

Article

Beamforming Based Full-Duplex for Millimeter-Wave Communication

Xiao Liu ^{1,2,3}, Zhenyu Xiao ^{1,2,3,*}, Lin Bai ^{1,2,3}, Jinho Choi ⁴, Pengfei Xia ⁵ and Xiang-Gen Xia ⁶

¹ School of Electronic and Information Engineering, Beihang University, Beijing 100191, China; eeliuxiao@buaa.edu.cn (X.L.); l.bai@buaa.edu.cn (L.B.)

² Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China

³ Beijing Key Laboratory for Network-Based Cooperative Air Traffic Management, and Beijing Laboratory for General Aviation Technology, Beijing 100191, China

⁴ School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Korea; jchoi0114@gist.ac.kr

⁵ School of Electronics and Information Engineering and the Key Laboratory of Embedded System and Service Computing, Tongji University, Shanghai 200092, China; pengfei.xia@gmail.com

⁶ Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA; xianggen@udel.edu

* Correspondence: xiaozzy@buaa.edu.cn; Tel.: +86-10-8231-7848 (ext. 6249)

Academic Editors: He (Henry) Chen, Yonghui Li, Kei Sakaguchi and Yong Li

Received: 26 April 2016; Accepted: 14 July 2016; Published: 21 July 2016

Abstract: In this paper, we study beamforming based full-duplex (FD) systems in millimeter-wave (mmWave) communications. A joint transmission and reception (Tx/Rx) beamforming problem is formulated to maximize the achievable rate by mitigating self-interference (SI). Since the optimal solution is difficult to find due to the non-convexity of the objective function, suboptimal schemes are proposed in this paper. A low-complexity algorithm, which iteratively maximizes signal power while suppressing SI, is proposed and its convergence is proven. Moreover, two closed-form solutions, which do not require iterations, are also derived under minimum-mean-square-error (MMSE), zero-forcing (ZF), and maximum-ratio transmission (MRT) criteria. Performance evaluations show that the proposed iterative scheme converges fast (within only two iterations on average) and approaches an upper-bound performance, while the two closed-form solutions also achieve appealing performances, although there are noticeable differences from the upper bound depending on channel conditions. Interestingly, these three schemes show different robustness against the geometry of Tx/Rx antenna arrays and channel estimation errors.

Keywords: full duplex; self-interference cancellation; beamforming; millimeter-wave; mmWave

1. Introduction

Full-duplex wireless communications (FDWC) for simultaneous transmission and reception (Tx/Rx) in the same frequency band [1–5] have attracted increasing attention recently due to the potential of doubling the spectrum efficiency. However, in order to achieve FDWC, the self-interference (SI) generated from a local transmitter to a local receiver must be mitigated for satisfactory performances [1,6], and this constitutes one of the critical challenges in FDWC.

There are basically three different approaches for SI cancellation. The first one is radio-frequency (RF) cancellation (or analog cancellation), where the RF signal to be transmitted at the Tx is exploited as a reference RF signal for SI cancellation in the Rx RF chain [1,3,4,7]. The second one is antenna cancellation, where multiple transmit (receive) antennas are carefully placed to generate two replicas with opposite phases [8,9] such that cancellation can be achieved by just adding these two replicas. The third one is digital cancellation, which is generally used together with RF or antenna cancellation

to further mitigate the residual SI at baseband [1,3–5,8,9]. Bharadia et al. [10] showed that full-duplex (FD) radio with a combined cancellation approach is able to cancel about 110 dB SI.

In this paper, we consider a beamforming-based approach to mitigate SI for FDWC. A distinct advantage of this approach is that with beamforming cancellation, some of the conventional RF, antenna and baseband cancellation operations may be avoided, and this greatly reduces the system complexity. Beamforming cancellation is particularly meaningful for millimeter wave (mmWave) wireless communications, where large antenna arrays are typically required to compensate for the high pass loss in the mmWave frequency band [11–19]. Note that although frequency re-use may be easier for mmWave communication since the pass loss is high for mmWave signal, more spectrum efficiency is almost always a plus. To improve the spectral efficiency for mmWave communication may be always favored for certain scenarios, e.g., mmWave backhaul applications, where high capacity is required to support high data rate. Also, in the case of mmWave cellular with dense users, e.g., stadiums and movie theaters, FD transmission can significantly increase the multi-user capacity.

On the other hand, the FD mmWave communication does not significantly increase the system complexity. To illustrate this, let us compare the complexity of an FD-mmWave node and that of a regular mmWave node with frequency-division duplex (FDD). There are a Tx RF chain and Tx antenna array, as well as a Rx RF chain and a Rx antenna array, at both the FD-mmWave node and the regular FDD mmWave node. The only difference is that at the FD-mmWave node the Rx needs to mitigate SI, while at the regular FDD mmWave node, the Rx does not need to mitigate SI. For FD-mmWave communication, SI cancellation can be done by using the beamforming technology proposed in this paper, which only needs to control the antenna weight vectors and almost does not increase the system complexity. In brief, if beamforming technology is adopted to mitigate the SI, the complexity of an FD-mmWave node is similar to that of a regular FDD mmWave node.

However, in order to realize FD in mmWave communications, we face a joint Tx/Rx beamforming (JTR-BF) problem to maximize the Tx/Rx achievable rate. As this problem is non-convex, suboptimal solutions are expected. The existing JTR-BF schemes proposed in mmWave communications are basically infeasible, because SI was not considered in these schemes [13,20–26]. Although the mitigation of loopback SI [27,28] as well as the utilization of loopback SI [29,30] are considered in multiple-input multiple-output (MIMO) relays, these methods cannot be employed in our setup, as their signal models are built for relay systems, but in this paper a bi-directional FD transmission system is considered. In [31], digital beamforming to cancel SI in FDWC was studied from an experimental perspective, where the JTR-BF with SI was not analytically investigated. Although some suboptimal distributed solutions proposed for the K -user interference alignment (IA) problem [32–35], e.g., the max-signal-to-interference-plus-noise-ratio (Max-SINR), maximum power (Max-Power), minimum leakage (Min-Leakage), and minimum-mean-square-error (MMSE) schemes [35], may be applicable, they are all iterative solutions designed for general K -user IA problems. Applying them to FD mmWave communications would result in high computational complexity for large antenna arrays. Moreover, the convergence of some of them is not yet proven in the literature [35] to the best of our knowledge.

In this paper, we propose several suboptimal solutions to the JTR-BF problem for FD mmWave communications. Firstly, an iterative algorithm, which iteratively maximizes the signal power with zero-forcing (ZF) SI (ZF-Max-Power), is proposed, and its convergence is proven. Next, two closed-form solutions are derived under MMSE, ZF, and maximum-ratio transmission (MRT) criteria, namely a lower bound based MMSE solution (LB-MMSE) and ZF SI with MRT (SI-ZF-MRT), where iterations are not required. Performance evaluations show that ZF-Max-Power approaches an upper bound on the joint achievable rate, and it needs only two iterations on average to achieve the convergence with random initial points. These performances of ZF-Max-Power are almost the same as those of the best baseline for the IA problem [32–35], namely Max-SINR. However, the convergence of Max-SINR is unproven yet [35] to the best of our knowledge, and the computational complexity of ZF-Max-Power is significantly lower than that of Max-SINR since matrix inversion is not needed.

The two closed-form solutions achieve suboptimal performances to the upper bound depending on channel conditions. In addition, ZF-Max-Power and SI-ZF-MRT are robust against the geometry of Tx/Rx antenna arrays due to the operation of ZF SI, while LB-MMSE is not. ZF-Max-Power and LB-MMSE are robust against channel estimation errors, while SI-ZF-MRT is not. These results verify the feasibility of FD mmWave communication and the effectiveness of beamforming cancellation.

The rest of this paper is organized as follows. In Section 2, we introduce the system model and formulate the problem. In Section 3, we study the optimization problem, propose the ZF-Max-Power approach, and conduct the convergence analysis and complexity comparison. In Section 4, we propose the two closed-form beamforming schemes, namely LB-MMSE and SI-ZF-MRT. In Section 5, we present performance evaluations. The conclusions are drawn lastly in Section 6.

Notation: a , \mathbf{a} , \mathbf{A} , and \mathcal{A} denote a scalar variable, a vector, a matrix, and a set, respectively. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote conjugate, transpose and conjugate transpose, respectively. In addition, $[x_1, x_2, \dots, x_M]$ denotes a row vector with its elements being x_i . Some other operations used in this paper are defined as follows.

- $\mathbb{E}\{\cdot\}$ Expectation operation.
- $|x|$ Absolute value of scalar variable x .
- $\|\mathbf{x}\|$ 2-norm of vector \mathbf{x} .
- $\langle \mathbf{x}, \mathbf{y} \rangle$ Inner product, equals to $\mathbf{y}^H \mathbf{x}$.
- \mathbf{x}^* Optimal value of variable \mathbf{x} .

2. System Model and Problem Formulation

2.1. System Model

An FD mmWave communication system consisting of two nodes, namely Node #1 and Node #2, is illustrated in Figure 1. Each node is equipped with a transmit antenna array and a receive antenna array, and supports only one data stream [21,22,36]. We denote by n_{t1} and n_{t2} the numbers of antenna elements of the transmit arrays at Node #1 and Node #2, respectively, while by n_{r1} and n_{r2} those of the receive arrays at Node #1 and Node #2, respectively. In our model, Node #1 transmits signals to Node #2 and receives signals from Node #2 simultaneously; thus both the nodes suffer from SI transmitted by the local transmitters.

It is noteworthy that although we depict separate antenna arrays for the Tx/Rx chains in Figure 1 (this structure is indeed common in FDWC [1,3–5,8]), the Tx/Rx chains may also share the same antenna array [1,10]. Fortunately, our signal model is suitable for both cases. Note that the SI with a shared array may be even higher than with separate arrays.

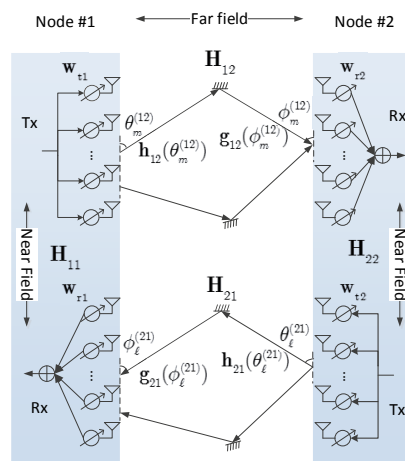


Figure 1. Illustration of the FD mmWave communication system.

2.2. Channel Model

As we can see from Figure 1, there are two types of channels. The first one is the communication channel, which represents the channel for information-bearing signals exchanged between Node #1 and Node #2, i.e., \mathbf{H}_{12} and \mathbf{H}_{21} , where \mathbf{H}_{ij} represents the channel from Node # i to Node # j . The other one is the SI channel. Clearly, \mathbf{H}_{11} and \mathbf{H}_{22} are the SI channels.

2.2.1. Communication Channel

We first consider the model of communication channel. Since the distance between these two nodes is generally much greater than the wavelength of mmWave, the commonly used far-field channel model, which has a plane wavefront, is suitable for \mathbf{H}_{12} and \mathbf{H}_{21} . According to channel measurement results for mmWave communication [13,37], mostly reflection contributes to generating multipath components (MPCs) besides the LOS component; scattering and diffraction effects are little due to the extremely short wavelength of mmWave communication. Thus, the MPCs in mmWave communication have a feature of directivity [22–24,38,39], i.e., different MPCs have different physical angles of departure (AoDs), i.e., $\theta_m^{(12)}$ and $\theta_\ell^{(21)}$, as well as angles of arrival (AoAs), e.g., $\phi_m^{(12)}$ and $\phi_\ell^{(21)}$, as shown in Figure 1. In general, mmWave signals have a wide band, and thus a frequency selective channel may be suitable [40]. However, with beamforming only a very small number (or even only one) of strong MPCs may be searched out to form beams between Tx and Rx. As a result, the effect of delay spread may be substantially mitigated [41]. Due to this reason, a frequency flat channel model is extensively used in mmWave communication [22–24,38,39], and the (narrow band) communication channels can be expressed as

$$\mathbf{H}_{12} = \sqrt{n_{t1}n_{r2}} \sum_{m=1}^M \alpha_m \mathbf{g}_{12}(\phi_m^{(12)}) \mathbf{h}_{12}^H(\theta_m^{(12)}) \quad (1)$$

and

$$\mathbf{H}_{21} = \sqrt{n_{t2}n_{r1}} \sum_{\ell=1}^L \beta_\ell \mathbf{g}_{21}(\phi_\ell^{(21)}) \mathbf{h}_{21}^H(\theta_\ell^{(21)}) \quad (2)$$

where M and L are the numbers of MPCs, α_m and β_ℓ are the coefficients of MPCs, $\mathbf{g}_{12}(\phi_m^{(12)})$ and $\mathbf{g}_{21}(\phi_\ell^{(21)})$ are receive steering vectors, $\mathbf{h}_{12}(\theta_m^{(12)})$ and $\mathbf{h}_{21}(\theta_\ell^{(21)})$ are transmit steering vectors of \mathbf{H}_{12} and \mathbf{H}_{21} , respectively. For uniform linear arrays (ULAs) with half-wavelength spacing, these steering vectors are defined as Equation (4), and they are all functions of the corresponding steering angles. Although ULA is adopted in this paper, the developed schemes are also feasible for other types of arrays, like uniform planar array (UPA) or circular array, because different types of arrays affect only the channel matrices. For convenience, we have the following normalization:

$$\sum_{m=1}^M \mathbb{E}\{|\alpha_m|^2\} = \sum_{\ell=1}^L \mathbb{E}\{|\beta_\ell|^2\} = 1. \quad (3)$$

In the case of Tx/Rx sharing the same antenna array at a node, we have $n_{r1} = n_{t1}$, $n_{r2} = n_{t2}$, $L = M$, $\theta_m^{12} = \phi_m^{21}$, $\theta_m^{21} = \phi_m^{12}$, and $\alpha_m = \beta_m$, $m = 1, 2, \dots, M$, i.e., parameters of \mathbf{H}_{12} are the same as those of \mathbf{H}_{21} . However, since Tx/Rx have different RF chains [1,10], the beamforming and combining vectors are basically different.

Note that for mmWave communications there may be other models. For instance, in [42] a clustered channel model was adopted, where the channel includes several clusters, and a cluster consists of many MPCs with small angle differences. Different models are suitable for different communication circumstances. As our schemes do not exploit the specific feature of the communication channel, they can be used for different models.

$$\begin{aligned}
\mathbf{g}_{12}(\phi_m^{(12)}) &= \left[\exp(j\pi 0 \cos(\phi_m^{(12)})), \exp(j\pi 1 \cos(\phi_m^{(12)})), \dots, \exp(j\pi(n_{r2} - 1) \cos(\phi_m^{(12)})) \right]^T / \sqrt{n_{r2}} \\
\mathbf{g}_{21}(\phi_\ell^{(21)}) &= \left[\exp(j\pi 0 \cos(\phi_\ell^{(21)})), \exp(j\pi 1 \cos(\phi_\ell^{(21)})), \dots, \exp(j\pi(n_{r1} - 1) \cos(\phi_\ell^{(21)})) \right]^T / \sqrt{n_{r1}} \\
\mathbf{h}_{12}(\theta_m^{(12)}) &= \left[\exp(j\pi 0 \cos(\theta_m^{(12)})), \exp(j\pi 1 \cos(\theta_m^{(12)})), \dots, \exp(j\pi(n_{t1} - 1) \cos(\theta_m^{(12)})) \right]^T / \sqrt{n_{t1}} \\
\mathbf{h}_{21}(\theta_\ell^{(21)}) &= \left[\exp(j\pi 0 \cos(\theta_\ell^{(21)})), \exp(j\pi 1 \cos(\theta_\ell^{(21)})), \dots, \exp(j\pi(n_{t2} - 1) \cos(\theta_\ell^{(21)})) \right]^T / \sqrt{n_{t2}}
\end{aligned} \tag{4}$$

2.2.2. SI Channel

Next, we consider the strength of SI and the SI channel. Note that even in mmWave band, where the center frequency is high and the signal strength attenuates rapidly, SI may be still much more significant than the background noise, because Tx/Rx antennas locate close to each other in FD mmWave communication. For instance, if the wavelength of the carrier frequency is 1 mm, the signal bandwidth is 1 GHz, and the transmission power is 20 dBm, according to the Friis formula, the SI at a position 10 cm (100λ) away from a transmit antenna is $20 - 20 \log_{10}(4\pi \times 100) = -42$ dBm, which is much greater than a typical noise power $\sigma^2 = 10 \log_{10}(\kappa TB) = 10 \log_{10}(1.38 \times 10^{-23} \times 300 \times 10^9 \times 10^3) = -83.83$ dBm, where κ , T , B are the Boltzmann constant, ambient temperature and bandwidth, respectively. Hence, SI should be taken into account in FD mmWave communication.

As the distance between the transmit and receive arrays at each node is short in portable devices, the far-field range condition, i.e., $R_0 \geq 2D^2/\lambda$ [43], may not hold for SI channel, where D is the diameter of the antenna aperture, λ is the wavelength of the carrier. For instance, considering a half-wavelength spaced ULA with 64 elements, the far-field range should satisfy $R_0 \geq 2(32\lambda)^2/\lambda = 2048\lambda$, which is basically too large for small-size devices like mobile phones or laptops even at the mmWave band. Thus, the SI channels may employ the near-field model, which has a spherical wavefront [43–47]. In such a case, the SI channels highly depend on the placement of the transmit and receive arrays as well as the circumstances.

In this paper we consider an antenna placement as shown in Figure 2. The distance between the first elements of the two arrays is d , and the angle between these two ULAs is ω . Although this placement together with the two parameters, i.e., d and ω , cannot cover all the possible antenna placements, it allows us to perform tractable analysis. Moreover, it can reflect the effects of two typical factors of antenna placement, i.e., the distance and angle between these two arrays, on the beamforming performance. With this antenna placement, the coefficient corresponding to the j th row and i th column of \mathbf{H}_{11} or \mathbf{H}_{22} is [43–47].

$$[\mathbf{H}_{11} \text{ or } \mathbf{H}_{22}]_{ij} = h_{ij} = \frac{\rho}{r_{ij}} \exp\left(-j2\pi \frac{r_{ij}}{\lambda}\right) \tag{5}$$

where r_{ij} is the distance between the i -th element of the transmit array and the j -th element of the receive array, and ρ is a constant for power normalization such that $\text{tr}(\mathbf{H}_{11}\mathbf{H}_{11}^H) = n_{t1}n_{r1}$ and $\text{tr}(\mathbf{H}_{22}\mathbf{H}_{22}^H) = n_{t2}n_{r2}$. The expression for r_{ij} is shown in Equation (6).

$$r_{ij} = \sqrt{\left(\frac{d}{\tan(\omega)} + (j-1)\frac{\lambda}{2}\right)^2 + \left(\frac{d}{\sin(\omega)} + (i-1)\frac{\lambda}{2}\right)^2 - 2\left(\frac{d}{\tan(\omega)} + (j-1)\frac{\lambda}{2}\right)\left(\frac{d}{\sin(\omega)} + (i-1)\frac{\lambda}{2}\right)\cos(\omega)} \tag{6}$$

In the case of Tx/Rx sharing the same antenna array, we have $d = 0$ and $\omega = 0$.

The SI channel model adopted in this paper is a simplified one, which is typical [46,47] but not necessarily accurate. In practice, the SI channel can be very complicated, including signal reflection, scattering and the coupling effects between adjacent antennas. Hence, it is very difficult, if not impossible, to elaborate an accurate SI channel. On the other hand, the proposed schemes can be used

for arbitrary SI channels, including the completely accurate one. Hence, we adopt the simple and typical model shown in Equation (6) here, which can reflect the robustness of the proposed schemes against the geometry of the Tx/Rx antenna arrays within a node.

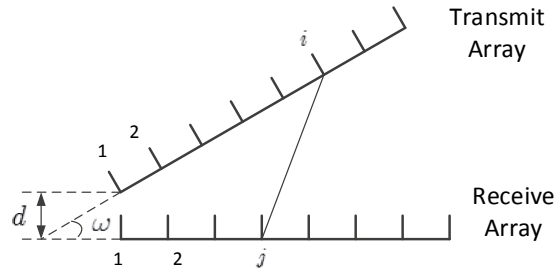


Figure 2. The transmit and receive antenna arrays of a node.

2.3. Problem Formulation

With the above system and channel models, the received signals at Node #1 and Node #2 are written as

$$y_1 = \mathbf{w}_{r1}^H (\sqrt{\varepsilon_{21}} \mathbf{H}_{21} \mathbf{w}_{t2} s_2 + \sqrt{\varepsilon_{11}} \mathbf{H}_{11} \mathbf{w}_{t1} s_1 + \mathbf{n}_1) \quad (7)$$

and

$$y_2 = \mathbf{w}_{r2}^H (\sqrt{\varepsilon_{12}} \mathbf{H}_{12} \mathbf{w}_{t1} s_1 + \sqrt{\varepsilon_{22}} \mathbf{H}_{22} \mathbf{w}_{t2} s_2 + \mathbf{n}_2) \quad (8)$$

respectively, where s_1 and s_2 are the transmitted symbols with unit power at Node #1 and Node #2, respectively, ε_{21} and ε_{12} are the average powers of the desired received signals, ε_{11} and ε_{22} are the average powers of the SI, \mathbf{n}_1 and \mathbf{n}_2 are the Gaussian white noise vectors with $\mathbb{E}\{\mathbf{n}_1 \mathbf{n}_1^H\} = \mathbf{I}_{n_{r1}}$ and $\mathbb{E}\{\mathbf{n}_2 \mathbf{n}_2^H\} = \mathbf{I}_{n_{r2}}$, respectively, \mathbf{w}_{t1} and \mathbf{w}_{t2} are the transmit antenna weight vectors (AWVs), and \mathbf{w}_{r1} and \mathbf{w}_{r2} are the receive AWVs. The two-norms of all these AWVs are normalized to 1.

With the transmit and receive AWVs, the joint achievable rate (JAR) can be expressed as

$$R = \log_2 \left(1 + \frac{\varepsilon_{21} |\mathbf{w}_{r1}^H \mathbf{H}_{21} \mathbf{w}_{t2}|^2}{1 + \varepsilon_{11} |\mathbf{w}_{r1}^H \mathbf{H}_{11} \mathbf{w}_{t1}|^2} \right) + \log_2 \left(1 + \frac{\varepsilon_{12} |\mathbf{w}_{r2}^H \mathbf{H}_{12} \mathbf{w}_{t1}|^2}{1 + \varepsilon_{22} |\mathbf{w}_{r2}^H \mathbf{H}_{22} \mathbf{w}_{t2}|^2} \right) \quad (9)$$

where $\varepsilon_{11} |\mathbf{w}_{r1}^H \mathbf{H}_{11} \mathbf{w}_{t1}|^2$ and $\varepsilon_{22} |\mathbf{w}_{r2}^H \mathbf{H}_{22} \mathbf{w}_{t2}|^2$ are the average powers of SI at Nodes #1 and #2, respectively. The JTR-BF problem is formulated as

$$\begin{aligned} & \underset{\mathbf{w}_{t1}, \mathbf{w}_{r1}, \mathbf{w}_{t2}, \mathbf{w}_{r2}}{\text{maximize}} && R \\ & \text{subject to} && \|\mathbf{w}_{t1}\|^2 = \|\mathbf{w}_{r1}\|^2 = 1 \\ & && \|\mathbf{w}_{t2}\|^2 = \|\mathbf{w}_{r2}\|^2 = 1 \end{aligned} \quad (10)$$

As the main purpose of this paper is to investigate the feasibility of FD mmWave communications and evaluate the performance of beamforming cancellation, the channel matrices and powers in Equation (10) are assumed known *a-priori*. In practice, these parameters can be estimated provided that the channel does not change too fast. For instance, as an mmWave communication channel has the feature of directivity and is sparse in the angle domain, schemes like AoD/AoA estimation [48], beam searching [26,39,49], iterative training [22,50], and even compressed sensing [51,52] can be adopted for the communication channel estimation. The estimation of SI channel can be more straightforward, e.g., one can estimate a scalar channel coefficient between a single Tx/Rx antenna pair (i.e., the i -th transmit antenna and the j -th receive antenna) once a time. After $n_{ri} n_{tj}$ ($i = 1, 2$) measurements, the SI channel matrix can be estimated. Although $n_{ri} n_{tj}$ is large in mmWave communication, each measurement may

require only one symbol thanks to the high strength of SI, rather than a long training sequence like those in [22,26,39,49–51]. Thus, the time cost of the SI channel estimation is yet affordable.

In addition, in our model both amplitudes and phases of the AWVs are controllable. In fact, a constant-amplitude (CA) array with only phases controllable has lower complexity [22,26,39,49,51]. An N -element CA array needs N phase shifters, which are not difficult to implement. In contrast, an N -element array with both amplitudes and phases controllable requires N phase shifters and N variable gain amplifier. Hence, a CA array has a lower complexity than an amplitude-and-phase-controllable array, especially when N is large, and thus is favored in mmWave communications, where N is large in general. However, the array structure without the CA constraint cannot be definitely ruled out, because it also arouses particular attention in both algorithm design [23,24,50,53,54] and implementation [55]. Moreover, the performance achieved with non-CA-constraint arrays can be seen as a bound achieved with the CA arrays, if the same scheme is adopted. Therefore, in this paper we adopt the arrays without the CA constraint to investigate the feasibility of FD mmWave communication and evaluate the beamforming performance, letting the problem with CA constraint be a further work.

3. The ZF-Max-Power Approach

It is clear that the JAR in Equation (10) is not a concave function, and the equality constraints are not affine. Thus, Equation (10) is not a convex/concave problem, and its globally optimal solution is hard to find. Consequently, we first give an upper bound JAR, and then propose the ZF-Max-Power approach in this section.

3.1. Upper Bound of the JAR

According to Equation (9), since $\varepsilon_{11}|\mathbf{w}_{r1}^H \mathbf{H}_{11} \mathbf{w}_{t1}|^2 \geq 0$ and $\varepsilon_{22}|\mathbf{w}_{r2}^H \mathbf{H}_{22} \mathbf{w}_{t2}|^2 \geq 0$, an *upper bound function* on the JAR, denoted by R_{ub} , in the presence of SI can be easily obtained as

$$R \leq \log_2 \left(1 + \varepsilon_{21} |\mathbf{w}_{r1}^H \mathbf{H}_{21} \mathbf{w}_{t2}|^2 \right) + \log_2 \left(1 + \varepsilon_{12} |\mathbf{w}_{r2}^H \mathbf{H}_{12} \mathbf{w}_{t1}|^2 \right) \triangleq R_{ub} \quad (11)$$

The corresponding optimal AWVs for R_{ub} are

$$\begin{aligned} \mathbf{w}_{t1} &= \text{RpSingVect}(\mathbf{H}_{12}), \quad \mathbf{w}_{t2} = \text{RpSingVect}(\mathbf{H}_{21}) \\ \mathbf{w}_{r1} &= \text{LpSingVect}(\mathbf{H}_{21}), \quad \mathbf{w}_{r2} = \text{LpSingVect}(\mathbf{H}_{11}) \end{aligned} \quad (12)$$

where $\text{LpSingVect}(\mathbf{X})$ and $\text{RpSingVect}(\mathbf{X})$ represent the left and right principal singular vectors of \mathbf{X} , respectively. Thus, we have

$$R \leq R_{ub} \leq R_{ub}^* < \infty \quad (13)$$

where R_{ub}^* is the maximum of the upper bound R_{ub} , and thus it is an upper bound on R .

3.2. The ZF-Max-Power Approach

As there are multiple coupled variables in Equation (10), we consider using the alternating optimization (AO) approach [56] to obtain a suboptimal solution of Equation (10). The basic idea of AO is to alternately optimize a few parameters, and assume the other parameters fixed and known [56]. In each round, a sub-problem with a few parameters are formulated and solved. Generally, this approach requires that the optimal solution to each sub-problem can be found. However, for the problem in Equation (10), the AO approach cannot be directly used, because, as we can see, even given \mathbf{w}_{r1} and \mathbf{w}_{r2} , optimal \mathbf{w}_{t1} and \mathbf{w}_{t2} still cannot be easily found.

To make the AO approach feasible, we propose the ZF-Max-Power scheme in this paper. The motivation of this scheme is as follows. Since the SI is usually significant in FD mmWave communication, we can force the SI to zero and maximize the signal power. In particular, we add

a constraint that the SI is completely mitigated. In such a case, we have $R = R_{ub}$, and the problem becomes

$$\begin{aligned} & \underset{\mathbf{w}_{t1}, \mathbf{w}_{r1}, \mathbf{w}_{t2}, \mathbf{w}_{r2}}{\text{maximize}} && R_{ub} \\ & \text{subject to} && \|\mathbf{w}_{t1}\|^2 = \|\mathbf{w}_{r1}\|^2 = 1 \\ & && \|\mathbf{w}_{t2}\|^2 = \|\mathbf{w}_{r2}\|^2 = 1 \\ & && \mathbf{w}_{r1}^H \mathbf{H}_{11} \mathbf{w}_{t1} = \mathbf{w}_{r2}^H \mathbf{H}_{22} \mathbf{w}_{t2} = 0 \end{aligned} \quad (14)$$

Since we have added a new constraint to the original problem in Equation (10), the solution of the new problem in Equation (14) is suboptimal to the original problem. To solve the new problem in Equation (14), we need the following result.

Lemma 1. Given an arbitrary set of linearly independent vectors $\{\mathbf{a}_i \in \mathbb{C}^{L \times 1}\}_{i=1,2,\dots,N}$; $N < L$ and an arbitrary vector $\mathbf{a} \in \mathbb{C}^{L \times 1}$, the vector $\mathbf{b} \in \mathbb{C}^{L \times 1}$ that maximizes $|\langle \mathbf{b}, \mathbf{a} \rangle|^2$ with the constraint that $\langle \mathbf{b}, \mathbf{a}_i \rangle = 0|_{i=1,2,\dots,N}$ and $\|\mathbf{b}\| = 1$ is (before normalization) $\mathbf{b} = \mathbf{a} - \sum_{i=1}^N \langle \mathbf{a}, \mathbf{b}_i \rangle \mathbf{b}_i$, where $\mathbf{b}_1 = \mathbf{a}_1 / \|\mathbf{a}_1\|$ and

$$\mathbf{b}_i = \frac{\mathbf{a}_i - \sum_{j=1}^{i-1} \langle \mathbf{a}_i, \mathbf{b}_j \rangle \mathbf{b}_j}{\|\mathbf{a}_i - \sum_{j=1}^{i-1} \langle \mathbf{a}_i, \mathbf{b}_j \rangle \mathbf{b}_j\|}, \quad i > 1 \quad (15)$$

Proof. See Appendix A. \square

Specifically, when $N = 1$ the optimal \mathbf{b} in Lemma 1 before normalization is

$$\mathbf{b}^* = \mathbf{a} - \langle \mathbf{a}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \quad (16)$$

and when $N = 2$ the optimal \mathbf{b} in Lemma 1 before normalization is

$$\mathbf{b}^* = \mathbf{a} - \langle \mathbf{a}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} - \langle \mathbf{a}, \mathbf{b}_2 \rangle \mathbf{b}_2 \quad (17)$$

where

$$\mathbf{b}_2 = \frac{\mathbf{a}_2 - \langle \mathbf{a}_2, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}}{\|\mathbf{a}_2 - \langle \mathbf{a}_2, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}} \quad (18)$$

Let us go back to the ZF-Max-Power scheme. According to Equation (16), given fixed \mathbf{w}_{t1} and \mathbf{w}_{t2} , the optimal \mathbf{w}_{r1} and \mathbf{w}_{r2} for Equation (14) (before normalization) are

$$\mathbf{w}_{r1} = \mathbf{H}_{21} \mathbf{w}_{t2} - \langle \mathbf{H}_{21} \mathbf{w}_{t2}, \frac{\mathbf{H}_{11} \mathbf{w}_{t1}}{\|\mathbf{H}_{11} \mathbf{w}_{t1}\|} \rangle \frac{\mathbf{H}_{11} \mathbf{w}_{t1}}{\|\mathbf{H}_{11} \mathbf{w}_{t1}\|} \quad (19)$$

and

$$\mathbf{w}_{r2} = \mathbf{H}_{12} \mathbf{w}_{t1} - \langle \mathbf{H}_{12} \mathbf{w}_{t1}, \frac{\mathbf{H}_{22} \mathbf{w}_{t2}}{\|\mathbf{H}_{22} \mathbf{w}_{t2}\|} \rangle \frac{\mathbf{H}_{22} \mathbf{w}_{t2}}{\|\mathbf{H}_{22} \mathbf{w}_{t2}\|} \quad (20)$$

respectively.

Similarly, given fixed \mathbf{w}_{r1} and \mathbf{w}_{r2} , the optimal \mathbf{w}_{t1} and \mathbf{w}_{t2} for Equation (14) (before normalization) are

$$\mathbf{w}_{t1} = \mathbf{H}_{12}^H \mathbf{w}_{r2} - \langle \mathbf{H}_{12}^H \mathbf{w}_{r2}, \frac{\mathbf{H}_{11}^H \mathbf{w}_{r1}}{\|\mathbf{H}_{11}^H \mathbf{w}_{r1}\|} \rangle \frac{\mathbf{H}_{11}^H \mathbf{w}_{r1}}{\|\mathbf{H}_{11}^H \mathbf{w}_{r1}\|} \quad (21)$$

and

$$\mathbf{w}_{t2} = \mathbf{H}_{21}^H \mathbf{w}_{r1} - \langle \mathbf{H}_{21}^H \mathbf{w}_{r1}, \frac{\mathbf{H}_{22}^H \mathbf{w}_{r2}}{\|\mathbf{H}_{22}^H \mathbf{w}_{r2}\|} \rangle \frac{\mathbf{H}_{22}^H \mathbf{w}_{r2}}{\|\mathbf{H}_{22}^H \mathbf{w}_{r2}\|} \quad (22)$$

respectively. Finally, the ZF-Max-Power scheme can be summarized as in Algorithm 1.

Algorithm 1 The ZF-Max-Power Scheme.**(1) Initialize:**

Initialize the transmit AWVs as $\mathbf{w}_{t1} = \text{RpSingVect}(\mathbf{H}_{12})$, $\mathbf{w}_{t2} = \text{RpSingVect}(\mathbf{H}_{21})$.

(2) Iteration:

Iterate the following process ρ times, then stop.

Compute \mathbf{w}_{r1} and \mathbf{w}_{r2} according to Equations (19) and (20), respectively; then normalize them.

Compute \mathbf{w}_{t1} and \mathbf{w}_{t2} according to Equations (21) and (22), respectively; then normalize them.

(3) Result:

\mathbf{w}_{t1} and \mathbf{w}_{t2} are the transmit AWVs, and \mathbf{w}_{r1} and \mathbf{w}_{r2} are the receive AWVs.

It is noted that there are various stopping rules for the ZF-Max-Power scheme. A simple one is to stop after a certain number of iterations, e.g., ρ iterations used in Algorithm 1. Another one is to compute R_{ub} after the n -th iteration and get $R_{\text{ub}}^{(n)}$ according to Equation (9). When $R_{\text{ub}}^{(n)} / R_{\text{ub}}^{(n-1)} < \mu$, stop the iteration, where μ is a predefined threshold slightly greater than 1, e.g., $\mu = 1.05$.

3.3. Convergence Analysis and Complexity Comparison

To prove the convergence of the ZF-Max-Power scheme, we need the following theorem.

Theorem 1. Let $R^{(n)}$ denote the value of R after the n -th iteration. Then $\{R^{(n)} | n = 1, 2, \dots\}$ is a non-descending sequence, i.e., $R^{(n+1)} \geq R^{(n)}$.

Proof. See Appendix B. \square

According to Equation (13), $R \leq R_{\text{ub}}^* < \infty$. Thus, $\{R^{(n)} | n = 1, 2, \dots\}$ converges to a suboptimal value, which guarantees the convergence of ZF-Max-Power.

On the other hand, the computational complexity of ZF-Max-Power is much lower than that of Max-SINR [35]. For simplicity, suppose $n_{t1} = n_{t2} = n_{r1} = n_{r2} = N$. According to Algorithm 1, in each iteration the main complexity of ZF-Max-Power lies in multiplications between a channel matrix and an AWV, as shown in Equations (19)–(22). Hence, the computational complexity of ZF-Max-Power is roughly $\mathcal{O}(N^2)$, because there are about N^2 scalar multiplications for a multiplication of a channel matrix and an AWV. In contrast, according to [35] matrix inversion is required in each iteration of Max-SINR. Hence, Max-SINR has a computational complexity of $\mathcal{O}(N^3)$, which is significantly higher than that of ZF-Max-Power, especially in FD mmWave communication where N is large.

4. Closed-Form Solutions

In this section, we consider different criteria for the JTR-BF problem that provide us with closed-form solutions.

4.1. LB-MMSE

It is natural to perform receive and transmit beamforming separately when considering closed-form solutions. According to Equation (9), an optimal solution can be achieved for receive beamforming by using the MMSE approach which is equivalent to maximizing the SINR. However, for transmit beamforming, the optimal solution is hard to find. Thus, we propose to optimize the lower bound on the JAR with MMSE, and the method is referred to as the lower bound MMSE (LB-MMSE) approach.

According to Equation (9), we have

$$\begin{aligned} R &\geq \log_2 \left(\frac{\varepsilon_{21} |\mathbf{w}_{r1}^H \mathbf{H}_{21} \mathbf{w}_{t2}|^2}{\mathbf{w}_{r1}^H \mathbf{w}_{r1} + \varepsilon_{11} |\mathbf{w}_{r1}^H \mathbf{H}_{11} \mathbf{w}_{t1}|^2} \frac{\varepsilon_{12} |\mathbf{w}_{r2}^H \mathbf{H}_{12} \mathbf{w}_{t1}|^2}{\mathbf{w}_{r2}^H \mathbf{w}_{r2} + \varepsilon_{22} |\mathbf{w}_{r2}^H \mathbf{H}_{22} \mathbf{w}_{t2}|^2} \right) \\ &= \log_2 \left(\frac{\varepsilon_{21} \mathbf{w}_{r1}^H \mathbf{H}_{21} \mathbf{w}_{t2} \mathbf{w}_{t2}^H \mathbf{H}_{21}^H \mathbf{w}_{r1}}{\mathbf{w}_{r1}^H (\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11} \mathbf{w}_{t1} \mathbf{w}_{t1}^H \mathbf{H}_{11}) \mathbf{w}_{r1}} \frac{\varepsilon_{12} \mathbf{w}_{r2}^H \mathbf{H}_{12} \mathbf{w}_{t1} \mathbf{w}_{t1}^H \mathbf{H}_{12}^H \mathbf{w}_{r2}}{\mathbf{w}_{r2}^H (\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22} \mathbf{w}_{t2} \mathbf{w}_{t2}^H \mathbf{H}_{22}) \mathbf{w}_{r2}} \right) \triangleq R_1 \end{aligned} \quad (23)$$

First, we find \mathbf{w}_{r1} and \mathbf{w}_{r2} to maximize the lower bound R_1 by temporally treating \mathbf{w}_{t1} and \mathbf{w}_{t2} as fixed parameters. This subproblem is actually to maximize SINR (or minimize MSE) at the two nodes. According to Lemma 2 in Appendix C, the optimal \mathbf{w}_{r1} and \mathbf{w}_{r2} (before normalization) can be respectively found as

$$\begin{aligned} \mathbf{w}_{r1} &= \left(\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11} \mathbf{w}_{t1} \mathbf{w}_{t1}^H \mathbf{H}_{11}^H \right)^{-1} \mathbf{H}_{21} \mathbf{w}_{t2} \\ \mathbf{w}_{r2} &= \left(\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22} \mathbf{w}_{t2} \mathbf{w}_{t2}^H \mathbf{H}_{22}^H \right)^{-1} \mathbf{H}_{12} \mathbf{w}_{t1} \end{aligned} \quad (24)$$

With these two receive AWVs, we further have Equation (25), where inequalities (a) and (b) are based on Lemmas 3 and 4, respectively, in Appendix C.

$$\begin{aligned} R_1 &= \log_2 \left(\varepsilon_{12} \varepsilon_{21} \left(\mathbf{w}_{t2}^H \mathbf{H}_{21}^H (\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11} \mathbf{w}_{t1} \mathbf{w}_{t1}^H \mathbf{H}_{11}^H)^{-1} \mathbf{H}_{21} \mathbf{w}_{t2} \right) \left(\mathbf{w}_{t1}^H \mathbf{H}_{12}^H (\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22} \mathbf{w}_{t2} \mathbf{w}_{t2}^H \mathbf{H}_{22}^H)^{-1} \mathbf{H}_{12} \mathbf{w}_{t1} \right) \right) \\ &\stackrel{(a)}{\geq} \log_2 \left(\varepsilon_{12} \varepsilon_{21} \frac{(\mathbf{w}_{t2}^H \mathbf{H}_{21}^H \mathbf{H}_{21} \mathbf{w}_{t2})^2}{\mathbf{w}_{t2}^H \mathbf{H}_{21}^H (\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11} \mathbf{w}_{t1} \mathbf{w}_{t1}^H \mathbf{H}_{11}^H) \mathbf{H}_{21} \mathbf{w}_{t2}} \frac{(\mathbf{w}_{t1}^H \mathbf{H}_{12}^H \mathbf{H}_{12} \mathbf{w}_{t1})^2}{\mathbf{w}_{t1}^H \mathbf{H}_{12}^H (\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22} \mathbf{w}_{t2} \mathbf{w}_{t2}^H \mathbf{H}_{22}^H) \mathbf{H}_{12} \mathbf{w}_{t1}} \right) \\ &\stackrel{(b)}{\geq} \log_2 \left(\varepsilon_{12} \varepsilon_{21} \frac{\mathbf{w}_{t2}^H \mathbf{H}_{21}^H \mathbf{H}_{21} \mathbf{w}_{t2}}{\mathbf{w}_{t1}^H (\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11}^H \mathbf{H}_{11}) \mathbf{w}_{t1}} \frac{\mathbf{w}_{t1}^H \mathbf{H}_{12}^H \mathbf{H}_{12} \mathbf{w}_{t1}}{\mathbf{w}_{t2}^H (\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22}^H \mathbf{H}_{22}) \mathbf{w}_{t2}} \right) \\ &= \log_2 \left(\varepsilon_{12} \varepsilon_{21} \frac{\mathbf{w}_{t2}^H \mathbf{H}_{21}^H \mathbf{H}_{21} \mathbf{w}_{t2}}{\mathbf{w}_{t2}^H (\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22}^H \mathbf{H}_{22}) \mathbf{w}_{t2}} \frac{\mathbf{w}_{t1}^H \mathbf{H}_{12}^H \mathbf{H}_{12} \mathbf{w}_{t1}}{\mathbf{w}_{t1}^H (\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11}^H \mathbf{H}_{11}) \mathbf{w}_{t1}} \right) \triangleq R_2 \end{aligned} \quad (25)$$

Next, let us find \mathbf{w}_{t1} and \mathbf{w}_{t2} to maximize the lower bound R_2 . This subproblem is equivalent to the following two optimization problems:

$$\arg \max_{\mathbf{w}_{t1}} \frac{\mathbf{w}_{t1}^H \mathbf{H}_{12}^H \mathbf{H}_{12} \mathbf{w}_{t1}}{\mathbf{w}_{t1}^H (\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11}^H \mathbf{H}_{11}) \mathbf{w}_{t1}} \quad (26)$$

and

$$\arg \max_{\mathbf{w}_{t2}} \frac{\mathbf{w}_{t2}^H \mathbf{H}_{21}^H \mathbf{H}_{21} \mathbf{w}_{t2}}{\mathbf{w}_{t2}^H (\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22}^H \mathbf{H}_{22}) \mathbf{w}_{t2}} \quad (27)$$

These are generalized Rayleigh quotient problems, and the optimal transmit AWVs (before normalization) are

$$\mathbf{w}_{t1} = \text{pEigVect} \left(\left(\mathbf{I} + \varepsilon_{11} \mathbf{H}_{11}^H \mathbf{H}_{11} \right)^{-1} \mathbf{H}_{12}^H \mathbf{H}_{12} \right) \quad (28)$$

and

$$\mathbf{w}_{t2} = \text{pEigVect} \left(\left(\mathbf{I} + \varepsilon_{22} \mathbf{H}_{22}^H \mathbf{H}_{22} \right)^{-1} \mathbf{H}_{21}^H \mathbf{H}_{21} \right) \quad (29)$$

respectively, where $\text{pEigVect}(\mathbf{X})$ denotes the principal eigenvector of \mathbf{X} .

In brief, by exploiting the proposed LB-MMSE, the transmit AWVs are computed as Equations (28) and (29), respectively. Based on the transmit AWVs, the receive AWVs (before normalization) are found as Equation (24).

4.2. SI-ZF-MRT

In this scheme, we first consider transmit beamforming by adopting MRT, and then use ZF to suppress the SI for receive beamforming. This approach is referred to as SI-ZF-MRT.

By exploiting MRT for transmit beamforming, we have

$$\mathbf{w}_{t1} = \mathbf{H}_{12}^H \mathbf{w}_{r2}; \quad \mathbf{w}_{t2} = \mathbf{H}_{21}^H \mathbf{w}_{r1} \quad (30)$$

To suppress the SI, we have

$$\mathbf{w}_{r1}^H \mathbf{H}_{11} \mathbf{H}_{12}^H \mathbf{w}_{r2} = 0 = \mathbf{w}_{r2}^H \mathbf{H}_{22} \mathbf{H}_{21}^H \mathbf{w}_{r1} \quad (31)$$

There are many solutions of \mathbf{w}_{r1} and \mathbf{w}_{r2} for Equation (31). Among those, we want to find the receive AWVs to maximize JAR while satisfying Equation (31). With Equations (30) and (31), the JAR becomes

$$R = \log_2 \left(1 + \varepsilon_{21} |\mathbf{w}_{r1}^H \mathbf{H}_{21} \mathbf{H}_{21}^H \mathbf{w}_{r1}|^2 \right) + \log_2 \left(1 + \varepsilon_{12} |\mathbf{w}_{r2}^H \mathbf{H}_{12} \mathbf{H}_{12}^H \mathbf{w}_{r2}|^2 \right) \quad (32)$$

There are two options to design \mathbf{w}_{r1} and \mathbf{w}_{r2} . One is to first derive \mathbf{w}_{r1} that maximizes Equation (32) without considering the constraint Equation (31); then derive \mathbf{w}_{r2} to optimize Equation (32) with Equation (31) satisfied. With this option, we have

$$\mathbf{w}_{r1} = \text{pEigVect} \left(\mathbf{H}_{21} \mathbf{H}_{21}^H \right) \quad (33)$$

and

$$\mathbf{w}_{r2} = \mathbf{a} - \left\langle \mathbf{a}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} - \langle \mathbf{a}, \mathbf{b}_2 \rangle \mathbf{b}_2 \quad (34)$$

where $\mathbf{a} = \text{pEigVect} \left(\mathbf{H}_{12} \mathbf{H}_{12}^H \right)$, $\mathbf{a}_1 = \mathbf{H}_{12} \mathbf{H}_{11}^H \mathbf{w}_{r1}$, and

$$\mathbf{b}_2 = \frac{\mathbf{H}_{22} \mathbf{H}_{21}^H \mathbf{w}_{r1} - \left\langle \mathbf{H}_{22} \mathbf{H}_{21}^H \mathbf{w}_{r1}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}}{\left\| \mathbf{H}_{22} \mathbf{H}_{21}^H \mathbf{w}_{r1} - \left\langle \mathbf{H}_{22} \mathbf{H}_{21}^H \mathbf{w}_{r1}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\|} \quad (35)$$

The other one is similar to the first one but with the positions of \mathbf{w}_{r1} and \mathbf{w}_{r2} exchanged. With this option, we have

$$\mathbf{w}_{r2} = \text{pEigVect} \left(\mathbf{H}_{12} \mathbf{H}_{12}^H \right) \quad (36)$$

and

$$\mathbf{w}_{r1} = \mathbf{a} - \left\langle \mathbf{a}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} - \langle \mathbf{a}, \mathbf{b}_2 \rangle \mathbf{b}_2 \quad (37)$$

where $\mathbf{a} = \text{pEigVect} \left(\mathbf{H}_{21} \mathbf{H}_{21}^H \right)$, $\mathbf{a}_1 = \mathbf{H}_{11} \mathbf{H}_{12}^H \mathbf{w}_{r2}$, and

$$\mathbf{b}_2 = \frac{\mathbf{H}_{21} \mathbf{H}_{22}^H \mathbf{w}_{r2} - \left\langle \mathbf{H}_{21} \mathbf{H}_{22}^H \mathbf{w}_{r2}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}}{\left\| \mathbf{H}_{21} \mathbf{H}_{22}^H \mathbf{w}_{r2} - \left\langle \mathbf{H}_{21} \mathbf{H}_{22}^H \mathbf{w}_{r2}, \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\rangle \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \right\|} \quad (38)$$

In brief, for SI-ZF-MRT, the two options are (i) to find the receive AWVs according to Equations (33) and (34), and then obtain the transmit AWVs according to Equation (30); (ii) to find the receive AWVs according to Equations (36) and (37), and then obtain the transmit AWVs according to Equation (30). Therefore, the one which has a higher JAR can be selected as the solution for SI-ZF-MRT.

Similar to SI-ZF-MRT, SI-ZF-maximum-ratio combining (MRC) is also applicable to the joint beamforming problem. With SI-ZF-MRC, receive beamforming is firstly performed by using MRC. Afterwards, transmit beamforming is carried out to maximize the JAR with the SI forced to zero. Since SI-ZF-MRC is similar to SI-ZF-MRT in formulation and performance, we do not present the details here.

4.3. Steering Beamforming

The conventional SBF in mmWave communication can also be introduced here to compare with the alternatives. SBF does not require full channel information. Instead, it only requires the knowledge of the transmit and steering vectors for the most significant MPC of the communication channel, and does not consider the SI. Suppose the m -th and ℓ -th multipath components are the most significant ones from Node #1 to Node #2 and from Node #2 to Node #1, respectively. By using SBF, the AWWs become

$$\begin{aligned} \mathbf{w}_{t2} &= \mathbf{h}_{21}(\theta_{\ell}^{(21)}), \quad \mathbf{w}_{t1} = \mathbf{h}_{12}(\theta_m^{(12)}) \\ \mathbf{w}_{r1} &= \mathbf{g}_{21}(\phi_{\ell}^{(21)}), \quad \mathbf{w}_{r2} = \mathbf{g}_{12}(\phi_m^{(12)}) \end{aligned} \quad (39)$$

By comparing the performance of SBF with those of the proposed schemes, we can see whether or not it is infeasible not to consider the SI in FD mmWave communication, and how much performance degradation it causes if we do not consider the SI in beamforming.

5. Simulation Results

In this section, we present the performances of all the involved schemes through numerical simulations, where the communication channel model and the SI channel model introduced in Section 2 are adopted. In all the evaluations, the near-field SI channels are deterministic, and are decided by Equation (5); while the far-field signal channels are random. Both LOS and non-LOS (NLOS) channels are considered for the communication channels. For NLOS channel, the transmit and receive steering angles are randomly generated within $[0, 2\pi)$, and the coefficients obey circularly symmetric complex Gaussian distribution with the same average power. For LOS channel, the LOS component has a fixed coefficient and fixed transmit and receive steering angles, while the other NLOS components have random steering angles and coefficients with average power 15 dB lower than the LOS component. The total power of a generated channel obeys Equation (3), and the total number of MPCs is 3 (We've also simulated with other numbers of MPCs, and similar results were obtained.). For each curve in all the figures in this section, we have generated 1000 realizations with the LOS or NLOS channel models, and computed the average JAR based on these realizations. Moreover, we have considered both types of array settings in the evaluations, namely Tx/Rx have separate antenna arrays and the same antenna array at a node. In all the simulations, $n_{t1} = n_{r1} = n_{t2} = n_{r2} = 32$.

Firstly, we consider the JAR and convergence performances of the ZF-Max-Power scheme with random initial transmit AWWs, which are shown in Figure 3 with relevant parameters listed in the caption. Both cases of separate arrays and the same array are included under LOS/NLOS channels. As ZF-Max-Power is an iterative method, we compare it with Max-SINR, which achieves the best performance within the typical solutions for the IA problem [35]. It is observed that ZF-Max-Power achieves a suboptimal performance close to the upper bound after convergence, under both LOS and NLOS channels, especially with separate arrays. The slight superiority of the case with separate arrays is due to that the different channel parameters of \mathbf{H}_{12} and \mathbf{H}_{21} provide more degrees of spatial freedoms for beamforming than the case of the same array, where the channel parameters are the same. Moreover, the convergence speed of ZF-Max-Power is fast in all the cases. Basically, only two iterations are required to achieve convergence. Interestingly, under the considered scenario, ZF-Max-Power achieves almost the same JAR and convergence performances as Max-SINR. However, it is noteworthy that ZF-Max-Power is a centralized iterative approach, where the iteration is performed at a certain node, and does not exploit channel reciprocity. By contrast, Max-SINR is a distributed iterative approach, where the iteration is performed at all distributed nodes by exploiting channel reciprocity. As a consequence, the convergence of ZF-Max-Power can be proven, while that of Max-SINR is still unproven in the literature [35] to the best of our knowledge. Moreover, ZF-Max-Power has a significantly lower computational complexity than Max-SINR.

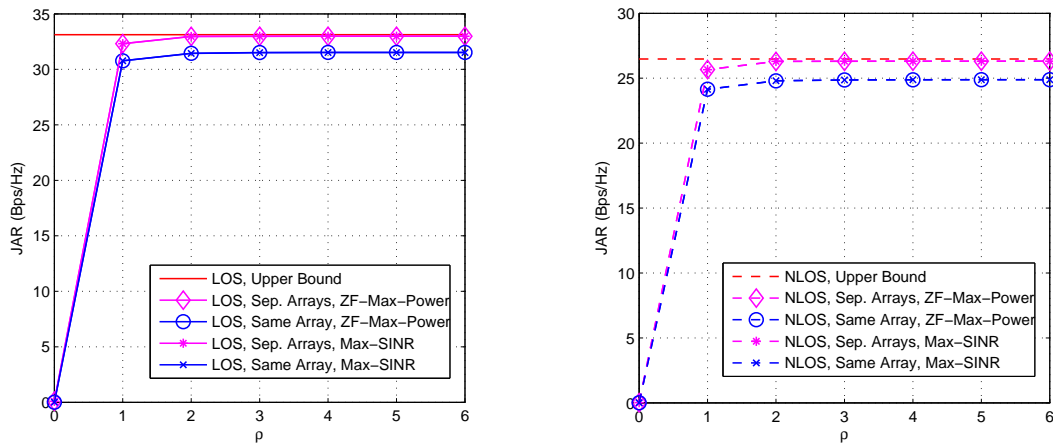


Figure 3. JAR and convergence performances of ZF-Max-Power with random initial transmit AWWs (**Left:** LOS channel, **Right:** NLOS channel). $\epsilon_{11} = \epsilon_{22} = 40$ dB. For LOS channel, $\epsilon_{12} = \epsilon_{21} = 20$ dB, while for NLOS channel, $\epsilon_{12} = \epsilon_{21} = 10$ dB. For the case of separate arrays, $d/\lambda = 1$ and $\omega = \pi/6$ rad, while for the case of sharing the same array, $d/\lambda = 0$ and $\omega = 0$ rad.

Next, let us see the JAR performances of the proposed schemes with respect to varying ω , d , i.e., the geometry of the arrays, when separate arrays are exploited at a node. Figure 4 shows the JAR performances with respect to varying ω under LOS and NLOS channels, respectively, where SI is fixed. The left hand side figure of Figure 5 shows the JAR performances with respect to d under LOS channel. Similar results can be observed under NLOS channel. As we can see, SI is assumed fixed in the left hand side figure of Figure 5, which may be not practically reasonable, because in practice d significantly affects SI. However, in order to adequately evaluate the effects of d on the JAR performance, we assume a fixed SI in the left hand side figure of Figure 5. For rigorousness, we also adopt varying SI in the right hand side figure of Figure 5, which is in accordance with the practice, where SI deteriorates with d^2 . It is noteworthy that the strength of SI is typically much higher than that of the desired signal, because SI comes from the local transmitter at the same node, while the designed signal comes from the remote transmitter at the other node. Relevant parameters for these figures are listed in the corresponding captions.

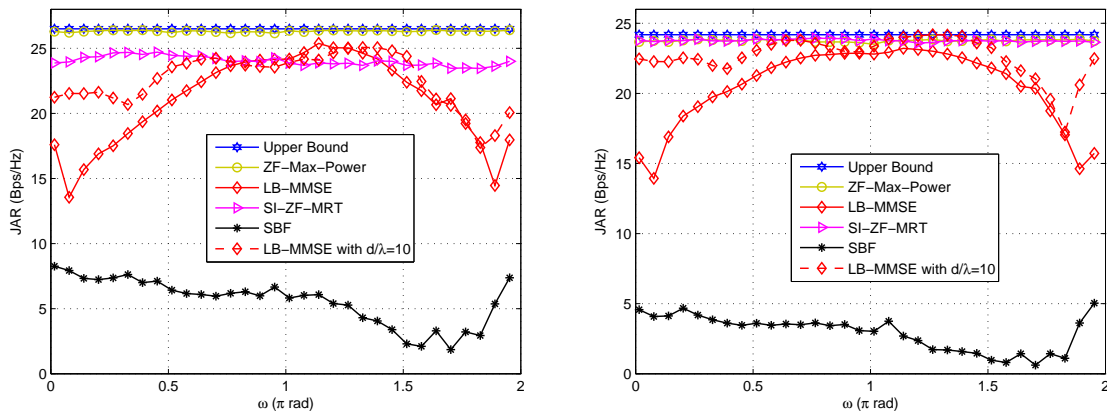


Figure 4. JAR performance of the involved schemes with respect to varying ω under LOS (**Left**) and NLOS (**Right**) channels in the case of separate arrays. $\epsilon_{11} = \epsilon_{22} = 40$ dB, $\epsilon_{12} = \epsilon_{21} = 10$ dB, $d/\lambda = 2$. For LB-MMSE, the JAR with $d/\lambda = 10$ is also plotted.

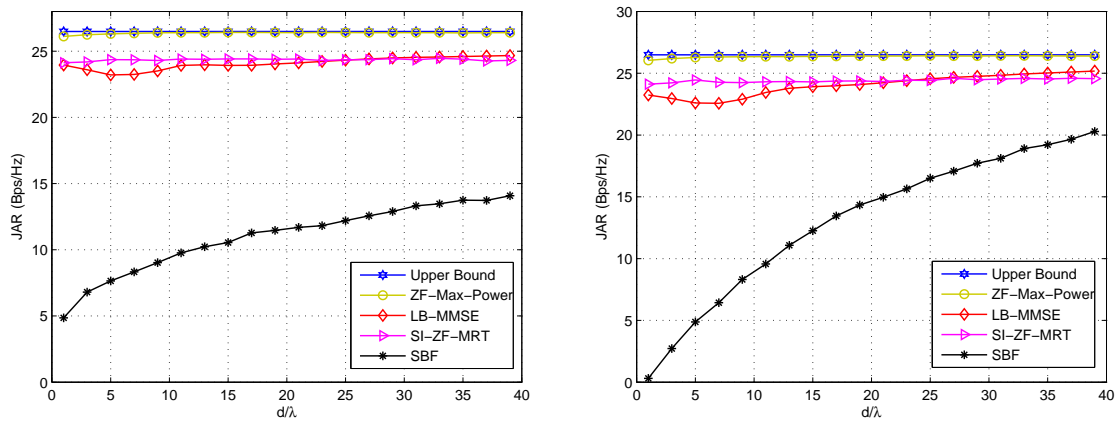


Figure 5. JAR performance of the involved schemes with respect to varying d under LOS channel in the case of separate arrays. $\omega = \pi$ rad. In the (Left) hand figure SI is assumed fixed, i.e., $\varepsilon_{11} = \varepsilon_{22} = 40$ dB, $\varepsilon_{12} = \varepsilon_{21} = 10$ dB; while in the (Right) hand figure SI varies with d/λ , i.e., $\varepsilon_{11} = \varepsilon_{22} = 60 - 20 \log_{10}(d/\lambda)$ dB, $\varepsilon_{12} = \varepsilon_{21} = 10$ dB.

By comparing these figures with each other, it can be observed that:

- (i) ZF-Max-Power is robust against ω , d and SI, and approaches the upper bound in all these cases. This is because ZF-Max-Power not only forces SI to zero, but also iteratively maximizes signal power. Thus, it achieves compelling performance that is insensitive to the geometry of the Tx/Rx arrays and SI.
- (ii) SI-ZF-MRT is also robust against ω , d and SI, thanks to its zero-forcing filtering to SI. In addition, it also achieves an acceptable performance, which is close to the upper bound. It is noted that the JAR gap between SI-ZF-MRT and the upper bound is greater under LOS channel than that under NLOS channel. This phenomenon can be explained by referring to Equation (34), where \mathbf{w}_{r2} is in fact set within an $(n_{r2} - 2)$ -dimension subspace, due to the two zero forcing equations shown in Equation (31). Clearly if \mathbf{a} in Equation (34), which represents the dimension with the largest power of the channel, has less energy projected on the $(n_{r2} - 2)$ -dimension subspace, the JAR performance will be poorer. Under LOS channel, the majority of the channel energy concentrates on a single path, or a single dimension. Once this dimension has a small projection on the subspace, the performance will be poor. In contrast, under NLOS channel the channel energy evenly disperses on multiple paths. Only when all of these paths have a small projection on the subspace, the performance will be poor. In other words, the probability of a poor performance is lower under NLOS channel than that under LOS channel. Hence, on average the JAR gap between SI-ZF-MRT and the upper bound is greater under LOS channel than that under NLOS channel.
- (iii) LB-MMSE is sensitive to ω and d . From Figure 4 we observe that the performance of LB-MMSE fluctuates as ω changes, and the fluctuation is different for different d . From Figure 5 we observe that the performance of LB-MMSE has a U-shape as d increases, but behaves stable when d is large. To understand these, we need to go back to Equations (28) and (29). From these two equations we can see that the transmit AWWs are decided to maximize the SINR rather than minimize SI based on the local information. Taking Equation (28) for illustration, since usually ε_{11} is big, when \mathbf{H}_{11} has a low rank, the eigenvector of $\mathbf{H}_{12}^H \mathbf{H}_{12}$ has a high probability to locate within the null space of \mathbf{H}_{11} . In such a case, a high signal power can be achieved while little SI locates within the signal subspace; thus good performance is achieved. Note that this statement is just for illustration. In practice, \mathbf{H}_{11} is generally with full rank except when $\omega = 0$ or π . However, when most energy of \mathbf{H}_{11} locates at a low-dimensional subspace, the situation will be similar to the statement that \mathbf{H}_{11} has a low rank. In comparison, when \mathbf{H}_{11} has a high or even full rank, SI will almost unavoidably locate within the signal subspace and affects the received SINR, and thus the performance will be

poor. When d is small, the energy dispersion of \mathbf{H}_{11} is sensitive to ω and d according to the SI channel model, and thus the JAR performance is also sensitive to ω and d . However, when d is large, the SI channel almost reduces to a directional channel with rank 1, and thus SI has a low probability to locate within the signal subspace. In such a case, LB-MMSE can stably achieve a near-optimal performance.

- (iv) SBF is also sensitive to ω , d and SI. This is because SBF does not even consider SI in the beamforming design. Meanwhile, from Figure 5 it is found that SBF becomes improved as d increases. In the right hand side of Figure 5 the improving speed of SBF is faster than that in Figure 5, because SI is reduced as d increases. This phenomenon suggests that when the near-field SI channel gradually reduces to a directional channel, the conventional beamforming schemes that to simply steer towards each other may also achieve good performance, because usually the communication channel and SI channel have difference steering angles. However, in practical FD mmWave communication where d is generally small, the SI channel does not have the feature of directivity; thus SBF is much poorer than the other candidates, and the performance of SBF does not show monotonicity with ω , as shown in Figure 4. Thus, SBF may not be a good choice for FD mmWave communication, where SI must be taken into account.

Then, we compare the JAR performances of the discussed schemes with separate arrays and the same array. Figure 6 shows the comparison results with respect to SI under LOS and NLOS channels, respectively, where relevant parameters are listed in the captions. From these two figures we observe that the schemes with separate arrays basically achieve better performance than those with the same array. This advantage is also due to that the different channel parameters of \mathbf{H}_{12} and \mathbf{H}_{21} when using separate arrays provide larger degrees of spatial freedom for beamforming than the case of using the same array, where the channel parameters are the same. Moreover, both ZF-Max-Power and SI-ZF-MRT are insensitive to the increase of SI, thanks to the operation of ZF SI, while the performance of SBF becomes poorer as the increase of SI, due to no operation of ZF SI. Interestingly, the JAR of LB-MMSE with separate arrays slowly decreases as the increase of SI whereas that with the same array changes little, which shows that LB-MMSE with the same array is more robust against the SI. To explain this, let us look at the third and fourth lines of Equation (25). In third line, $\frac{\mathbf{w}_{t2}^H \mathbf{H}_{21}^H \mathbf{H}_{21} \mathbf{w}_{t2}}{\mathbf{w}_{t1}^H (\mathbf{I} + \epsilon_{11} \mathbf{H}_{11}^H \mathbf{H}_{11}) \mathbf{w}_{t1}}$ and $\frac{\mathbf{w}_{t1}^H \mathbf{H}_{12}^H \mathbf{H}_{12} \mathbf{w}_{t1}}{\mathbf{w}_{t2}^H (\mathbf{I} + \epsilon_{22} \mathbf{H}_{22}^H \mathbf{H}_{22}) \mathbf{w}_{t2}}$ can be roughly seen as the receive SINRs at Node #1 and Node #2 without considering the receive AWWs, respectively. However, in order to obtain closed-form expressions of the transmission AWWs, the denominators (or numerators) of these two components are exchanged and optimized respectively, as shown in Equations (26) and (27). This means that the optimizations in Equations (26) and (27) are not to directly optimize the receive SINRs at Node #1 and Node #2. Hence, in general LB-MMSE is not so robust against the SI. However, in the case with the same array, the link from Node #1 to Node #2 is symmetric; thus the denominators (or numerators) of the two components in the third line of Equation (25) can be seen equal or at least proportional to each other. In such a case, Equations (26) and (27) are in fact to optimize the receive SINRs at Node #1 and Node #2 without considering the receive AWWs. Hence, LB-MMSE with the same array is relatively more robust against the SI.

Finally, we evaluate the effects of channel estimation errors on the proposed schemes. For the SI channel, there exists Gaussian error; while for the communication channel, it is possible to miss some MPCs during the beam search process. Figure 7 shows the effects of these estimation errors on the proposed schemes with separate arrays (the results are similar with the same array) under LOS and NLOS channels, respectively, where the parameters are specified in the captions. From these two figures we can observe that both ZF-Max-Power and LB-MMSE are relatively robust against the channel estimation error. Even only one MPC is acquired, they can achieve promising performance, especially under LOS channel. However, SI-ZF-MRT is not robust against the estimation error of the communication channel, i.e., if only one MPC is acquired, the performance of SI-ZF-MRT

becomes rather poor, this is because the full channel information is involved in the SI ZF operation according to Equation (31).

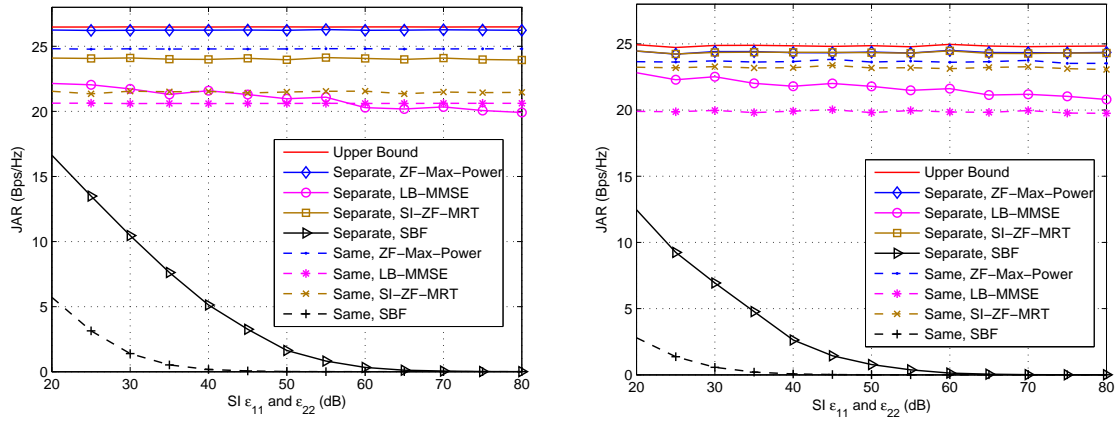


Figure 6. JAR comparison between different array settings (separate arrays versus the same array) under LOS (Left) and NLOS (Right) channels with varying SI. $\varepsilon_{12} = \varepsilon_{21} = 10$ dB. For the case of separate arrays, $\omega = 0.6\pi$ rad, $d/\lambda = 1$.

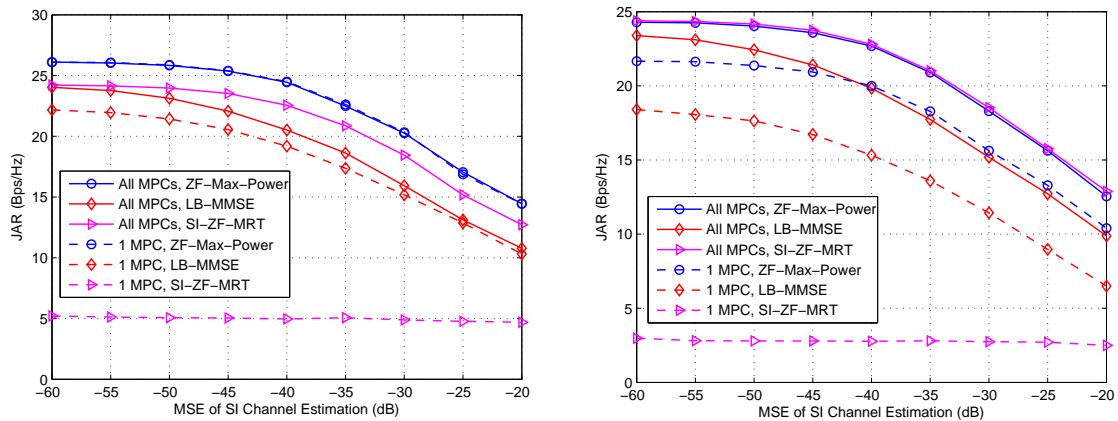


Figure 7. Effects of channel estimation errors on the proposed schemes with separate arrays under LOS (Left) and NLOS (Right) channels. $\varepsilon_{12} = \varepsilon_{21} = 10$ dB, $\varepsilon_{11} = \varepsilon_{22} = 40$ dB, $\omega = \pi$ rad, $d/\lambda = 1$.

It is noteworthy that circuit imperfections are not taken into consideration in our system model, i.e., in Equations (7) and (8). In a practical FD system, there are always Tx/Rx hardware and implementation imperfections, including low-noise amplifier (LNA) noise figure, phase noise, in-phase and quadrature (IQ) mismatch, nonlinear distortion of power amplifier, etc. These imperfections may be more severe in mmWave communication systems than in low-frequency systems, because of the higher carrier frequency and larger bandwidth. Since SI is strong, the Tx imperfections, which are carried by the transmitted signals s_1 and s_2 in Equations (7) and (8), will also arrive at the Rx. When beamforming cancellation is adopted to force the SI to zero, the performance will be affected little by the Tx circuit imperfections provided that the AWV control is perfect, because all the SI, including the imperfections, can be filtered out by beamforming. However, in practice, the AWV control may have error. In such a case the SI cannot be completely filtered out by beamforming, and the residual SI, which contains Tx imperfections, will degrade the system performance. Although baseband (BB) cancellation can be used to deal with the residual SI after beamforming cancellation, it basically cannot effectively cancel the residual Tx imperfections.

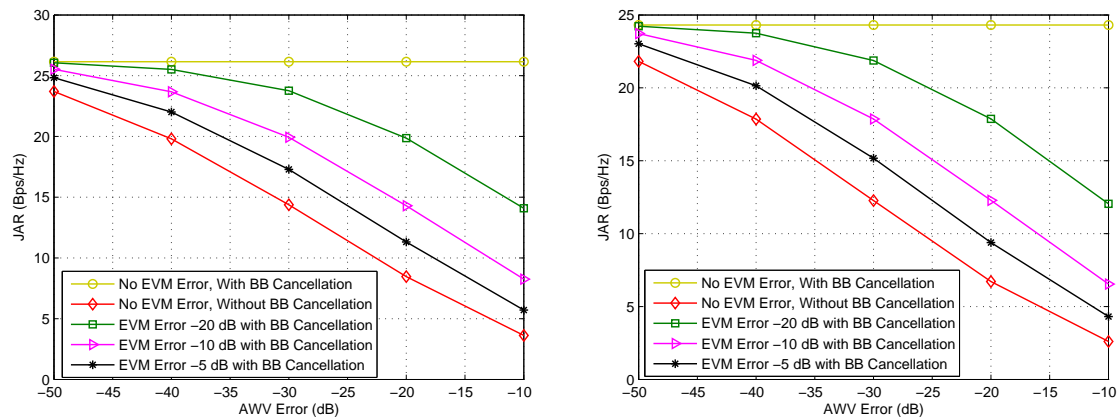


Figure 8. Effects of AWW error and EVM error on the JAR performance of ZF-Max-Power under LOS (Left) and NLOS (Right) channels. $\varepsilon_{12} = \varepsilon_{21} = 10$ dB, $\varepsilon_{11} = \varepsilon_{22} = 50$ dB, $\omega = 0.8\pi$ rad, $d/\lambda = 1$.

To further illustrate the effect of the circuit imperfections on the system performance, we model all the typical Tx imperfections as a zero-mean Gaussian distributed error vector magnitude (EVM) noise [57], and its average power can be measured in dB with respect to the transmission power. On the other hand, we also need to consider AWW control error, which can also be modeled as a zero-mean Gaussian variable, and its average power can also be measured in dB with respect to the 2-norm of its corresponding weight. By exploiting this model, we can evaluate the effects of AWW error and EVM error on the JAR performance of ZF-Max-Power as shown in Figure 8, where relevant parameters are listed in the caption. The effects are similar to the performances of the other schemes. From this figure we can find that when AWW control is perfect or AWW error is small enough, the SI as well as the EVM noise can be mitigated successfully by beamforming. However, when there is significant AWW error, the system performance deteriorates as the AWW error becomes greater. In such a case, BB cancellation is needed to cancel the residual SI. From the figure we observe that with BB cancellation, the performance is greatly improved and does not depend on the AWW error. On the other hand, BB cancellation can hardly mitigate the residual EVM noise, because it is difficult to estimate relevant parameters of the EVM noise [57]. Hence, we can observe that even when BB cancellation is adopted, if EVM noise exists, the performance still deteriorates as the AWW error becomes greater.

6. Conclusions

In this paper, we investigate FD mmWave communications, where we employ beamforming to cancel SI, and study a JTR-BF problem in the presence of significant SI. As the problem of finding the optimal beamforming vectors to maximize the JAR is non-convex, several suboptimal solutions are proposed. Firstly, ZF-Max-Power, which restricts the original problem by ZF SI and alternatively optimizes the desired power, is proposed, and its convergence is proven. It is shown that the computational complexity of ZF-Max-Power is lower than that of Max-SINR by one order of magnitude. Next, two closed-form solutions, namely LB-MMSE and SI-ZF-MRT are proposed, by jointly using MMSE, ZF and MRT criteria. Performance evaluations show that ZF-Max-Power approaches an upper bound on the JAR, and it needs only 2 iterations on average to achieve the convergence with random initial points. LB-MMSE and SI-ZF-MRT achieve suboptimal performances. In addition, we find that ZF-Max-Power and SI-ZF-MRT are robust against the geometry of Tx/Rx antenna arrays due to the operation of ZF SI, while LB-MMSE is not. ZF-Max-Power and LB-MMSE are robust against channel estimation error, while SI-ZF-MRT is not. Furthermore, these schemes basically achieve better JAR performance with a separate-array setting than those sharing the same array. The results demonstrate the feasibility of FD mmWave communications and the effectiveness of beamforming cancellation.

Acknowledgments: The authors would like to thank the funding support from the National Natural Science Foundation of China (NSFC) under grant Nos. 61571025, 91538204, 91338106, and 61231013, National Key R&D Program under grant No. 2016YFB1200100, National Basic Research Program of China under grant No.2011CB707000, and Foundation for Innovative Research Groups of the National Natural Science Foundation of China under grant No. 61221061.

Author Contributions: Z.X., J.C., P.X. and X.-G.X. conceived and designed the idea and simulations; X.L. and L.B. performed the simulations; X.L. and Z.X. analyzed the performance; all the authors contributed to writing and improving the paper.

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

Appendix A. Proof of Lemma 1

Since $\{\mathbf{a}_i \in \mathbb{C}^{L \times 1}\}_{i=1,2,\dots,N}; N < L$ is a group of linear independent vectors, they span a linear subspace denoted by $\mathbb{S} \triangleq \mathcal{S}\{\mathbf{a}_i | i=1,2,\dots,N\}$. Suppose $\{\mathbf{v}_i\}_{i=1,2,\dots,N}$ is an orthogonal basis of the linear subspace \mathbb{S} . Then constraint $\langle \mathbf{a}, \mathbf{a}_i \rangle = 0 |_{i=1,2,\dots,N}$ is equivalent to $\langle \mathbf{a}, \mathbf{v}_i \rangle = 0 |_{i=1,2,\dots,N}$.

Remark The vector $\mathbf{b} \in \mathbb{C}^{L \times 1}$ to maximize $|\langle \mathbf{b}, \mathbf{a} \rangle|^2$ with the constraint that $\langle \mathbf{b}, \mathbf{v}_i \rangle = 0 |_{i=1,2,\dots,N}$ and $\|\mathbf{b}\| = 1$ is (before normalization) $\mathbf{b} = \mathbf{a} - \sum_{i=1}^N \langle \mathbf{a}, \mathbf{v}_i \rangle \mathbf{v}_i$.

According to the above remark, the next thing we need to prove is that $\{\mathbf{b}_i | i=1,2,\dots,N\}$ shown in Equation (15) is an orthogonal basis of $\{\mathbf{b}_i | i=1,2,\dots,N\}$. This proposition holds because $\{\mathbf{b}_i | i=1,2,\dots,N\}$ is obtained according to the standard Gram-Schmidt process from $\{\mathbf{a}_i | i=1,2,\dots,N\}$. Therefore, Lemma 1 holds.

Appendix B. Proof of Theorem 1

Under the constraints in Equation (14), the JAR becomes

$$R(\mathbf{w}_{t1}, \mathbf{w}_{t2}, \mathbf{w}_{r1}, \mathbf{w}_{r2}) = \log_2 \left(1 + \varepsilon_{21} |\mathbf{w}_{r1}^H \mathbf{H}_{21} \mathbf{w}_{t2}|^2 \right) + \log_2 \left(1 + \varepsilon_{12} |\mathbf{w}_{r2}^H \mathbf{H}_{12} \mathbf{w}_{t1}|^2 \right) \triangleq R_{ub} \quad (B1)$$

which shows that R is a function of \mathbf{w}_{t1} , \mathbf{w}_{t2} , \mathbf{w}_{r1} , and \mathbf{w}_{r2} .

Let $\mathbf{w}_{t1}^{(n)}$ and $\mathbf{w}_{t2}^{(n)}$ denote the transmit AWVs after the n -th iteration, and $\mathbf{w}_{r1}^{(n)}$ and $\mathbf{w}_{r2}^{(n)}$ the receive AWVs after the n -th iteration. Then we have $R^{(n)} = R(\mathbf{w}_{t1}^{(n)}, \mathbf{w}_{t2}^{(n)}, \mathbf{w}_{r1}^{(n)}, \mathbf{w}_{r2}^{(n)})$ and $R^{(n+1)} = R(\mathbf{w}_{t1}^{(n+1)}, \mathbf{w}_{t2}^{(n+1)}, \mathbf{w}_{r1}^{(n+1)}, \mathbf{w}_{r2}^{(n+1)})$.

According to Algorithms 1, as well as Equations (19) and (20), we have

$$R(\mathbf{w}_{t1}^{(n)}, \mathbf{w}_{t2}^{(n)}, \mathbf{w}_{r1}^{(n+1)}, \mathbf{w}_{r2}^{(n+1)}) \geq R(\mathbf{w}_{t1}^{(n)}, \mathbf{w}_{t2}^{(n)}, \mathbf{w}_{r1}^{(n)}, \mathbf{w}_{r2}^{(n)}) = R^{(n)} \quad (B2)$$

Also, according to Algorithms 1, as well as Equations (21) and (22), we have

$$R(\mathbf{w}_{t1}^{(n+1)}, \mathbf{w}_{t2}^{(n+1)}, \mathbf{w}_{r1}^{(n+1)}, \mathbf{w}_{r2}^{(n+1)}) = R^{(n+1)} \geq R(\mathbf{w}_{t1}^{(n)}, \mathbf{w}_{t2}^{(n)}, \mathbf{w}_{r1}^{(n+1)}, \mathbf{w}_{r2}^{(n+1)}) \quad (B3)$$

Therefore, we have $R^{(n+1)} \geq R^{(n)}$.

Appendix C. Lemmas 2, 3, 4

Lemma 2. Given $\mathbf{R} \in \mathbb{C}^{M \times M}$ a positive definite Hermitian matrix, the solution to maximize $\frac{\mathbf{x}^H \mathbf{a} \mathbf{a}^H \mathbf{x}}{\mathbf{x}^H \mathbf{R} \mathbf{x}}$ is $\mathbf{x} = \mathbf{R}^{-1} \mathbf{a}$.

Proof. Since \mathbf{R} is a positive definite Hermitian matrix, $\mathbf{R}^{1/2}$ is also an invertible Hermitian matrix. Thus, we have

$$\frac{\mathbf{x}^H \mathbf{a} \mathbf{a}^H \mathbf{x}}{\mathbf{x}^H \mathbf{R} \mathbf{x}} = \frac{\mathbf{x}^H \mathbf{a} \mathbf{a}^H \mathbf{x}}{(\mathbf{R}^{1/2} \mathbf{x})^H (\mathbf{R}^{1/2} \mathbf{x})} \quad (C1)$$

Let $\tilde{\mathbf{x}} = \mathbf{R}^{1/2}\mathbf{x}$, the problem becomes to maximize $\frac{\tilde{\mathbf{x}}^H(\mathbf{R}^{-H/2}\mathbf{a})(\mathbf{R}^{-H/2}\mathbf{a})^H\tilde{\mathbf{x}}}{\tilde{\mathbf{x}}^H\tilde{\mathbf{x}}}$. Therefore, the solution is $\tilde{\mathbf{x}} = \mathbf{R}^{-H/2}\mathbf{a} = \mathbf{R}^{-1/2}\mathbf{a}$, i.e., $\mathbf{x} = \mathbf{R}^{-1}\mathbf{a}$. \square

Lemma 3. Given $\mathbf{R} \in \mathbb{C}^{M \times M}$ a positive definite Hermitian matrix, $\mathbf{a}^H \mathbf{R}^{-1} \mathbf{a} \geq \frac{(\mathbf{a}^H \mathbf{a})^2}{\mathbf{a}^H \mathbf{R} \mathbf{a}}$.

Proof. Since $\mathbf{R} \in \mathbb{C}^{M \times M}$ is a positive definite Hermitian matrix, it can be expressed as $\mathbf{R} = \sum_{i=1}^M \lambda_i \mathbf{v}_i \mathbf{v}_i^H$, where $\lambda_i > 0$ and $\{\mathbf{v}_i\}_{i=1,2,\dots,M}$ constitutes an orthogonal base in $\mathbb{C}^{M \times M}$. Thus, \mathbf{a} can be expressed as $\mathbf{a} = \sum_{i=1}^M \alpha_i \mathbf{v}_i$. Besides, we have $\mathbf{R}^{-1} = \sum_{i=1}^M \frac{1}{\lambda_i} \mathbf{v}_i \mathbf{v}_i^H$. Thus,

$$\begin{aligned} & (\mathbf{a}^H \mathbf{R}^{-1} \mathbf{a}) (\mathbf{a}^H \mathbf{R} \mathbf{a}) \\ &= \left(\sum_{i=1}^M \frac{|\alpha_i|^2}{\lambda_i} \right) \left(\sum_{i=1}^M |\alpha_i|^2 \lambda_i \right) \\ &= \sum_{i=1}^M |\alpha_i|^4 + \sum_{i=1}^M \sum_{j=1}^{i-1} |\alpha_i|^2 |\alpha_j|^2 \left(\frac{\lambda_j}{\lambda_i} + \frac{\lambda_i}{\lambda_j} \right) \\ &\geq \sum_{i=1}^M |\alpha_i|^4 + 2 \sum_{i=1}^M \sum_{j=1}^{i-1} |\alpha_i|^2 |\alpha_j|^2 \\ &= (\mathbf{a}^H \mathbf{a})^2 \end{aligned} \tag{C2}$$

\square

Lemma 4. Given $\mathbf{R} \in \mathbb{C}^{M \times M}$ a positive definite Hermitian matrix with the formulation $\mathbf{R} = \mathbf{I} + \epsilon \mathbf{b} \mathbf{b}^H$, $\frac{\mathbf{a}^H \mathbf{a}}{\mathbf{a}^H \mathbf{R} \mathbf{a}} \geq \frac{1}{1 + \epsilon \mathbf{b}^H \mathbf{b}}$.

Proof. Let $\mathbf{v} = \mathbf{a} / \|\mathbf{a}\|$. We have

$$\begin{aligned} \frac{\mathbf{a}^H \mathbf{a}}{\mathbf{a}^H \mathbf{R} \mathbf{a}} &= \frac{\mathbf{a}^H \mathbf{a}}{\mathbf{a}^H (\mathbf{I} + \epsilon \mathbf{b} \mathbf{b}^H) \mathbf{a}} = \frac{1}{\mathbf{v}^H (\mathbf{I} + \epsilon \mathbf{b} \mathbf{b}^H) \mathbf{v}} \\ &= \frac{1}{1 + \epsilon |\mathbf{b}^H \mathbf{v}|^2} \geq \frac{1}{1 + \epsilon \frac{\mathbf{b}^H \mathbf{b}}{\|\mathbf{b}\|^2}} = \frac{1}{1 + \epsilon \mathbf{b}^H \mathbf{b}} \end{aligned} \tag{C3}$$

\square

References

1. Sabharwal, A.; Schniter, P.; Guo, D.; Bliss, D.; Rangarajan, S.; Wichman, R. In-band full-duplex wireless: Challenges and opportunities. *IEEE J. Sel. Areas Commun.* **2014**, *32*, 1637–1652.
2. Ahmed, E.; Eltawil, A.; Sabharwal, A. Simultaneous transmit and sense for cognitive radios using full-duplex: A first study. In Proceedings of the IEEE Antennas and Propagation Society International Symposium (APSURSI), Chicago, IL, USA, 8–14 July 2012; IEEE: Chicago, IL, USA, 2012; pp. 1–2.
3. Duarte, M.; Dick, C.; Sabharwal, A. Experiment-driven characterization of full-duplex wireless systems. *IEEE Trans. Wirel. Commun.* **2011**, *11*, 4296–4307.
4. Duarte, M.; Sabharwal, A.; Aggarwal, V.; Jana, R.; Ramakrishnan, K.; Rice, C.; Shankaranarayanan, N. Design and characterization of a full-duplex multiantenna system for WiFi networks. *IEEE Trans. Veh. Technol.* **2014**, *63*, 1160–1177.
5. Jain, M.; Choi, J.I.; Kim, T.; Bharadia, D.; Seth, S.; Srinivasan, K.; Levis, P.; Katti, S.; Sinha, P. Practical, real-time, full duplex wireless. In Proceedings of the 17th Annual International Conference on Mobile Computing and Networking, Las Vegas, NV, USA, 19–23 September 2011; ACM: New York, NY, USA, 2011; pp. 301–312.

6. Li, N.; Zhu, W.; Han, H. Digital interference cancellation in single channel, full duplex wireless communication. In Proceedings of the 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), Shanghai, China, 21–23 September 2012; IEEE: Shanghai, China, 2012; pp. 1–4.
7. Li, S.; Murch, R.D. Full-duplex wireless communication using transmitter output based echo cancellation. In Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM 2011), Houston, TX, USA, 5–9 December 2011; IEEE: Houston, TX, USA, 2011; pp. 1–5.
8. Choi, J.I.; Jain, M.; Srinivasan, K.; Levis, P.; Katti, S. Achieving single channel, full duplex wireless communication. In Proceedings of the 16th Annual International Conference on Mobile Computing and Networking (MobiCom 2010), Chicago, IL, USA, 20–24 September 2010; IEEE: Chicago, IL, USA, 2010; pp. 301–312.
9. Khojastepour, M.A.; Sundaresan, K.; Rangarajan, S.; Zhang, X.; Barghi, S. The case for antenna cancellation for scalable full-duplex wireless communications. In Proceedings of the 10th ACM Workshop on Hot Topics in Networks, Cambridge, MA, USA, 14–15 November 2011; ACM: New York, NY, USA, 2011; pp. 1–6.
10. Bharadia, D.; McMilin, E.; Katti, S. Full duplex radios. In Proceedings of the Special Interest Group on Data Communication (SIGCOMM), Hong Kong, China, 12–16 August 2013; ACM: Hong Kong, China, 2013.
11. Daniels, R.C.; Murdock, J.N.; Rappaport, T.S.; Heath, R.W. 60 GHz wireless: Up close and personal. *IEEE Microw. Mag.* **2010**, *11*, 44–50.
12. Perahia, E.; Cordeiro, C.; Park, M.; Yang, L.L. IEEE 802.11 ad: Defining the next generation multi-Gbps Wi-Fi. In Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, USA, 9–12 January 2010; IEEE: Las Vegas, NV, USA, 2010; pp. 1–5.
13. Yong, S.K.; Xia, P.; Valdes-Garcia, A. *60GHz Technology for Gbps WLAN and WPAN: From Theory to Practice*; Wiley: West Sussex, UK, 2011.
14. Wang, P.; Li, Y.; Yuan, X.; Song, L.; Vucetic, B. Tens of gigabits wireless communications over E-band LoS MIMO channels with uniform linear antenna arrays. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 3791–3805.
15. Wang, P.; Li, Y.; Song, L.; Vucetic, B. Multi-gigabit millimeter wave wireless communications for 5G: From fixed access to cellular networks. *IEEE Commun. Mag.* **2015**, *53*, 168–178.
16. Gao, X.; Dai, L.; Han, S.; Chih-Lin, I.; Heath, R.W. Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 998–1009.
17. Gao, Z.; Dai, L.; Wang, Z.; Chen, S. Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO. *IEEE Trans. Signal Process.* **2015**, *63*, 6169–6183.
18. Gao, Z.; Dai, L.; Lu, Z.; Yuen, C.; Wang, Z. Super-resolution sparse MIMO-OFDM channel estimation based on spatial and temporal correlations. *IEEE Commun. Lett.* **2014**, *18*, 1266–1269.
19. Niu, Y.; Li, Y.; Jin, D.; Su, L.; Vasilakos, A.V. A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges. *Wirel. Netw.* **2015**, *21*, 2657–2676.
20. Tang, Y.; Vucetic, B.; Li, Y. An iterative singular vectors estimation scheme for beamforming transmission and detection in MIMO systems. *IEEE Commun. Lett.* **2005**, *9*, 505–507.
21. Xiao, Z. Suboptimal Spatial Diversity Scheme for 60 GHz Millimeter-Wave WLAN. *IEEE Commun. Lett.* **2013**, *17*, 1790–1793.
22. Xiao, Z.; Bai, L.; Choi, J. Iterative Joint Beamforming Training with Constant-Amplitude Phased Arrays in Millimeter-Wave Communications. *IEEE Commun. Lett.* **2014**, *18*, 829–832.
23. Park, M.; Pan, H. A Spatial Diversity Technique for IEEE 802.11 ad WLAN in 60 GHz Band. *IEEE Commun. Lett.* **2012**, *16*, 1260–1262.
24. Nsenga, J.; Van Thillo, W.; Horlin, F.; Ramon, V.; Bourdoux, A.; Lauwereins, R. Joint transmit and receive analog beamforming in 60 GHz MIMO multipath channels. In Proceedings of the 2009 IEEE International Conference on Communications (ICC), Dresden, Germany, 14–18 June 2009; IEEE: Dresden, Germany, 2009; pp. 1–5.
25. Li, B.; Zhou, Z.; Zhang, H.; Nallanathan, A. Efficient beamforming training for 60-GHz millimeter-wave communications: A novel numerical optimization framework. *IEEE Trans. Veh. Technol.* **2014**, *63*, 703–717.
26. Li, B.; Zhou, Z.; Zou, W.; Sun, X.; Du, G. On the Efficient Beam-Forming Training for 60GHz Wireless Personal Area Networks. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 504–515.

27. Chun, B.; Lee, Y.H. A spatial self-interference nullification method for full duplex amplify-and-forward MIMO relays. In Proceedings of the 2010 IEEE Wireless Communications and Networking Conference (WCNC), Sydney, Australia, 18–21 April 2010; IEEE: Sydney, Australia, 2010; pp. 1–6.
28. Riihonen, T.; Werner, S.; Wichman, R. Mitigation of loopback self-interference in full-duplex MIMO relays. *IEEE Trans. Signal Process.* **2011**, *59*, 5983–5993.
29. Liu, Y.; Xia, X.G.; Zhang, H. Distributed space-time coding for full-duplex asynchronous cooperative communications. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 2680–2688.
30. Liu, Y.; Xia, X.G.; Zhang, H. Distributed linear convolutional space-time coding for two-relay full-duplex asynchronous cooperative networks. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 6406–6417.
31. Snow, T.; Fulton, C.; Chappell, W.J. Transmit–receive duplexing using digital beamforming system to cancel self-interference. *IEEE Trans. Microw. Theory Tech.* **2011**, *59*, 3494–3503.
32. Gomadam, K.; Cadambe, V.R.; Jafar, S.A. A distributed numerical approach to interference alignment and applications to wireless interference networks. *IEEE Trans. Inf. Theory* **2011**, *57*, 3309–3322.
33. Santamaria, I.; Gonzalez, O.; Heath, R.; Peters, S.W. Maximum sum-rate interference alignment algorithms for MIMO channels. In Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM 2010), Miami, FL, USA, 6–10 December 2010; IEEE: Miami, FL, USA, 2010; pp. 1–6.
34. Peters, S.W.; Heath, R.W. Cooperative algorithms for MIMO interference channels. *IEEE Trans. Veh. Technol.* **2011**, *60*, 206–218.
35. Schmidt, D.A.; Shi, C.; Berry, R.A.; Honig, M.L.; Utschick, W. Comparison of distributed beamforming algorithms for MIMO interference networks. *IEEE Trans. Signal Process.* **2013**, *61*, 3476–3489.
36. Wang, J.; Lan, Z.; Pyo, C.; Baykas, T.; Sum, C.; Rahman, M.; Gao, J.; Funada, R.; Kojima, F.; Harada, H. Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems. *IEEE J. Sel. Areas Commun.* **2009**, *27*, 1390–1399.
37. Maltsev, A.; Maslennikov, R.; Sevastyanov, A.; Lomayev, A.; Khoryaev, A.; Davydov, A.; Ssorin, V. Characteristics of indoor millimeter-wave channel at 60 GHz in application to perspective WLAN system. In Proceedings of the Fourth European Conference on Antennas and Propagation (EuCAP), Barcelona, Spain, 12–16 April 2010; IEEE: Barcelona, Spain, 2010; pp. 1–5.
38. Alkhateeb, A.; Mo, J.; González-Prelcic, N.; Heath, R. MIMO precoding and combining solutions for millimeter-wave systems. *IEEE Commun. Mag.* **2014**, *52*, 122–131.
39. Hur, S.; Kim, T.; Love, D.J.; Krogmeier, J.V.; Thomas, T.A.; Ghosh, A. Millimeter wave beamforming for wireless backhaul and access in small cell networks. *IEEE Trans. Commun.* **2013**, *61*, 4391–4403.
40. Xiao, Z.; Xia, P.; Xia, X.G. Enabling UAV cellular with millimeter-wave communication: Potentials and approaches. *IEEE Commun. Mag.* **2016**, *54*, 66–73.
41. Rangan, S.; Rappaport, T.S.; Erkip, E. Millimeter-wave cellular wireless networks: Potentials and challenges. *Proc. IEEE* **2014**, *102*, 366–385.
42. El Ayach, O.; Rajagopal, S.; Abu-Surra, S.; Pi, Z.; Heath, R. Spatially Sparse Precoding in Millimeter Wave MIMO Systems. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 1499–1513.
43. Fenn, A.J. Evaluation of adaptive phased array antenna, far-field nulling performance in the near-field region. *IEEE Trans. Antennas Propag.* **1990**, *38*, 173–185.
44. Yaghjian, A.D. An overview of near-field antenna measurements. *IEEE Trans. Antennas Propag.* **1986**, *34*, 30–45.
45. Sayers, A.; Dorsey, W.; O’Haver, K.; Valenzi, J. Planar near-field measurement of digital phased arrays using near-field scan plane reconstruction. *IEEE Trans. Antennas Propag.* **2012**, *60*, 2711–2718.
46. Liao, Y.P.; Han, H.; Guo, Q. Design of robust near-field multi-beam forming based on improved LCMV algorithm. *J. Inf. Hiding Multimed. Signal Process.* **2015**, *6*, 783–791.
47. Lin, J.R.; Peng, Q.C.; Shao, H.Z. Near-field robust adaptive beamforming based on worst-case performance optimization. *Int. J. Electr. Comput. Eng.* **2006**, *1*, 712–718.
48. Lee, M.S.; Park, J.Y.; Katkovnik, V.; Itoh, T.; Kim, Y.H. Adaptive robust DOA estimation for a 60-GHz antenna-array system. *IEEE Trans. Veh. Technol.* **2007**, *56*, 3231–3237.
49. Tsang, Y.M.; Poon, A.S.; Addepalli, S. Coding the beams: Improving beamforming training in mmwave communication system. In Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM 2011), Houston, TX, USA, 5–9 December 2011; IEEE: Houston, TX, USA, 2011; pp. 1–6.

50. Xiao, Z.; Xia, X.G.; Jin, D.; Ge, N. Iterative Eigenvalue Decomposition and Multipath-Grouping Tx/Rx Joint Beamformings for Millimeter-Wave Communications. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 1595–1607.
51. Alkhateeb, A.; El Ayach, O.; Leus, G.; Heath, R. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 831–846.
52. Peng, Y.; Li, Y.; Wang, P. An Enhanced Channel Estimation Method for Millimeter Wave Systems with Massive Antenna Arrays. *IEEE Commun. Lett.* **2015**, *19*, 1592–1595.
53. Vía, J.; Santamaría, I.; Elvira, V.; Eickhoff, R. A general criterion for analog Tx-Rx beamforming under OFDM transmissions. *IEEE Trans. Signal Process.* **2010**, *58*, 2155–2167.
54. Kokshoorn, M.; Wang, P.; Li, Y.; Vucetic, B. Fast channel estimation for millimetre wave wireless systems using overlapped beam patterns. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; IEEE: London, UK, 2015; pp. 1304–1309.
55. Kang, D.W.; Kim, J.G.; Min, B.W.; Rebeiz, G.M. Single and four-element-Ka-band transmit/receive phased-array silicon RFICs with 5-bit amplitude and phase control. *IEEE Trans. Microw. Theory Tech.* **2009**, *57*, 3534–3543.
56. Li, Q.; Hong, M.; Wai, H.T.; Liu, Y.F.; Ma, W.K.; Luo, Z.Q. Transmit Solutions for MIMO Wiretap Channels using Alternating Optimization. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 1714–1727.
57. Li, W.; Lilleberg, J.; Rikkinen, K. On rate region analysis of half-and full-duplex OFDM communication links. *IEEE J. Sel. Areas Commun.* **2014**, *32*, 1688–1698.



© 2016 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 (<http://creativecommons.org/licenses/by/4.0/>).