Full-Duplex mmWave Communication with Hybrid Precoding and Combining

Roberto López-Valcarce, Marcos Martínez-Cotelo
atlanTTic Research Center, University of Vigo, Spain
{valcarce,mmcotelo}@gts.uvigo.es

***Abstract*—We investigate the design of hybrid precoders and combiners for a millimeter wave (mmWave) point-to-point bidirectional link in which both nodes transmit and receive simultaneously and on the same carrier frequency. In such full-duplex configuration, mitigation of self-interference (SI) becomes critical. Large antenna arrays provide an opportunity for spatial SI suppression in mmWave. We assume a phase-shifter based, fully connected architecture for the analog part of the precoder and combiner. The proposed design, which aims at cancelling SI in the analog domain to avoid frontend saturation, significantly improves on the performance of previous approaches.**

***Index Terms*—Millimeter wave communication, full-duplex, hybrid precoding and combining.**

I. Introduction

Millimeter wave (mmWave) multiple-input multiple-output (MIMO) communication is envisioned as a key technology for future high data rate systems [1], [2]. The large propagation losses experienced at such frequencies need to be compensated by directional transmission using large antenna arrays. The traditional approach at lower frequencies, in which precoding and combining are implemented in the digital domain at baseband, is not suitable for large arrays since it requires a dedicated radio frequency (RF) chain per antenna, resulting in high cost and power consumption. Instead, hybrid solutions are preferred, in which the processing is split between the digital baseband and analog RF domains, so that the number of RF chains can be significantly reduced [3], [4].

Recently, there has been interest in the application of full-duplex (FD) to mmWave communication. In contrast with traditional half-duplex (HD) approaches, in which time-bandwidth resources are split to ensure transmit-receive (TX-RX) orthogonality, with FD these take place simultaneously and in the same frequency [5]–[7]. Besides the potential to double spectral efficiency, FD may add flexibility to point-to-point handshaking and multiple access methods [8]. However, it also results in *self-interference* (SI), since an FD node's transmission will leak to its own receiver and will overwhelm the much weaker signal of interest from a remote node. Promising results in SI mitigation have been reported for a single-antenna FD node operating in the sub-6 GHz bands [9], [10], by properly combining propagation domain, analog-circuit domain and digital domain methods. The first are passive and hinge on antenna design/placement; the second are needed to avoid saturation of the receiver RF chain and A/D converter, whereas the third estimate and substract the residual SI signal in the baseband. Unfortunately, analog-circuit domain methods do not scale well with the number of antennas, so their extension to MIMO FD is difficult [7]; on the other hand, the availability of multiple antennas can be exploited in MIMO FD to mitigate SI [11]–[13]. In microwave-band systems, the downside of such spatial suppression approach is a spectral efficiency loss, because some of the available spatial degrees of freedom (DoF) are spent in mitigating SI. In contrast, the large arrays used in mmWave result in a much larger available DoF, so that spatial suppression becomes particularly attractive.

In this context, a design was proposed in [14] for the digital beamforming weights of an FD mmWave single-stream bidirectional link. Although it provided close to optimal performance, its applicability was limited by the need to incorporate one RF chain per antenna [6]. This issue was overcome in [15], where a suitable modification of the design from [14] was introduced to allow for analog beamforming weights implemented via phase shifters. The resulting analog beamforming design effectively cancels SI with a modest spectral efficiency loss with respect to the all-digital design from [14]. We present an extension of the single-stream analog design from [15] to the multistream case, using hybrid precoders and combiners. The multistream scenario has also been recently addressed in [16], which proposed a two-step scheme in which a set of all-digital precoders and combiners is designed first, and then a hybrid decomposition is obtained using least-squares (LS) approximation to the all-digital solution. This approach has two significant drawbacks: first, the method from [16] attempts to suppress SI at the baseband, whereas in practice it must be placed in the analog domain to avoid saturation of the analog frontend; second, such hybrid LS approximations need not cancel SI accurately [6], so that their performance degrades with high SI levels. These two issues are explicitly taken into account in our proposed design.

II. Signal Model

Consider a network consisting of 2 FD nodes, as shown in Fig. 1. Node $i \in \{1,2\}$ is equipped with $N_{t,i}$ transmit antennas, and supports the transmission of $N_{s,i}$ data streams towards node $j \in \{1,2\}$, $j \neq i$, which is equipped with $N_{r,j}$ receive antennas. The data vectors $\boldsymbol{s}_i \in \mathbb{C}^{N_{s,i}}$ are assumed

Work supported in part by the Agencia Estatal de Investigación (Spain), in part by the European Regional Development Fund (ERDF) through projects WINTER (TEC2016-76409-C2-2-R) and RODIN (PID2019-105717RB-C21), and in part by the Xunta de Galicia (atlanTTic accreditation 2020-2023).

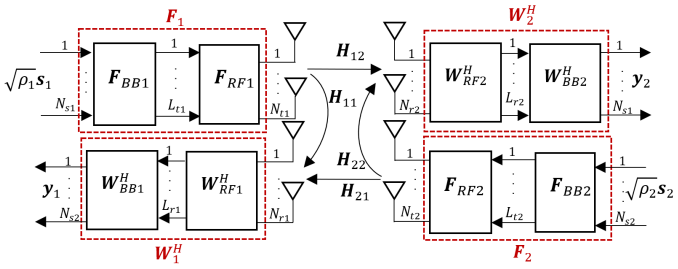

Fig. 1. Two-node MIMO FD network with hybrid precoding and combining.

zero-mean with covariance $\boldsymbol{I}_{N_{s,i}}$, $i \in \{1,2\}$. Channels are assumed approximately frequency-flat. The $N_{r,j} \times N_{t,i}$ channel matrix from the TX array of node i to the RX array of node j is denoted as $\boldsymbol{H}_{ij}$. In particular, $\boldsymbol{H}_{jj}$ corresponds to the SI channel affecting node j. Our design is model-independent, so a discussion on specific channel models is deferred to the description of simulation results in Sec. V.

The number of RF chains at the TX and RX frontends of node i are $L_{t,i}$ and $L_{r,i}$ respectively. For an all-digital system, $L_{t,i} = N_{t,i}$ and $L_{r,i} = N_{r,i}$, $i \in \{1,2\}$, so that a dedicated RF chain per antenna is available. However, to reduce cost and power consumption, hybrid architectures with a smaller number of RF chains are preferred. Thus, in general $N_{s,i} \leq L_{t,i} \leq N_{t,i}$ and $N_{s,i} \leq L_{r,j} \leq N_{r,j}$, $i,j \in \{1,2\}$, $i \neq j$.

The transmitter of node i applies a precoder $\boldsymbol{F}_i = \boldsymbol{F}_{\mathrm{RF},i}\boldsymbol{F}_{\mathrm{BB},i}$ to the transmitted data vector $\boldsymbol{s}_i$, where $\boldsymbol{F}_{\mathrm{BB},i} \in \mathbb{C}^{L_{t,i} \times N_{s,i}}$ is the baseband digital precoder, and $\boldsymbol{F}_{\mathrm{RF},i} \in \mathbb{V}^{N_{t,i} \times L_{t,i}}$ is the RF analog precoder. Here, $\mathbb{V}^{N \times L} \subset \mathbb{C}^{N \times L}$ denotes the feasible set for the RF matrices as dictated by hardware (HW) constraints; in this paper we assume that $\mathbb{V}^{N \times L}$ is the set of $N \times L$ matrices such that their entries have all the same magnitude, as corresponds to a phase-shifter based implementation. Similarly, the receiver of node j applies a combiner $\boldsymbol{W}_j = \boldsymbol{W}_{\mathrm{RF},j}\boldsymbol{W}_{\mathrm{BB},j}$ to the received vector, where $\boldsymbol{W}_{\mathrm{BB},j} \in \mathbb{C}^{L_{r,j} \times N_{s,i}}$ is the baseband combiner, and $\boldsymbol{W}_{\mathrm{RF},j} \in \mathbb{V}^{N_{r,j} \times L_{r,j}}$ is the RF combiner. In this way, the vector at the output of node j's combiner is given by

$$\boldsymbol{y}_j = \sqrt{\rho_i}\boldsymbol{W}_j^H\boldsymbol{H}_{ij}\boldsymbol{F}_i\boldsymbol{s}_i + \underbrace{\sqrt{\eta_j}\boldsymbol{W}_j^H\boldsymbol{H}_{jj}\boldsymbol{F}_j\boldsymbol{z}_j + \boldsymbol{W}_j^H\boldsymbol{n}_j}_{\text{SI + noise}}, \quad (1)$$

where ρ_i is the transmit power of node i, η_j is the SI power at node j, $\boldsymbol{n}_j$ is zero-mean white Gaussian noise with covariance $\sigma_j^2\boldsymbol{I}_{N_{r,j}}$, and $\boldsymbol{z}_j$ represents the SI at node j, which is to be understood as residual SI if any propagation and/or analog-circuit domain SI mitigation is applied. In general, $\boldsymbol{z}_j$ will be a distorted version of the signal transmitted by node j, and is modeled as a zero-mean vector with covariance $\boldsymbol{I}_{N_{s,j}}$. Hence, the SI + noise covariance matrix at node j is given by

$$\boldsymbol{R}_j = \eta_j\boldsymbol{W}_j^H\boldsymbol{H}_{jj}\boldsymbol{F}_j\boldsymbol{F}_j^H\boldsymbol{H}_{jj}^H\boldsymbol{W}_j + \sigma_j^2\boldsymbol{W}_j^H\boldsymbol{W}_j. \quad (2)$$

The effective channel from node i to node j is given by $\widetilde{\boldsymbol{H}}_{ij} = \boldsymbol{W}_j^H\boldsymbol{H}_{ij}\boldsymbol{F}_i$, so that, treating SI as noise, the spectral efficiency of the $i \to j$ link is given by

$$\mathcal{R}_{ij} = \log_2\left|\boldsymbol{I}_{N_{s,i}} + \rho_i\widetilde{\boldsymbol{H}}_{ij}^H\boldsymbol{R}_j^{-1}\widetilde{\boldsymbol{H}}_{ij}\right|. \quad (3)$$

Using the SVD of $\boldsymbol{W}_j$, it is seen that $\mathcal{R}_{ij}$ depends on $\boldsymbol{W}_j$ only through its left singular vectors, so that $\boldsymbol{W}_j$ can be assumed semi-unitary ($\boldsymbol{W}_j^H\boldsymbol{W}_j = \boldsymbol{I}_{N_{s,i}}$) w.l.o.g. Then $\mathcal{R}_{ij}$ can be written as in (4) at the top of next page, where $\epsilon_{ij} \triangleq \frac{\rho_i}{\sigma_j^2}$ and $\epsilon_{jj} \triangleq \frac{\eta_j}{\sigma_j^2}$ are respectively the SNR and INR at the receiver of node j. Similarly, $\mathcal{R}_{ij}$ is independent of the right singular vectors of $\boldsymbol{F}_i$; in the sequel, we assume uniform power allocation across streams, so that $\boldsymbol{F}_i$ is semi-unitary as well.

The goal is to maximize the overall spectral efficiency, given by $\mathcal{R} = \mathcal{R}_{12} + \mathcal{R}_{21}$. For benchmarking purposes, it is useful to have a performance upper bound, which can be obtained by assuming no SI ($\epsilon_{11} = \epsilon_{22} = 0$) and no HW constraints (so that $\{\boldsymbol{F}_i, \boldsymbol{W}_i\}_{i=1}^2$ are only constrained to be semi-unitary). Then the optimum precoders and combiners are respectively given by the dominant right and left singular vectors of the channel matrices, yielding $\mathcal{R} \leq \mathcal{R}^\star$ where

$$\mathcal{R}^\star \triangleq \sum_{k=1}^{N_{s,1}} \log_2(1+\epsilon_{12}\sigma_k^2(\boldsymbol{H}_{12})) + \sum_{\ell=1}^{N_{s,2}} \log_2(1+\epsilon_{21}\sigma_\ell^2(\boldsymbol{H}_{21})), \quad (5)$$

with $\{\sigma_k(\boldsymbol{H})\}$ the singular values of $\boldsymbol{H}$, in descending order.

III. All-Digital Design

In this section we focus on an all-digital design for the precoders and combiners, which will serve as starting point for the hybrid design in Sec. IV as well as a reference for its performance. Note that even in this case, i.e., with no constraints on $\{\boldsymbol{F}_i, \boldsymbol{W}_i\}_{i=1}^2$ beyond being semi-unitary, maximizing $\mathcal{R}$ is a non-convex problem for which no closed-form solution is known. Following similar steps to those in [14] for the single-stream case, we adopt a suboptimal approach in which a zero-forcing (ZF) constraint is imposed on the SI at both nodes, resulting in the following problem:

$$\max_{\{\boldsymbol{F}_i,\boldsymbol{W}_i\}_{i=1}^2} \log_2\left|\boldsymbol{I}_{N_{s,1}} + \epsilon_{12}\boldsymbol{F}_1^H\boldsymbol{H}_{12}^H\boldsymbol{W}_2\boldsymbol{W}_2^H\boldsymbol{H}_{12}\boldsymbol{F}_1\right| + \log_2\left|\boldsymbol{I}_{N_{s,2}} + \epsilon_{21}\boldsymbol{F}_2^H\boldsymbol{H}_{21}^H\boldsymbol{W}_1\boldsymbol{W}_1^H\boldsymbol{H}_{21}\boldsymbol{F}_2\right| \quad (6)$$

$$\text{s.t.} \begin{cases} \boldsymbol{W}_j^H\boldsymbol{W}_j &= \boldsymbol{I}_{N_{s,i}} \\ \boldsymbol{F}_i^H\boldsymbol{F}_i &= \boldsymbol{I}_{N_{s,i}} \\ \boldsymbol{W}_i^H\boldsymbol{H}_{ii}\boldsymbol{F}_i &= \boldsymbol{0} \end{cases} \quad i,j \in \{1,2\},\ i \neq j.$$

Problem (6) is difficult due to the coupling between variables introduced by the ZF constraints $\boldsymbol{W}_i^H\boldsymbol{H}_{ii}\boldsymbol{F}_i = \boldsymbol{0}$. However, if $\boldsymbol{F}_i$ is held fixed, then it is possible to maximize the objective w.r.t. $\boldsymbol{W}_j$, $j \neq i$, in closed form, and vice versa. This suggests the following cyclic maximization procedure to tackle (6):

- Given $\boldsymbol{F}_1$, $\boldsymbol{F}_2$, for $j \in \{1,2\}$ and $i \neq j$ solve

$$\max_{\boldsymbol{W}_j} \log_2\left|\boldsymbol{I}_{N_{s,i}} + \epsilon_{ij}\boldsymbol{W}_j^H\boldsymbol{H}_{ij}\boldsymbol{F}_i\boldsymbol{F}_i^H\boldsymbol{H}_{ij}^H\boldsymbol{W}_j\right| \quad (7)$$
$$\text{s.t.} \quad \boldsymbol{W}_j^H\boldsymbol{W}_j = \boldsymbol{I}_{N_{s,i}} \quad \text{and} \quad \boldsymbol{W}_j^H\boldsymbol{H}_{jj}\boldsymbol{F}_j = \boldsymbol{0}.$$

- Given $\boldsymbol{W}_1$, $\boldsymbol{W}_2$, for $i \in \{1,2\}$ and $j \neq i$ solve

$$\max_{\boldsymbol{F}_i} \log_2\left|\boldsymbol{I}_{N_{s,i}} + \epsilon_{ij}\boldsymbol{F}_i^H\boldsymbol{H}_{ij}^H\boldsymbol{W}_j\boldsymbol{W}_j^H\boldsymbol{H}_{ij}\boldsymbol{F}_i\right| \quad (8)$$
$$\text{s.t.} \quad \boldsymbol{F}_i^H\boldsymbol{F}_i = \boldsymbol{I}_{N_{s,i}} \quad \text{and} \quad \boldsymbol{F}_i^H\boldsymbol{H}_{ii}^H\boldsymbol{W}_i = \boldsymbol{0}.$$

$$\mathcal{R}_{ij} = \log_2 \left| \boldsymbol{I}_{N_{s,i}} + \epsilon_{ij} \boldsymbol{F}_i^H \boldsymbol{H}_{ij}^H \boldsymbol{W}_j (\boldsymbol{I}_{N_{s,i}} + \epsilon_{jj} \boldsymbol{W}_j^H \boldsymbol{H}_{jj} \boldsymbol{F}_j \boldsymbol{F}_j^H \boldsymbol{H}_{jj}^H \boldsymbol{W}_j)^{-1} \boldsymbol{W}_j^H \boldsymbol{H}_{ij} \boldsymbol{F}_i \right| \tag{4}$$

These steps are then iterated until convergence. For initialization, $\boldsymbol{F}_1$, $\boldsymbol{F}_2$ can be set to the dominant right singular vectors of $\boldsymbol{H}_{12}$ and $\boldsymbol{H}_{21}$, respectively. Note that the problems in (7) and (8) share the following generic structure:

$$\max_{\boldsymbol{X}} \log_2 \left| \boldsymbol{I}_N + \epsilon \boldsymbol{X}^H \boldsymbol{A} \boldsymbol{A}^H \boldsymbol{X} \right| \text{ s.t. } \begin{cases} \boldsymbol{X}^H \boldsymbol{X} &= \boldsymbol{I}_N, \\ \boldsymbol{X}^H \boldsymbol{C} &= \boldsymbol{0}, \end{cases} \tag{9}$$

with $\boldsymbol{X} \in \mathbb{C}^{M \times N}$ the optimization variable, and $\boldsymbol{A} \in \mathbb{C}^{M \times N}$, $\boldsymbol{C} \in \mathbb{C}^{M \times P}$ given matrices. Assuming that $M \geq N + P$ (meaning that the corresponding number of antennas M is sufficiently large to sustain the transmission of N streams after spending P degrees of freedom to meet the ZF constraint), the solution to (9) is given by the N dominant left singular vectors of $\boldsymbol{P}_\perp \boldsymbol{A}$, where $\boldsymbol{P}_\perp$ is the projection matrix onto the subspace orthogonal to the columns of $\boldsymbol{C}$ (see Appendix A for the proof). Thus, steps (7)-(8) can be solved in closed form.

IV. Hybrid design

The precoding and combining matrices must be decomposed now into their baseband and RF factors, with the latter restricted to the set of feasible values as per HW constraints. We adopt a cyclic maximization procedure analogous to that from Sec. III, after incorporating these additional constraints:

- Given $\boldsymbol{F}_1$, $\boldsymbol{F}_2$, for $j \in \{1, 2\}$ and $i \neq j$ solve

$$\begin{aligned} \max_{\boldsymbol{W}_{\mathrm{BB},j}, \boldsymbol{W}_{\mathrm{RF},j}} \; & \log_2 \left| \boldsymbol{I}_{N_{s,i}} + \epsilon_{ij} \boldsymbol{W}_j^H \boldsymbol{H}_{ij} \boldsymbol{F}_i \boldsymbol{F}_i^H \boldsymbol{H}_{ij}^H \boldsymbol{W}_j \right| \\ \text{s.t. } & \boldsymbol{W}_j = \boldsymbol{W}_{\mathrm{RF},j} \boldsymbol{W}_{\mathrm{BB},j}, \quad \boldsymbol{W}_{\mathrm{RF},j} \in \mathbb{V}^{N_{r,j} \times L_{r,j}}, \\ & \boldsymbol{W}_j^H \boldsymbol{W}_j = \boldsymbol{I}_{N_{s,i}}, \qquad \boldsymbol{W}_{\mathrm{RF},j}^H \boldsymbol{H}_{jj} \boldsymbol{F}_{\mathrm{RF},j} = \boldsymbol{0}. \end{aligned} \tag{10}$$

- Given $\boldsymbol{W}_1$, $\boldsymbol{W}_2$, for $i \in \{1, 2\}$ and $j \neq i$ solve

$$\begin{aligned} \max_{\boldsymbol{F}_{\mathrm{BB},i}, \boldsymbol{F}_{\mathrm{RF},i}} \; & \log_2 \left| \boldsymbol{I}_{N_{s,i}} + \epsilon_{ij} \boldsymbol{F}_i^H \boldsymbol{H}_{ij}^H \boldsymbol{W}_j \boldsymbol{W}_j^H \boldsymbol{H}_{ij} \boldsymbol{F}_i \right| \\ \text{s.t. } & \boldsymbol{F}_i = \boldsymbol{F}_{\mathrm{RF},i} \boldsymbol{F}_{\mathrm{BB},i}, \quad \boldsymbol{F}_{\mathrm{RF},i} \in \mathbb{V}^{N_{t,i} \times L_{t,i}}, \\ & \boldsymbol{F}_i^H \boldsymbol{F}_i = \boldsymbol{I}_{N_{s,i}}, \qquad \boldsymbol{F}_{\mathrm{RF},i}^H \boldsymbol{H}_{ii}^H \boldsymbol{W}_{\mathrm{RF},i} = \boldsymbol{0}. \end{aligned} \tag{11}$$

Note that in (10)-(11) the ZF constraints only involve the RF precoders and combiners, and not the baseband ones. The reason for this is twofold: first, it guarantees that the SI is cancelled before downconversion, sampling, and quantization take place in the receivers, so that saturation of the corresponding RF chains due to large SI levels is avoided. Second, it adds robustness to any transmitter noise that may be generated in the upconversion/amplification steps and therefore not present in the baseband, since the SI generated by such transmitter noise would not be affected by the baseband precoder.

It is seen that problems (10)-(11) share a common structure:

$$\begin{aligned} \max_{\boldsymbol{X}_{\mathrm{BB}}, \boldsymbol{X}_{\mathrm{RF}}} \; & \log_2 \left| \boldsymbol{I}_N + \epsilon \boldsymbol{X}_{\mathrm{BB}}^H \boldsymbol{X}_{\mathrm{RF}}^H \boldsymbol{A} \boldsymbol{A}^H \boldsymbol{X}_{\mathrm{RF}} \boldsymbol{X}_{\mathrm{BB}} \right| \\ \text{s.t. } & \begin{cases} \boldsymbol{X}_{\mathrm{BB}}^H \boldsymbol{X}_{\mathrm{RF}}^H \boldsymbol{X}_{\mathrm{RF}} \boldsymbol{X}_{\mathrm{BB}} = \boldsymbol{I}_N, \\ \boldsymbol{X}_{\mathrm{RF}}^H \boldsymbol{C} = \boldsymbol{0}, \\ \boldsymbol{X}_{\mathrm{RF}} \in \mathbb{V}^{M \times L}, \end{cases} \end{aligned} \tag{12}$$

where $\boldsymbol{X}_{\mathrm{RF}} \in \mathbb{C}^{M \times L}$, $\boldsymbol{X}_{\mathrm{BB}} \in \mathbb{C}^{L \times N}$ are the optimization variables, and $\boldsymbol{A} \in \mathbb{C}^{M \times N}$, $\boldsymbol{C} \in \mathbb{C}^{M \times P}$ are given matrices.

The optimal baseband factor $\boldsymbol{X}_{\mathrm{BB}}$ can be found in terms of the RF factor $\boldsymbol{X}_{\mathrm{RF}}$ as follows. First, consider the SVD $\boldsymbol{X}_{\mathrm{RF}} = \boldsymbol{U}_{\mathrm{RF}} \boldsymbol{S}_{\mathrm{RF}} \boldsymbol{V}_{\mathrm{RF}}^H$. Then $\boldsymbol{X}_{\mathrm{BB}} = \boldsymbol{V}_{\mathrm{RF}} \boldsymbol{S}_{\mathrm{RF}}^{-1} \boldsymbol{Q}_\star$, where the columns of $\boldsymbol{Q}_\star \in \mathbb{C}^{M \times N}$ comprise the N dominant left singular vectors of $\boldsymbol{U}_{\mathrm{RF}}^H \boldsymbol{A}$ (see Appendix B for the proof).

On the other hand, no analytical solution exists for the RF factor $\boldsymbol{X}_{\mathrm{RF}}$. Inaccuracies in SI cancellation have a large impact on the spectral efficiency, and thus we propose to find a feasible RF factor which simultaneously satisfies the ZF condition $\boldsymbol{X}_{\mathrm{RF}}^H \boldsymbol{C} = \boldsymbol{0}$ and the HW-related constraints $\boldsymbol{X}_{\mathrm{RF}} \in \mathbb{V}^{M \times L}$ by applying the method of alternating projections [17], which has been shown in [15] to provide good results in the single-stream case. Specifically, at iteration k, the previous estimate $\boldsymbol{X}_{\mathrm{RF}}^{(k-1)}$ is projected onto the subspace orthogonal to the columns of $\boldsymbol{C}$, and then the result is projected onto $\mathbb{V}^{M \times L}$ to obtain $\boldsymbol{X}_{\mathrm{RF}}^{(k)}$. This is repeated multiple times until convergence is achieved.

To increase the likelihood that the RF factor obtained in this way yields a large value of the objective in (12), proper initialization of the alternate projections method is crucial. We propose to set $\boldsymbol{X}_{\mathrm{RF}}^{(0)}$ to the maximizer of the objective in (12) neglecting the ZF and HW-related constraints, after the baseband factor $\boldsymbol{X}_{\mathrm{BB}}$ has been optimized. Then, $\boldsymbol{X}_{\mathrm{RF}}^{(0)}$ is given by the L dominant left singular vectors of $\boldsymbol{A}$ (see Appendix B). The overall hybrid design is summarized in Algorithm 1.

V. Results

Consider a setting with both nodes equipped with 32-antenna, $\frac{\lambda}{2}$-spaced uniform linear arrays (ULAs), with 4 TX and 4 RX RF chains, and with 4 streams to be transmitted in each direction. For the $1 \to 2$ and $2 \to 1$ channels, the Saleh-Valenzuela narrowband clustered model from [3] is assumed, with N_{cl} scattering clusters and N_{ray} rays per cluster:

$$\mathbf{H}_{ij} = \sum_{n=1}^{N_{\mathrm{cl}}} \sum_{m=1}^{N_{\mathrm{ray}}} \alpha_{ij}^{m,n} \mathbf{a}_r \left(\phi_{ij}^{m,n} \right) \mathbf{a}_t^T \left(\theta_{ij}^{m,n} \right) \tag{13}$$

where for the m-th ray in the n-th cluster, $\mathbf{a}_t$ and $\mathbf{a}_r$ are the antenna array steering vectors at the transmitter and receiver, respectively, evaluated at the corresponding azimuth angles of departure from transmitter, $\theta_{ij}^{m,n}$, or arrival at receiver, $\phi_{ij}^{m,n}$; and $\alpha_{ij}^{m,n}$ is the complex gain. The SI channels have a near-field line-of-sight (LOS) component, as well as a far-field component due to SI reflections in nearby scatterers:

$$\mathbf{H}_{ii} = \sqrt{\frac{\kappa}{\kappa + 1}} \mathbf{H}_{\mathrm{LOS}}^{(i)} + \sqrt{\frac{1}{\kappa + 1}} \mathbf{H}_{\mathrm{REF}}^{(i)} \tag{14}$$

Algorithm 1 Hybrid Full-Duplex precoder-combiner design

function ALTPROJ($\boldsymbol{A}$, $\boldsymbol{C}$, L)
  $\boldsymbol{X}_{\mathrm{RF}} \leftarrow L$ dominant left singular vectors of $\boldsymbol{A}$
  $\boldsymbol{P}_{\perp} \leftarrow \boldsymbol{I} - \boldsymbol{C}\boldsymbol{C}^{\dagger}$
  for $k \leftarrow 1, N_{\mathrm{inner}}$ **do**
    $\boldsymbol{Y} \leftarrow \boldsymbol{P}_{\perp}\boldsymbol{X}_{\mathrm{RF}}$
    for $i \leftarrow 1, M$ and $j \leftarrow 1, L$ **do**
      $(\boldsymbol{X}_{\mathrm{RF}})_{ij} \leftarrow \frac{(\boldsymbol{Y})_{ij}}{|(\boldsymbol{Y})_{ij}|}$
    end for
  end for
  return $\boldsymbol{X}_{\mathrm{RF}}$
end function

function BASEBANDFACTOR($\boldsymbol{X}_{\mathrm{RF}}$, $\boldsymbol{A}$, N)
  Compute SVD $\boldsymbol{X}_{\mathrm{RF}} = \boldsymbol{U}_{\mathrm{RF}}\boldsymbol{S}_{\mathrm{RF}}\boldsymbol{V}_{\mathrm{RF}}^{H}$
  $\boldsymbol{Q} \leftarrow N$ dominant left singular vectors of $\boldsymbol{U}_{\mathrm{RF}}^{H}\boldsymbol{A}$
  $\boldsymbol{X}_{\mathrm{BB}} \leftarrow \boldsymbol{V}_{\mathrm{RF}}\boldsymbol{S}_{\mathrm{RF}}^{-1}\boldsymbol{Q}$
  return $\boldsymbol{X}_{\mathrm{BB}}$
end function

Input: $\boldsymbol{H}_{12}$, $\boldsymbol{H}_{21}$, $\boldsymbol{H}_{11}$, $\boldsymbol{H}_{22}$
Initialize $\boldsymbol{F}_{\mathrm{BB},i}$, $\boldsymbol{F}_{\mathrm{RF},i}$; set $\boldsymbol{F}_i = \boldsymbol{F}_{\mathrm{RF},i}\boldsymbol{F}_{\mathrm{BB},i}$, $i \in \{1,2\}$
for $t \leftarrow 1, N_{\mathrm{outer}}$ **do**
  $\boldsymbol{W}_{\mathrm{RF},1} \leftarrow$ ALTPROJ($\boldsymbol{H}_{21}\boldsymbol{F}_2, \boldsymbol{H}_{11}\boldsymbol{F}_{\mathrm{RF},1}, L_{r,1}$)
  $\boldsymbol{W}_{\mathrm{BB},1} \leftarrow$ BASEBANDFACTOR($\boldsymbol{W}_{\mathrm{RF},1}, \boldsymbol{H}_{21}\boldsymbol{F}_2, N_{s,2}$)
  $\boldsymbol{W}_{\mathrm{RF},2} \leftarrow$ ALTPROJ($\boldsymbol{H}_{12}\boldsymbol{F}_1, \boldsymbol{H}_{22}\boldsymbol{F}_{\mathrm{RF},2}, L_{r,2}$)
  $\boldsymbol{W}_{\mathrm{BB},2} \leftarrow$ BASEBANDFACTOR($\boldsymbol{W}_{\mathrm{RF},2}, \boldsymbol{H}_{12}\boldsymbol{F}_1, N_{s,1}$)
  Set $\boldsymbol{W}_j = \boldsymbol{W}_{\mathrm{RF},j}\boldsymbol{W}_{\mathrm{BB},j}$, $j \in \{1,2\}$
  $\boldsymbol{F}_{\mathrm{RF},1} \leftarrow$ ALTPROJ($\boldsymbol{H}_{12}^{H}\boldsymbol{W}_2, \boldsymbol{H}_{11}^{H}\boldsymbol{W}_{\mathrm{RF},1}, L_{t,1}$)
  $\boldsymbol{F}_{\mathrm{BB},1} \leftarrow$ BASEBANDFACTOR($\boldsymbol{F}_{\mathrm{RF},1}, \boldsymbol{H}_{12}^{H}\boldsymbol{W}_2, N_{s,1}$)
  $\boldsymbol{F}_{\mathrm{RF},2} \leftarrow$ ALTPROJ($\boldsymbol{H}_{21}^{H}\boldsymbol{W}_1, \boldsymbol{H}_{22}^{H}\boldsymbol{W}_{\mathrm{RF},2}, L_{t,2}$)
  $\boldsymbol{F}_{\mathrm{BB},2} \leftarrow$ BASEBANDFACTOR($\boldsymbol{F}_{\mathrm{RF},2}, \boldsymbol{H}_{21}^{H}\boldsymbol{W}_1, N_{s,2}$)
  Set $\boldsymbol{F}_i = \boldsymbol{F}_{\mathrm{RF},i}\boldsymbol{F}_{\mathrm{BB},i}$, $i \in \{1,2\}$
end for

with κ the Rice factor. For the far-field term $\mathbf{H}_{\mathrm{REF}}$, the same model as in (13) is adopted, whereas the LOS component follows the near-field model [6], [14], [16]

$$\left(\mathbf{H}_{\mathrm{LOS}}^{(i)}\right)_{mn} = \frac{1}{r_{mn}^{(i)}} \exp\left(-j2\pi \frac{r_{mn}^{(i)}}{\lambda}\right) \tag{15}$$

where $r_{mn}^{(i)}$ is the distance from the m-th TX antenna to the n-th RX antenna of the i-th node, and λ is the wavelength. For both nodes, the array geometry of [6, Fig. 2] is adopted, with a TX-RX array distance of $d = 2\lambda$ and angle $\omega = \frac{\pi}{2}$.

We assumed $N_{\mathrm{cl}} = 6$, $N_{\mathrm{ray}} = 10$ and $\kappa = 10$ dB. Departure/arrival angles are random, with mean cluster angle uniformly distributed in $[0, 360°]$ and angular spreads of $16°$. Path gains are i.i.d. complex circular Gaussian with the same variance. Channel matrices are normalized so that their squared Frobenius norms equal the number of their entries.

The spectral efficiency was computed by averaging over 300 channel realizations at each point. We compare the proposed two-step hybrid FD design (with randomly initialized precoders) with that from [16] under the same setting. For reference, a hybrid Half-Duplex design was included as well, for which the hybrid precoders and combiners are obtained by iteratively minimizing their Euclidean distance to the corresponding channel singular vectors while keeping the semi-unitary and HW-related constraints.

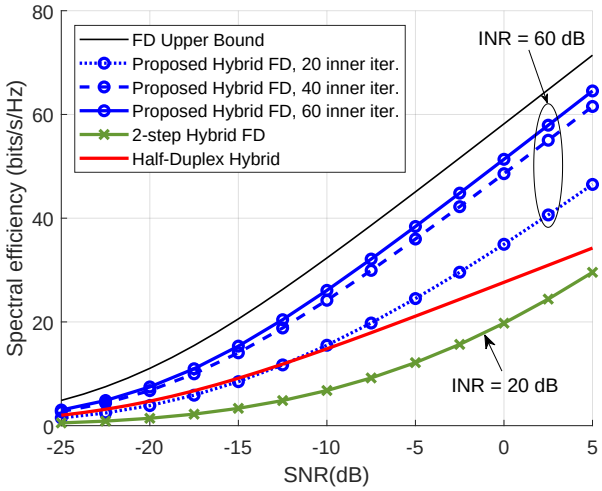

Fig. 2. Spectral efficiency vs. SNR .

Fig. 2 shows the spectral efficiency $\mathcal{R}$ as a function of the SNR $\epsilon_{12} = \epsilon_{21}$ assumed to be the same at both nodes, and for a given value of the INR $\epsilon_{11} = \epsilon_{22}$, also identical at both nodes. The performance of the proposed design under INR $= 60$ dB is seen to be sensitive to the number of "inner" alternating projection iterations (parameter N_{inner} in Algorithm 1; the number of "outer" iterations was set to $N_{\mathrm{outer}} = 20$ in all cases). With too few iterations, the final RF precoder/combiner is not yet in the null space of the corresponding matrix, so that the ZF constraint is only approximately satisfied. In the current scenario, 60 inner iterations are sufficient and performance does not improve beyond this point, getting close to the upper bound $\mathcal{R}_{\star}$ from (5). In contrast, the performance of the two-step approach from [16], even under a much lower INR value (20 dB), is seen to be significantly worse, dropping below the performance of the HD hybrid scheme. This is due to the fact that the hybrid design from [16] obtains the baseband and RF matrices by LS approximation to the corresponding all-digital matrix; the approximation errors that ensue translate into imperfect SI cancellation by the hybrid network, with a large impact in spectral efficiency. Additionally, note that in contrast with our method, the design from [16] does not attempt to cancel the SI in the analog domain; thus, it is likely to result in frontend saturation in a practical setting.

Fig. 3 shows the behavior of the spectral efficiency in terms of INR, for a fixed value of SNR $= 0$ dB. The FD all-digital design of Sec. III remains very close to the upper bound (5) and is insensitive to the INR level. With a sufficient number of inner iterations, the proposed hybrid FD design achieves more than 90% of the spectral efficiency of the all-digital design for a wide range of INR values. In contrast, the method from

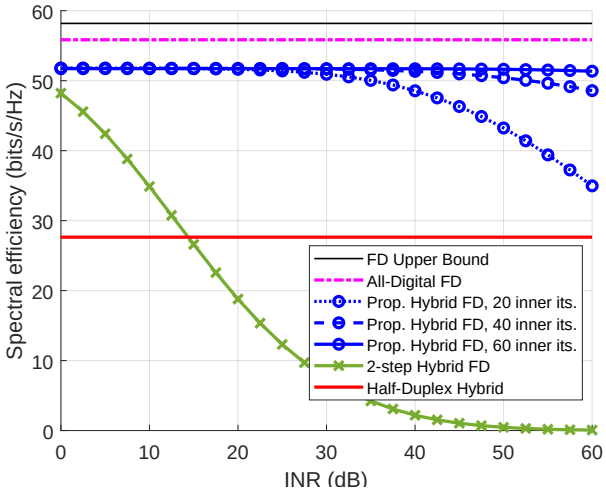

Fig. 3. Spectral efficiency vs. INR for SNR $= 0$ dB.

[16] is seen to degrade quickly as the INR increases, for the reasons mentioned above.

VI. Conclusion

An iterative joint design for the precoders and combiners of a hybrid FD architecture in a multistream bi-directional mmWave two-node network was proposed, with the SI cancelled in the analog domain. The key feature of the design is the specification of a zero-forcing constraint at every iteration of the alternating-projections procedure lying at the core of the method. This avoids inaccuracies in previous schemes which obtain the hybrid factorizations by directly approximating the all-digital matrices. Numerical results attest to the effectivity of the proposed design even with high SI levels.

Appendix

A. Solution to Problem (9)

Assume $\boldsymbol{C}$ has full column rank P (if not, it can be replaced in the constraint by another matrix $\boldsymbol{C}' \in \mathbb{C}^{M\times P'}$ with the same range space as $\boldsymbol{C}$ and of full rank $P' < P$). Let the columns of $\boldsymbol{U}_0 \in \mathbb{C}^{M\times(M-P)}$ constitute an orthonormal basis for the subspace orthogonal to $\boldsymbol{C}$. Then $\boldsymbol{P}_\perp = \boldsymbol{U}_0\boldsymbol{U}_0^H$; and $\boldsymbol{X}^H\boldsymbol{C} = \boldsymbol{0}$ implies that $\boldsymbol{X} = \boldsymbol{U}_0\boldsymbol{Y}$ for some $\boldsymbol{Y} \in \mathbb{C}^{(M-P)\times N}$. Further, $\boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_N$ and $M - P \geq N$ imply that $\boldsymbol{Y}^H\boldsymbol{Y} = \boldsymbol{I}_N$, since $\boldsymbol{U}_0^H\boldsymbol{U}_0 = \boldsymbol{I}_{M-P}$. The problem becomes

$$\max_{\boldsymbol{Y}^H\boldsymbol{Y}=\boldsymbol{I}_N} \log_2\left|\boldsymbol{I}_N + \epsilon\boldsymbol{Y}^H\boldsymbol{U}_0^H\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{U}_0\boldsymbol{Y}\right|, \qquad (16)$$

whose solution $\boldsymbol{Y}_\star$ is given by the N dominant left singular vectors of $\boldsymbol{U}_0^H\boldsymbol{A}$. Thus, $\boldsymbol{U}_0^H\boldsymbol{A}$ admits an SVD of the form $\boldsymbol{U}_0^H\boldsymbol{A} = \boldsymbol{Y}_\star\boldsymbol{S}\boldsymbol{V}^H$, so that $\boldsymbol{U}_0\boldsymbol{U}_0^H\boldsymbol{A} = \boldsymbol{U}_0\boldsymbol{Y}_\star\boldsymbol{S}\boldsymbol{V}^H = \boldsymbol{X}_\star\boldsymbol{S}\boldsymbol{V}^H$ constitutes an SVD of $\boldsymbol{U}_0\boldsymbol{U}_0^H\boldsymbol{A} = \boldsymbol{P}_\perp\boldsymbol{A}$. Therefore, the solution $\boldsymbol{X}_\star = \boldsymbol{U}_0\boldsymbol{Y}_\star$ to problem (9) is given by the N dominant left singular vectors of $\boldsymbol{P}_\perp\boldsymbol{A}$.

B. Optimal baseband factor in (12)

For $\boldsymbol{X}_{\mathrm{RF}}$ given, consider the SVD $\boldsymbol{X}_{\mathrm{RF}} = \boldsymbol{U}_{\mathrm{RF}}\boldsymbol{S}_{\mathrm{RF}}\boldsymbol{V}_{\mathrm{RF}}^H$, and let $\boldsymbol{Q} = \boldsymbol{S}_{\mathrm{RF}}\boldsymbol{V}_{\mathrm{RF}}^H\boldsymbol{X}_{\mathrm{BB}} \in \mathbb{C}^{L\times N}$. Then $\boldsymbol{X}_{\mathrm{RF}}\boldsymbol{X}_{\mathrm{BB}} = \boldsymbol{U}_{\mathrm{RF}}\boldsymbol{Q}$ so that $\boldsymbol{X}_{\mathrm{BB}}^H\boldsymbol{X}_{\mathrm{RF}}^H\boldsymbol{X}_{\mathrm{RF}}\boldsymbol{X}_{\mathrm{BB}} = \boldsymbol{Q}^H\boldsymbol{U}_{\mathrm{RF}}^H\boldsymbol{U}_{\mathrm{RF}}\boldsymbol{Q} = \boldsymbol{Q}^H\boldsymbol{Q}$. Then the problem can be cast in terms of $\boldsymbol{Q}$ as

$$\max_{\boldsymbol{Q}^H\boldsymbol{Q}=\boldsymbol{I}_N} \log_2\left|\boldsymbol{I}_N + \epsilon\boldsymbol{Q}^H\boldsymbol{U}_{\mathrm{RF}}^H\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{U}_{\mathrm{RF}}\boldsymbol{Q}\right|, \qquad (17)$$

whose solution $\boldsymbol{Q}_\star$ is given by the N dominant left singular vectors of $\boldsymbol{U}_{\mathrm{RF}}^H\boldsymbol{A}$. By undoing the change of variable, the solution to the overall problem $\boldsymbol{X}_{\mathrm{BB}} = \boldsymbol{V}_{\mathrm{RF}}\boldsymbol{S}_{\mathrm{RF}}^{-1}\boldsymbol{Q}_\star$ is found. For $\boldsymbol{Q} = \boldsymbol{Q}_\star$, the objective in (17) becomes

$$\log_2\left|\boldsymbol{I}_N + \epsilon\boldsymbol{A}^H\boldsymbol{U}_{\mathrm{RF}}\boldsymbol{U}_{\mathrm{RF}}^H\boldsymbol{A}\right| \leq \log_2\left|\boldsymbol{I}_N + \epsilon\boldsymbol{A}^H\boldsymbol{A}\right|, \qquad (18)$$

with the bound in (18) applying to any semi-unitary $\boldsymbol{U}_{\mathrm{RF}}$. This bound holds with equality if the columns of $\boldsymbol{U}_{\mathrm{RF}}$ are taken as the L dominant left singular vectors of $\boldsymbol{A}$.

References

[1] S. Rangan, T. S. Rappaport and E. Erkip, "Millimeter wave cellular wireless networks: Potentials and challenges," *Proc. IEEE,* vol. 102, no. 3, pp. 366–385, Mar. 2014.

[2] 3GPP "Physical channels and modulation. Release 15," TS 38.211, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213, Mar. 2018.

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

[4] R. W. Heath Jr., N. González-Prelcic, S. Rangan, W. Roh and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.,* vol. 10, no. 3, pp. 436–453, Apr. 2016.

[5] Z. Zhang, K. Long, A. V. Vasilakos and L. Hanzo, "Full-duplex wireless communications: Challenges, solutions, and future research directions," *Proc. IEEE,* vol. 104, no. 7, pp. 1369–1409, Jul. 2016.

[6] Z. Xiao, P. Xia and X.-G. Xia, "Full-Duplex millimeter-wave communication," *IEEE Wireless Commun.,* pp. 136–143, Dec. 2017.

[7] L. Song, R. Wichman, Y. Li and Z. Han, *Full-Duplex Communications and Networks,* Cambridge University Press, 2017.

[8] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.,* vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[9] D. Bharadia, E. McMilin and S. Katti, "Full duplex radios," in *Proc. ACM SIGCOMM,* Hong Kong, China, Aug. 2013, pp. 375–386.

[10] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, K. Haneda, R. Wichman and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full-duplex relays," *IEEE Commun. Mag.,* vol. 53, no. 5, pp. 91–101, May 2015.

[11] B. Day, A. Margetts, D. Bliss and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.,* vol. 60, no. 7, pp. 3702–3713, Jul. 2012.

[12] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. Wireless Commun.,* vol. 14, no. 4, pp. 2205–2217, Apr. 2015.

[13] E. Everett, C. Shepard, L. Zhong and A. Sabharwal, "SoftNull: Many-antenna full-duplex wireless via digital beamforming," *IEEE Trans. Wireless Commun.,* vol. 12, no. 15, pp. 8077–8092, Dec. 2016.

[14] X. Liu, Z. Xiao, L. Bai, J. Choi, P. Xia and X.-G. Xia, "Beamforming based Full-Duplex for millimeter-wave communication," *Sensors,* 16, 1130, Jul. 2016.

[15] R. López-Valcarce and N. González-Prelcic, "Analog beamforming for Full-Duplex millimeter wave communication," *Proc. Int. Symp. Wireless Commun. Syst.,* pp. 687–691, 2019.

[16] K. Satyanarayana, M. El-Hajjar, P.-H. Kuo, A. Mourad and L. Hanzo, "Hybrid beamforming design for Full-Duplex millimeter Wave communication," *IEEE Trans. Veh. Technol.,* vol. 68, pp. 1394–1404, Feb. 2019.

[17] R. Escalante and M. Raydan, *Alternating projection methods.* Society for Industrial and Applied Mathematics, vol. 8, 2011.