{G}} \bar{\mathbf{H}}_1 \bar{\mathbf{F}}$.

The proposed MO based HBF algorithm for the FD relay system is summarized in Algorithm 3.

Algorithm 3 HBF-MO-FD-R Algorithm.**Input:** $\bar{\mathbf{H}}_1, \bar{\mathbf{H}}_2, \bar{\mathbf{H}}_{\text{SI}}, \sigma_{\text{R}}^2, \sigma_{\text{D}}^2, \Phi_1, \Phi_2, \Theta_1, \Theta_2$.**Output:** $\gamma_1, \gamma_2, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{U}}, \mathbf{G}_1, \mathbf{G}_2^{\text{SIC}}, \mathbf{W}_{\text{RF}}, \mathbf{W}_{\text{B}}$.

- 1: **Initialize** $\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{U}}, \mathbf{G}_{\text{RF}1}, \mathbf{G}_{\text{RF}2}, \mathbf{G}_{\text{B}1}, \mathbf{G}_{\text{U}2}, \mathbf{W}_{\text{RF}}$ and \mathbf{W}_{B} randomly, and set $\gamma_1 = \gamma_2 = 1, k = 1$.
- 2: **while** $\text{IMSE}_k - \text{IMSE}_{k-1} > \varepsilon$ **do**
- 3: Calculate $\nabla \text{J1}(\mathbf{F}_{\text{RF},k})$ according to (42) and compute $\mathbf{F}_{\text{RF},k+1}$ by using the MO method.
- 4: Update $\gamma_{1,k+1}, \mathbf{F}_{\text{U},k+1}$ and $\bar{\mathbf{F}}_{k+1}$ according to (38) and (39).
- 5: Calculate $\nabla \text{Q1}(\mathbf{G}_{\text{RF}1,k})$ according to (47) and compute $\mathbf{G}_{\text{RF}1,k+1}$ by using the MO method.
- 6: Update $\mathbf{G}_{\text{B}1,k+1}$ and $\mathbf{G}_{1,k+1}$ according to (44).
- 7: Calculate $\nabla \text{Q2}(\mathbf{G}_{\text{RF}2,k})$ according to (54) and compute $\mathbf{G}_{\text{RF}2,k+1}$ by using the MO method.
- 8: Update $\gamma_{2,k+1}, \mathbf{G}_{\text{U}2,k+1}$ and $\bar{\mathbf{G}}_{2,k+1}$ according to (49) and (51).
- 9: Calculate $\nabla \text{J3}(\mathbf{W}_{\text{RF},k})$ according to (66) and compute $\mathbf{W}_{\text{RF},k+1}$ by using the MO method.
- 10: Update $\mathbf{W}_{\text{B},k+1}$ and \mathbf{W}_{k+1} according to (62).
- 11: $k \leftarrow k + 1$.
- 12: **end while**
- 13: Compute $\mathbf{G}_2^{\text{SIC}}$ by using the SIC-NP algorithm and update \mathbf{G} according to (60).
- 14: Power normalization.

5. Simulation Results

To evaluate the performance of the proposed HBF algorithms, simulation results are presented in this section. In the simulations, it is assumed that the SNRs between the relay–destination link and the source-relay link are identical. Unless otherwise mentioned, the number of antennas at the source, relay and destination nodes are $N_{\text{T}} = 64, N_{\text{R}} = 48$ and $N_{\text{D}} = 32$, respectively. The number of RF chains and data streams are set as $N_{\text{RF}} = N_{\text{S}} = 2$. Due to the limited scattering characteristics of mmWave channels, we assume $L = 10$. The Rician factor is chosen as 20 dB. All the AoDs and AoAs are independently and uniformly distributed in $[0, 2\pi]$. Moreover, we define the same source and relay transmit power, i.e., $P_{\text{B}} = P_{\text{R}} = 2$. In this paper, all simulation results are averaged over 1000 mmWave channels, which are generated randomly.

5.1. HBF-MO

Firstly, we present the performance of the HBF-MO algorithm (denoted as *HBF-MO*) for the HD relay system. In this subsection, we will consider a benchmark approach and two other different algorithms, which are the optimal digital processing based on SVD (denoted as *DBF*), the HBF algorithm in [14] (denoted as *HBF*) and the OMP based algorithm in [7] (denoted as *OMP*). Unless otherwise specified, the convergence parameter of the outer loop is set to 10^{-4} , and that of the inner loop is set to 10^{-6} .

Figure 3 validates the convergence of SE and MSE with SNR set to 6dB. To be specific, the number of outer iterations is fixed to 15. As can be observed, our proposed algorithm converges to an optimal solution within about four iterations. SE increases with the growth of iteration number and converges to almost 5.485 bit/s/Hz, while MSE decreases and maintain at 0.0449 after 15 iterations.

Next, the SE performance versus SNR of our proposed HBF-MO algorithm and three other approaches is investigated in Figure 4. As shown in the figure, the SE of all algorithms increases with growing SNR. Taking the SE of DBF method as a benchmark, the proposed algorithm prevails over two other methods by approaching the DBF algorithm, which is more obvious as SNR becomes higher. When SNR = 10 dB, our HBF-MO algorithm can provide a gain of 7.99% over HBF [14] and 11% over OMP [7].

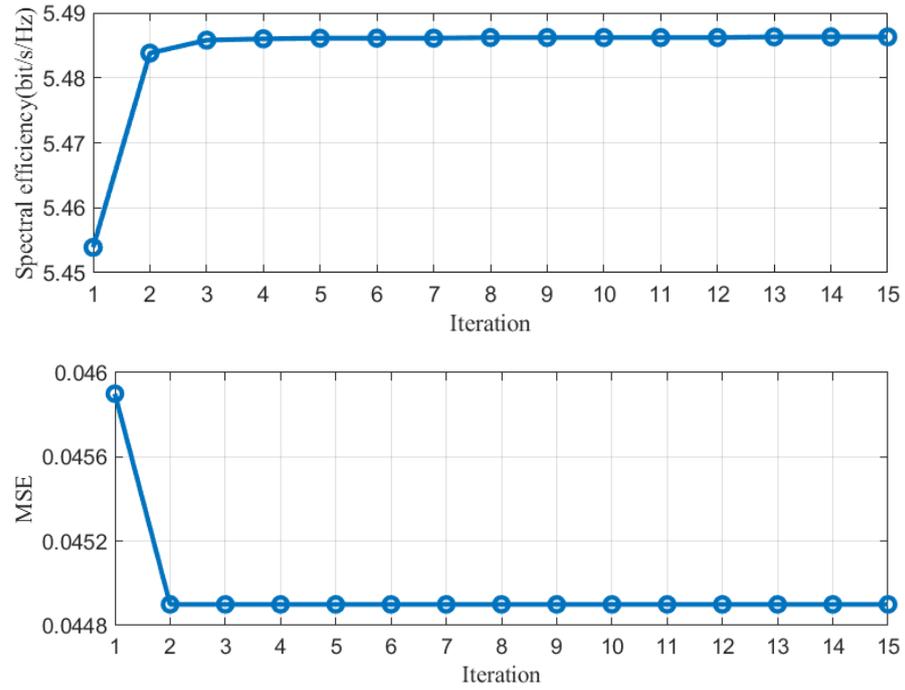


Figure 3. Convergence of SE and MSE for the HBF-MO algorithm.

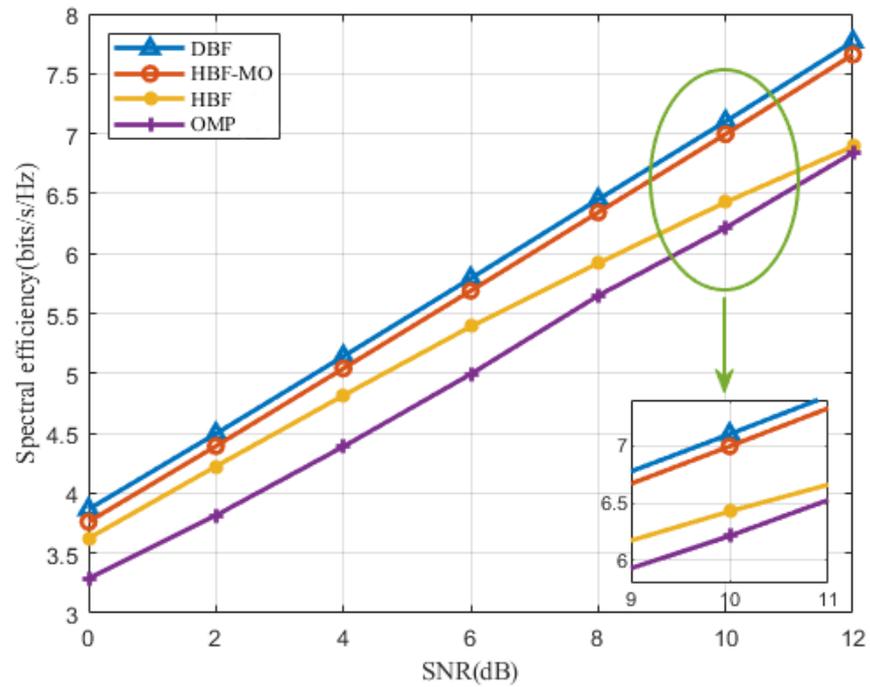


Figure 4. SE versus SNR for different beamforming algorithms.

By varying the number of relay and destination antennas, the SE performance of all algorithms are evaluated when SNR = 6dB in Figures 5 and 6. Likewise, the DBF method is adopted as a benchmark. Due to the additional antenna gain, the spectral efficiency of all algorithms improves with the number of antennas increasing. Still, our proposed algorithm yields the best performance over the others.

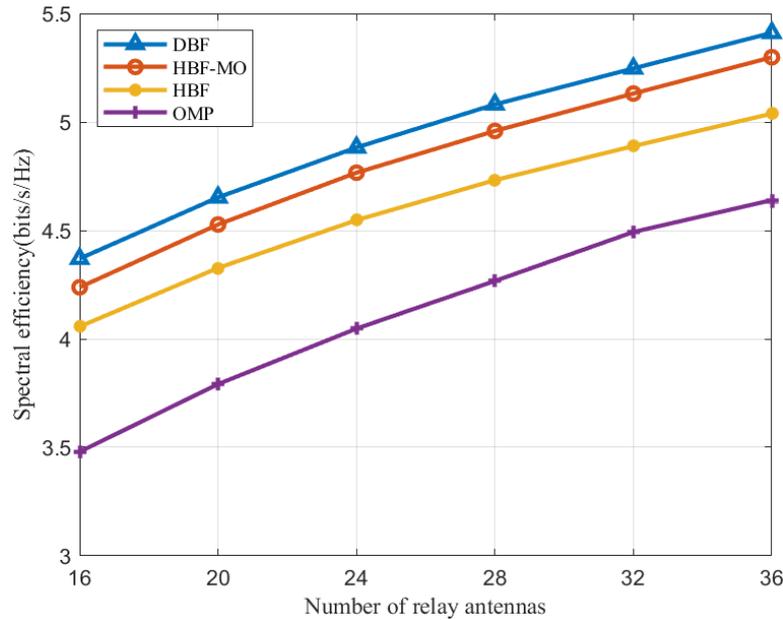


Figure 5. SE comparison with different relay antennas with $N_T = 64$ and $N_D = 32$.

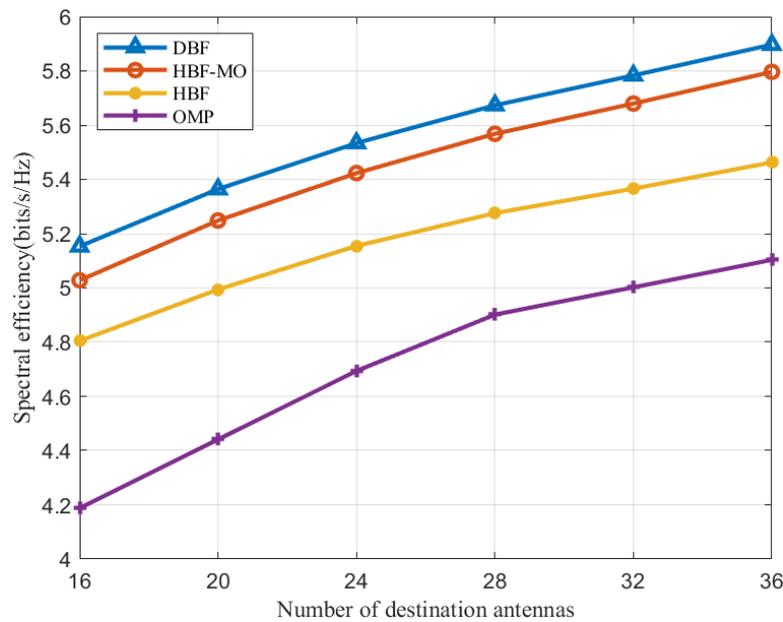


Figure 6. SE comparison with different destination antennas with $N_T = 64$ and $N_R = 48$.

5.2. HBF-MO-FD-R

In this subsection, simulation results for the proposed robust HBF scheme for the FD relay system is presented. For notation convenience, “HBF-MO-FD-R” represents our proposed method, “HBF-FD-SIC” denotes the HBF-MO-FD-R algorithm with SIC-NP replaced by the SIC method in [30], “HBF-FD-NSIC” stands for the proposed method without SIC-NP, while “HBF-MO-HD” denotes the proposed HD algorithm with imperfect CSI. Additionally, it is assumed that $a_0 = b_0 = \psi = 0$ [15]. Unless otherwise stated, the parameters related to channel estimation errors are set as $\alpha = 0.6$, $\beta = 0.4$ and $\sigma_e^2 = 0.5$. Meanwhile, the Interference to Noise Ratio (INR) is defined as $INR = 10 \lg(P_{th}/\sigma_R^2)$ [30,35].

In Figure 7, we show the convergence performance of HBF-MO-FD-R with respect to iterations at SNR = 6dB and INR = 10dB, where the maximum number of iterations is set to 15. It can be observed that SE converges to a maximum of above 2.5 bits/s/Hz in

about seven iterations. On the other hand, MSE decreases to a minimum of almost 0.18 after about five iterations.

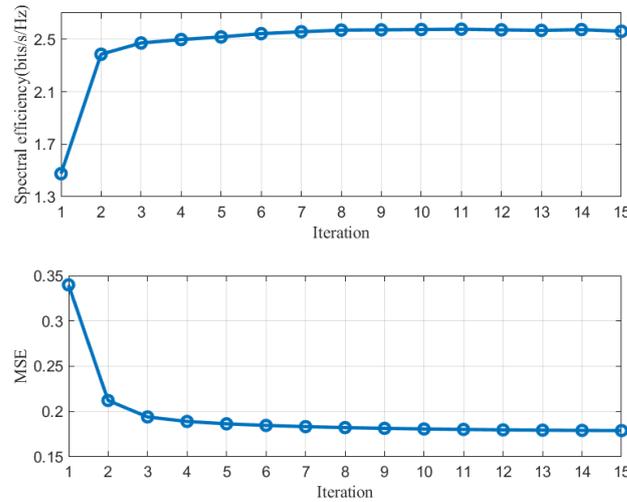


Figure 7. Convergence of SE and MSE for the HBF-MO-FD-R algorithm.

Figure 8 compares the SE performance of four different algorithms. As expected, SE of all approaches improves monotonically with increasing SNR. It can be obviously seen that both the HBF-MO-FD-R algorithm and the HBF-FD-SIC algorithm obtain higher SE by suppressing the LOS SI in contrast with the method without SIC. Additionally, the performance of the two methods is still affected by the residual SI because the NLOS channel cannot be accurately estimated. As is shown in the figure, with the increase in INR, the SE of HBF-MO-FD-R and HBF-FD-SIC decreases, but that is still higher than the non-SIC method. Meanwhile, we find that our proposed algorithm outperforms the HBF-FD-SIC approach. The average gap between two approaches is about 0.913 bit/s/Hz when INR = 10 dB and 0.791 bit/s/Hz for INR = 20 dB. Moreover, the non-robust algorithm can obtain more SE than the non-SIC algorithm when SI is severe (INR = 20 dB). It is noteworthy that the value of INR has no impact on the performance of HBF-MO-HD because there is no SI channel in HD mode.

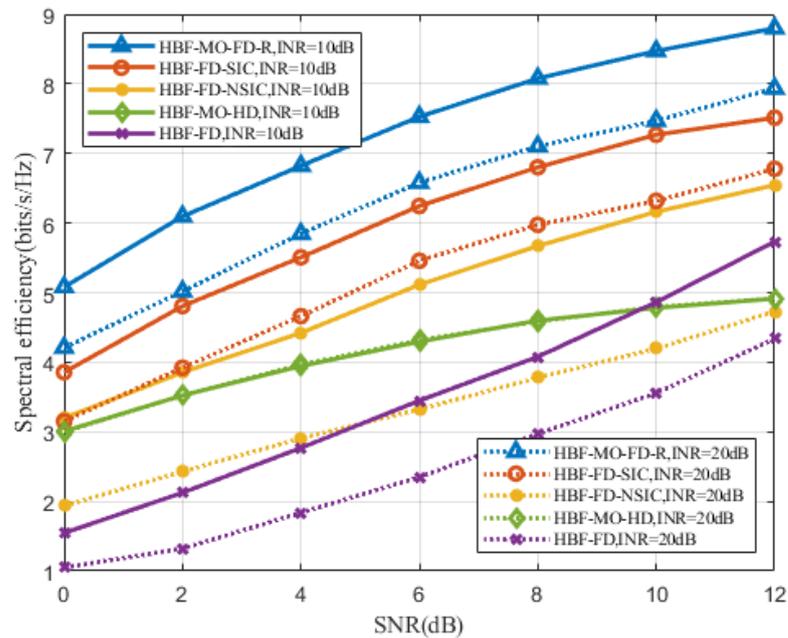


Figure 8. SE versus SNR for four HBF algorithms with different INR settings.

Figure 9 depicts the SE performance with various CSI errors. We analysis how the SE performance is affected by estimation error variance with $\alpha = 0.6$, $\beta = 0.4$ and INR = 10 dB. As we can see, imperfect CSI will lead to performance degradation in all approaches, whereas our algorithm still has the best performance, which illustrates the effectiveness of the robust design and SIC.

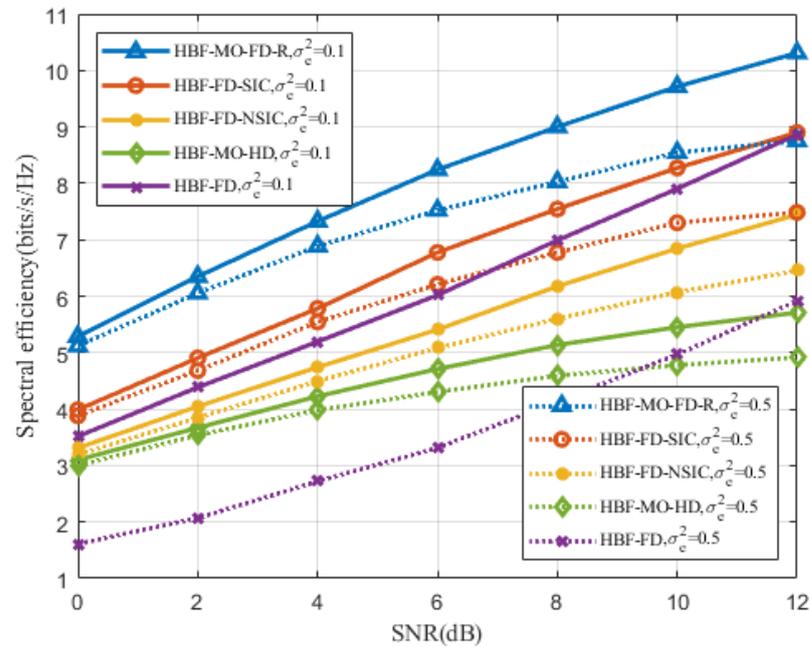


Figure 9. SE versus SNR with different σ_c^2 .

6. Conclusions

We have investigated the joint HBF design for mmWave AF relay communication systems in this paper. Specifically, the optimization-based HBF design for HD relay systems with perfect CSI has been addressed. We have designed an alternating optimization algorithm based on MO, where the modified MSE was adopted as the optimization objective. Simulation results have shown that the performance prevailed other HBF algorithms in the literature. Moreover, we have further extended the scenario to more general cases with channel estimation errors, where the relay system was operated in FD mode. A robust HBF design has been proposed, with SI mitigated by the NP based algorithm. It has been demonstrated via numerical results that the proposed scheme had an improved robustness against SI.

Author Contributions: Conceptualization, J.Z. and H.W.; methodology, J.Z. and D.J.; validation, H.W., B.L. and Y.Z.; formal analysis, B.L. and H.Y.; investigation, Y.Z. and X.L.; writing—original draft preparation, J.Z. and Y.Z.; writing—review and editing, H.W. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Derivation of (13)

In accordance with [28], the differential of $T1(\mathbf{F}_{RF})$ is given by

$$d[T1(\mathbf{F}_{RF})] = \text{tr} \left[\nabla T1(\mathbf{F}_{RF}) d(\mathbf{F}_{RF}^H) \right]. \quad (\text{A1})$$

By defining $\mathbf{B}_{12} = \mathbf{I}_{N_S} + (\mathbf{W}^H \mathbf{H}_2 \bar{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF}) \mathbf{B}_{11}^{-1} (\mathbf{W}^H \mathbf{H}_2 \bar{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF})^H$, (A1) can be rewritten as

$$d[T1(\mathbf{F}_{RF})] = d \left[\text{tr}(\mathbf{B}_{12}^{-1}) \right] = -\text{tr}(\mathbf{B}_{12}^{-2} d\mathbf{B}_{12}). \quad (\text{A2})$$

The differential of \mathbf{B}_{12} can be expressed as

$$d\mathbf{B}_{12} = \mathbf{W}^H \mathbf{H}_2 \bar{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \left[d\mathbf{B}_{11}^{-1} (\mathbf{W}^H \mathbf{H}_2 \bar{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF})^H + \mathbf{B}_{11}^{-1} d\mathbf{F}_{RF}^H (\mathbf{W}^H \mathbf{H}_2 \bar{\mathbf{G}} \mathbf{H}_1)^H \right]. \quad (\text{A3})$$

Due to the fact that $d\mathbf{X}^{-1} = -\mathbf{X}^{-1}(d\mathbf{X})\mathbf{X}^{-1}$, we have the differential of \mathbf{B}_{11}^{-1} as $d\mathbf{B}_{11}^{-1} = -\mathbf{B}_{11}^{-1}(d\mathbf{B}_{11})\mathbf{B}_{11}^{-1}$. In addition, $d\mathbf{B}_{11}$ can be expressed as

$$d\mathbf{B}_{11} = b_{11} d\mathbf{F}_{RF}^H \mathbf{F}_{RF}. \quad (\text{A4})$$

Substituting (A3) and (A4) into (A2) yields

$$d[T1(\mathbf{F}_{RF})] = \text{tr} \left[\left(b_{11} \mathbf{F}_{RF} \mathbf{B}_{11}^{-1} \mathbf{F}_{RF}^H - \mathbf{I}_{N_T} \right) \mathbf{H}_1^H \bar{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} \mathbf{B}_{12}^{-2} \mathbf{W}^H \mathbf{H}_2 \bar{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \mathbf{B}_{11}^{-1} d\mathbf{F}_{RF}^H \right]. \quad (\text{A5})$$

Finally, comparing (A5) with (A1), we can obtain the Euclidean gradient of $T1(\mathbf{F}_{RF})$ with respect to \mathbf{F}_{RF} as

$$\nabla T1(\mathbf{F}_{RF}) = \left(b_{11} \mathbf{F}_{RF} \mathbf{B}_{11}^{-1} \mathbf{F}_{RF}^H - \mathbf{I}_{N_T} \right) \mathbf{H}_1^H \bar{\mathbf{G}}^H \mathbf{H}_2^H \mathbf{W} \mathbf{B}_{12}^{-2} \mathbf{W}^H \mathbf{H}_2 \bar{\mathbf{G}} \mathbf{H}_1 \mathbf{F}_{RF} \mathbf{B}_{11}^{-1}. \quad (\text{A6})$$

References

1. Busari, S.A.; Huq, K.M.S.; Mumtaz, S.; Dai, L.; Rodriguez, J. Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey. *IEEE Commun. Surv. Tutorials* **2018**, *20*, 836–869. [\[CrossRef\]](#)
2. Wang, X.; Kong, L.; Kong, F.; Qiu, F.; Xia, M.; Arnon, S.; Chen, G. Millimeter Wave Communication: A Comprehensive Survey. *IEEE Commun. Surv. Tutorials* **2018**, *20*, 1616–1653. [\[CrossRef\]](#)
3. Rangan, S.; Rappaport, T.S.; Erkip, E. Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges. *Proc. IEEE* **2014**, *102*, 366–385. [\[CrossRef\]](#)
4. Lin, X.; Andrews, J.G. Connectivity of Millimeter Wave Networks With Multi-Hop Relaying. *IEEE Wirel. Commun. Lett.* **2015**, *4*, 209–212. [\[CrossRef\]](#)
5. Zhang, Z.; Long, K.; Vasilakos, A.V.; Hanzo, L. Full-Duplex Wireless Communications: Challenges, Solutions, and Future Research Directions. *Proc. IEEE* **2016**, *104*, 1369–1409. [\[CrossRef\]](#)
6. Kolodziej, K.E.; Perry, B.T.; Herd, J.S. In-Band Full-Duplex Technology: Techniques and Systems Survey. *IEEE Trans. Microw. Theory Tech.* **2019**, *67*, 3025–3041. [\[CrossRef\]](#)
7. Lee, J.; Lee, Y.H. AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing. In Proceedings of the 2014 IEEE International Conference on Communications (ICC), Sydney, NSW, Australia, 10–14 June 2014; pp. 5838–5842.
8. Mustafa, H.M.T.; Baik, J.I.; You, Y.H.; Abbasi, Z.; Song, H.K. Hybrid Wideband Millimeter Wave Transceiver for Single-User Multi-Relay MIMO Systems. *IEEE Access* **2023**, *11*, 93600–93618. [\[CrossRef\]](#)
9. Xu, W.; Wang, Y.; Xue, X. ADMM for Hybrid Precoding of Relay in Millimeter-Wave Massive MIMO System. In Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 27–30 August 2018; pp. 1–5.
10. Xue, X.; Wang, Y.; Dai, L.; Masouros, C. Relay Hybrid Precoding Design in Millimeter-Wave Massive MIMO Systems. *IEEE Trans. Signal Process.* **2018**, *66*, 2011–2026. [\[CrossRef\]](#)
11. Jiang, L.; Liu, X.L.; Jafarkhani, H. Hybrid Precoding/Combining Design in mmWave Amplify-and-Forward MIMO Relay Networks. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–6.
12. Han, M.; Du, J.; Zhang, Y.; Chen, Y.; Li, X.; Rabie, K.M.; Kara, F. Hybrid Beamforming With Sub-Connected Structure for MmWave Massive Multi-User MIMO Relay Systems. *IEEE Trans. Green Commun. Netw.* **2023**, *7*, 772–786. [\[CrossRef\]](#)

13. Luo, Z.; Zhao, L.; Liu, H.; Zhan, C. Robust Hybrid Beamforming Designs for Multi-user MmWave Relay Systems. In Proceedings of the 2022 IEEE Wireless Communications and Networking Conference (WCNC), Austin, TX, USA, 10–13 April 2022; pp. 1551–1556.
14. Jiang, L.; Jafarkhani, H. mmWave Amplify-and-Forward MIMO Relay Networks With Hybrid Precoding/Combining Design. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 1333–1346. [[CrossRef](#)]
15. Ding, Q.; Gao, X.; Deng, Y.; Wu, Z. Discrete Phase Shifters-Based Hybrid Precoding for Full-Duplex mmWave Relaying Systems. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 3698–3709. [[CrossRef](#)]
16. Cai, Y.; Xu, Y.; Shi, Q.; Champagne, B.; Hanzo, L. Robust Joint Hybrid Transceiver Design for Millimeter Wave Full-Duplex MIMO Relay Systems. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 1199–1215. [[CrossRef](#)]
17. Hei, Y.; Yu, S.; Liu, C.; Li, W.; Yang, J. Energy-Efficient Hybrid Precoding for mmWave MIMO Systems With Phase Modulation Array. *IEEE Trans. Green Commun. Netw.* **2020**, *4*, 678–688. [[CrossRef](#)]
18. Chen, C.H.; Tsai, C.R.; Liu, Y.H.; Hung, W.L.; Wu, A.Y. Compressive Sensing (CS) Assisted Low-Complexity Beamspace Hybrid Precoding for Millimeter-Wave MIMO Systems. *IEEE Trans. Signal Process.* **2017**, *65*, 1412–1424. [[CrossRef](#)]
19. Jin, J.; Zheng, Y.R.; Chen, W.; Xiao, C. Hybrid Precoding for Millimeter Wave MIMO Systems: A Matrix Factorization Approach. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 3327–3339. [[CrossRef](#)]
20. Alkhateeb, A.; El Ayach, O.; Leus, G.; Heath, R.W. Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 831–846. [[CrossRef](#)]
21. Lin, T.; Cong, J.; Zhu, Y.; Zhang, J.; Ben Letaief, K. Hybrid Beamforming for Millimeter Wave Systems Using the MMSE Criterion. *IEEE Trans. Commun.* **2019**, *67*, 3693–3708. [[CrossRef](#)]
22. Joham, M.; Utschick, W.; Nosske, J. Linear transmit processing in MIMO communications systems. *IEEE Trans. Signal Process.* **2005**, *53*, 2700–2712. [[CrossRef](#)]
23. Stankovic, V.; Haardt, M. Generalized Design of Multi-User MIMO Precoding Matrices. *IEEE Trans. Wirel. Commun.* **2008**, *7*, 953–961. [[CrossRef](#)]
24. Nguyen, D.H.N.; Le, L.B.; Le-Ngoc, T. Hybrid MMSE precoding for mmWave multiuser MIMO systems. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–6.
25. Sohrabi, F.; Yu, W. Hybrid Analog and Digital Beamforming for mmWave OFDM Large-Scale Antenna Arrays. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1432–1443. [[CrossRef](#)]
26. Nguyen, D.H.N.; Le, L.B.; Le-Ngoc, T.; Heath, R.W. Hybrid MMSE Precoding and Combining Designs for mmWave Multiuser Systems. *IEEE Access* **2017**, *5*, 19167–19181. [[CrossRef](#)]
27. Yu, X.; Shen, J.C.; Zhang, J.; Letaief, K.B. Alternating Minimization Algorithms for Hybrid Precoding in Millimeter Wave MIMO Systems. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 485–500. [[CrossRef](#)]
28. hjørungnes, A. *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*; Cambridge University Press: Cambridge, UK, 2011.
29. Absil, P.A.; Mahony, R.; Sepulchre, R. *Optimization Algorithms on Matrix Manifolds*; Princeton University Press: Princeton, NJ, USA, 2008.
30. Luo, Z.; Zhao, L.; Tonghui, L.; Liu, H.; Zhang, R. Robust Hybrid Precoding/Combining Designs for Full-Duplex Millimeter Wave Relay Systems. *IEEE Trans. Veh. Technol.* **2021**, *70*, 9577–9582. [[CrossRef](#)]
31. Zhang, Y.; Xiao, M.; Han, S.; Skoglund, M.; Meng, W. On Precoding and Energy Efficiency of Full-Duplex Millimeter-Wave Relays. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 1943–1956. [[CrossRef](#)]
32. Rong, Y. Robust Design for Linear Non-Regenerative MIMO Relays With Imperfect Channel State Information. *IEEE Trans. Signal Process.* **2011**, *59*, 2455–2460. [[CrossRef](#)]
33. Zhang, X.; Palomar, D.P.; Ottersten, B. Statistically Robust Design of Linear MIMO Transceivers. *IEEE Trans. Signal Process.* **2008**, *56*, 3678–3689. [[CrossRef](#)]
34. Satyanarayana, K.; El-Hajjar, M.; Kuo, P.H.; Mourad, A.; Hanzo, L. Hybrid Beamforming Design for Full-Duplex Millimeter Wave Communication. *IEEE Trans. Veh. Technol.* **2019**, *68*, 1394–1404. [[CrossRef](#)]
35. Luo, Z.; Zhang, X.; Gou, L.; Liu, H. Full-Duplex mmWave Communications With Robust Hybrid Beamforming. In Proceedings of the 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 7–11 December 2021; pp. 1–6.
36. Gupta, A.K.; Nagar, D.K. *Matrix Variate Distributions*; Chapman and Hall/CRC: Boca Raton, FL, USA, 1999.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.