Multi-Stage/Hybrid BF under Limited Dynamic Range for OFDM FD Backhaul with MIMO SI Nulling

Christo Kurisummoottil Thomas, Chandan Kumar Sheemar, Dirk Slock
EURECOM, Sophia-Antipolis, France, Email:{kurisumm,Chandan.Sheemar,slock}@eurecom.fr

***Abstract*—This paper considers a bidirectional full-duplex Multi-Input Multi-Output (MIMO) OFDM system. We consider the more realistic noise model called limited dynamic range (LDR) model which takes into account the hardware impairments in the analog sections of the transceiver chains. At the transmit side, we introduce a two stage beamformer (BF) with an inner BF of lower dimension and an outer BF of higher dimension, both BFs being at the digital (baseband) side. The inner BF in OFDM domain handles transmission, while the outer BF in time domain handles self interference (SI). At the receive side, we propose a hybrid combiner which involves an analog phase shifter based BF, with fewer RF chains compared to the number of receive antennas and a digital (baseband) BF in OFDM domain. The analog BF helps reduce SI before analog-to-digital conversion (ADC). All BFs optimize alternatingly a weighted sum mean squared error (WSMSE) equivalent of the weighted sum rate over uplink (UL) and downlink (DL) channels. The proposed multi-stage BF architecture allows to reduce the coupling between classical transceiver design and MIMO SI nulling, and guarantee SI reduction during OFDM cyclic prefixes, with UL/DL possibly using different numerology or being asynchronous, allowing proper ADC operation.**

***Keywords*—** Full-Duplex, Hybrid Beamforming, Millimeter Wave, Weighted Sum Rate, Limited Dynamic Range.

I. Introduction

In this paper, Tx and Rx may denote transmit/transmitter/transmission and receive/receiver/reception. In-band full-duplex (FD) wireless, which allows each node to transmit and receive simultaneously has the potential to double the spectral efficiency and is one of the prominent candidates for 5G. It avoids the use of two independent channels for bi-directional communication, by allowing more flexibility in spectrum utilization, improving data security and reduces the air interface latency and delay issues. Unfortunately, it suffers from severe self-interference (SI) which could be 110 dB higher than the Rx signal power and canceling it is not a trivial task due to non-linearities and imperfections in the Tx chains, as identified in [1].

However, advancement in cancellation techniques have made FD operation possible. A combination of analog, digital and passive SIC techniques is required to reduce SI near the noise floor, by allowing signal reception with high signal-to-self-interference-plus-noise ratio. The first design and implementation of FD WiFi radio was introduced in [2]. In [3], SIC in FD is investigated experimentally and a practical FD system is proposed. In [4], the authors combine analog and digital SIC techniques and study the effect of residual SI together with clipping plus-quantization noise due to the limited dynamic range (LDR) of ADCs is studied. The analog cancellation stage is fundamental to reduce the SI sufficiently to ensure that it does not saturate the ADCs in the RX chains. It's complexity remains a serious challenge for upcoming massive MIMO FD scenarios, as it scales very poorly with the number of antennas. As discussed in [5], the next generation base stations (BS) will deploy 64-256 elements antenna elements. Therefore, the analog cancellation stage may become infeasible for upcoming communication scenarios, due to the large complexity associated. Also the cost of hardware components required to mimic the SI signal may become unattractive.

The use of separate Tx/Rx antenna arrays combined with various spatial precoding techniques has also been proposed to mitigate SI. In [6], two sequential convex programming (SCP) based algorithms for the joint optimization of beamforming (BF) and SIC are proposed. Recent studies on fully digital BF schemes under LDR using weighted sum rate (WSR) criteria for FD systems can be found in [7], [8].

Hybrid BF (HBF) [9] involves a two stage BF architecture providing both BF gain (by the analog phase shifting network) and spatial multiplexing gain by the digital BF. Various relevant studies on HBF designs for single user systems can be found in [10], [11].

A. Contributions of this paper

- We propose a two stage BF design for a bidirectional FD MIMO OFDM system based on the WSR criterion which is solved using the weighted sum mean square error (WSMSE) approach. At the Tx side, we propose to use a two stage BF at the baseband where the higher dimensional precoder is applied to the time domain signal, which aims to mitigates the SI and the lower dimensional precoder in the OFDM domain provides spatial multiplexing gain. At the Rx side, we introduce a HBF design. The objective of the time domain phase shifter analog BF stage is to reduce the SI significantly before the ADCs (such that the signal to the ADC is below the saturation level) while preserving the dimension of the desired signal space.
- Compared to the only existing state of the art design which introduces HBF for FD systems [12], we consider

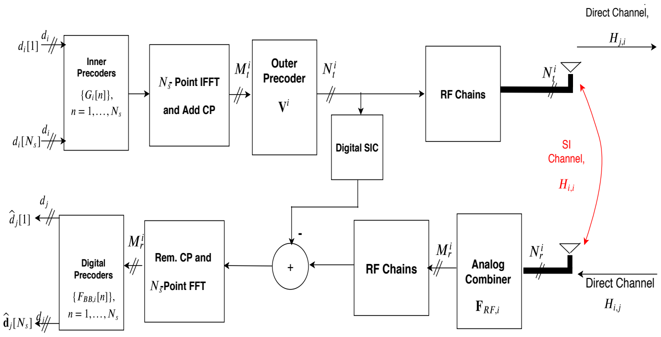

Fig. 1. Bidirectional FD MIMO OFDM System with Multi-Stage/Hybrid BF. Only a single node is shown for simplicity in the figure.

a more realistic LDR noise model at both the Tx and Rx.

- Through Monte Carlo simulations, we validate the performance of our proposed two-stage BF at Tx or hybrid combiner at Rx side. Simulations demonstrate that using an analog combiner stage at Rx (which operates before the Rx side LDR noise) has better sum rate performance compared to using a two-stage BF at Tx side.
- We also discuss the BF/combiner matrix dimensions and the number of Tx/Rx antennas such that the LDR noise (thus the residual SI also) gets reduced significantly.

Notation: In the following, boldface lower-case and upper-case characters denote vectors and matrices respectively. the operators $E\{\cdot\}$, tr$\{\cdot\}$, $(\cdot)^H$, $(\cdot)^T$ and $(\cdot)^*$ represent expectation, trace, conjugate transpose, transpose and complex conjugate, respectively. A circularly complex Gaussian random vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix Θ is distributed as $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \Theta)$. $\mathbf{V}_{1:d_k}(\mathbf{A}, \mathbf{B})$ represents the matrix formed by the (normalized) d_k dominant generalized eigenvectors of $\mathbf{A}$ and $\mathbf{B}$. $\mathbf{x} = vec(\mathbf{X})$ represents the vector obtained by stacking each of the columns of $\mathbf{X}$ and $unvec(\mathbf{x})$ represents the inverse operation of $vec(.)$. The operator $\otimes$ represents the Kronecker product.

II. Full-Duplex Bidirectional MIMO System Model

In this paper, we shall consider a multi-stream approach with d_j streams intended for the base station (BS) j. Two BSs are represented by the indices i and j respectively. So, consider a single user bidirectional FD backhaul system as depicted in Figure 1, with N_t^i or N_t^j Tx antennas at the BS i or j, respectively. We may also use the index 1 or 2 instead of i or j in the paper. Furthermore, we consider an OFDM system with N_s subcarriers. BSs are equipped with N_r^1 or N_r^2 receive antennas. $\mathbf{H}_{i,j}$, $i \neq j$ represents the $N_r^i \times N_t^j$ MIMO direct channel between node i and node j. Let $\mathbf{H}_{i,i}$ represent the SI channel from the Tx of node i to the Rx of node i. User i receives

$$\begin{aligned}\mathbf{y}_i[n] = &\mathbf{F}_{RF,i}\mathbf{H}_{i,j}[n](\mathbf{V}^j\,\mathbf{G}_j[n]\mathbf{d}_j[n] + \mathbf{c}_j[n]) + \\ &\mathbf{F}_{RF,i}\mathbf{H}_{i,i}[n](\mathbf{V}^i\,\mathbf{G}_i[n]\mathbf{d}_i[n] + \mathbf{c}_i[n]) + \mathbf{e}_i[n] + \mathbf{F}_{RF,i}\mathbf{n}_i[n],\end{aligned} \tag{1}$$

where $\mathbf{d}_j[n]$, of size $d_j \times 1$, is the intended signal stream vector (all entries are white, unit variance) to node i. At the Tx side, we have a two stage beamformer (inner BF, $\mathbf{G}_j$ of lower dimension and an outer BF, $\mathbf{V}^j$ of higher dimension), both the beamformers being at the digital (baseband) side. The outer BF will be applied to the time domain signal at the Tx side, so after the IFFT and it will be common to all the subcarriers. The inner BF will be different for different subcarriers. We are considering a noise whitened signal representation so that we get for the noise $\mathbf{n}_i \sim \mathcal{CN}(0, \mathbf{I}_{N_r^i})$. The higher dimensional outer precoder $\mathbf{V}^j$ at Tx of node j is of dimension $N_t^j \times M_t^j$. The digital beamformer is $\mathbf{G}_j$ which has dimensions $M_t^j \times d_j$, where $\mathbf{G}_j = \left[\mathbf{g}_j^{(1)} \ldots \mathbf{g}_j^{(d_j)}\right]$ and $\mathbf{g}_j^{(s)}$ represents the beamformer for stream s. $\mathbf{c}_i, \mathbf{e}_i$ represents the noise at the Tx or Rx antennas of node i respectively, which models the effect of LDR. LDR noise at Tx or Rx closely approximates the effects of non-ideal amplifiers, oscillators and ADCs/DACs. The covariance matrix of $\mathbf{c}_i$ is given by $\alpha_i(\alpha_i \ll 1)$ times the energy of the transmitted signal at each antenna. $\mathbf{c}_i$ is approximated as the Gaussian model, $\mathbf{c}_i[n] \sim \mathcal{CN}(\mathbf{0}, \frac{\alpha_i}{N_s}\,\text{diag}\,(\sum_{n=1}^{N_s}\mathbf{Q}_i[n]))$, where $\mathbf{Q}_i[n]$ is the Tx signal covariance matrix at subcarrier n of node i and can be written as $\mathbf{Q}_i[n] = \mathbf{V}^i\mathbf{G}_i[n]\mathbf{G}_i^H[n]\mathbf{V}^{i\,H}$ and $\mathbf{c}_i[n]$ is statistically independent of $\mathbf{x}_i[n]$. $\mathbf{e}_i[n]$ is the LDR noise at the Rx side and can be approximated as $\mathbf{e}_i[n] \sim \mathcal{CN}(\mathbf{0}, \frac{\beta_i}{N_s}\,\text{diag}\,(\mathbf{Z}))$, where $\mathbf{Z}$ is sum of the covariance matrix of the undistorted Rx signal across all subcarriers [13] assuming the subcarrier signals are decorrelated, $\mathbf{Z} = \sum_{n=1}^{N_s} \text{E}(\mathbf{z}_i[n]\mathbf{z}_i^H[n]), \mathbf{z}_i[n] = \mathbf{y}_i[n] - \mathbf{e}_i[n]$ and $\mathbf{e}_i[n]$ is statistically independent of $\mathbf{z}_i[n]$. Also, $\beta_i \ll 1$. The Tx power (sum of all subcarrier powers) constraint at node j can be written as $\sum_{n=1}^{N_s} \text{tr}\{\,\mathbf{V}^{j\,H}\mathbf{V}^j\;\mathbf{G}_j[n]\,\mathbf{G}_j^H[n]\,\} \leq P_j$. We introduce a digital self interference canceller at the base band which subtracts the residual interference signal $\mathbf{H}_{i,i}\mathbf{x}_i$ from the received signal. Assuming that $\mathbf{H}_{i,i}$ is perfectly estimated at the baseband and since $\mathbf{x}_i$ is already known to node i, we can rewrite the received signal at the baseband as,

$$\mathbf{y}'_i[n] = \mathbf{y}_i[n] - \mathbf{F}_{RF,i}\mathbf{H}_{i,i}[n]\mathbf{x}_i[n] = \mathbf{F}_{RF,i}\mathbf{H}_{i,j}[n]\mathbf{x}_j[n] + \mathbf{v}_i[n], \tag{2}$$

where $\mathbf{v}_i[n] = \mathbf{F}_{RF,i}\mathbf{H}_{i,j}[n]\mathbf{c}_j[n] + \mathbf{F}_{RF,i}\mathbf{H}_{i,i}[n]\mathbf{c}_i[n] + \mathbf{e}_i[n] + \mathbf{F}_{RF,i}\mathbf{n}_i[n]$ is the unknown interference plus noise component after SI cancellation. In this paper, for our BF design, we assume that all the channel matrices and scaling factors in (1) are known. Also, another point worth noting here is that the dependence of the signal model (2) on the SI power is only through the LDR noise and the BF design in the next section try to reduce the LDR noise significantly.

A. Channel Model

In this sub-section, we omit the node indices for simplicity. Considering a delay-d geometric direct channel model for a mmWave propagation environment [14] with L_s scattering clusters and L_r scatterers or rays in each cluster, we have,

$$\mathbf{H}_d = \sum_{s=1}^{L_s}\sum_{l=1}^{L_r} \alpha_{s,l}\mathbf{h}_r(\theta_{s,l})\mathbf{h}_t(\phi_{s,l})^H p(dT_s - \tau_s - \tau_{rl}) \tag{3}$$

Here $\theta_{s,l}, \phi_{s,l}$ represent the angle of arrival (AoA) and angle of departure (AoD) respectively for the l^{th} path in the s^{th} cluster.

$\mathbf{h}_r(\cdot), \mathbf{h}_t(\cdot)$ represent the antenna array responses at Rx and Tx respectively. The complex path gain, $\alpha_{s,l} \sim \mathcal{CN}(0, \frac{N_t N_r}{L_s L_r})$ and $p(\tau)$ represents the band-limited pulse shaping filter response evaluated at τ seconds. Each cluster has a time delay $\tau_s \in \mathcal{R}$ and each ray $l = 1, ..L_r$ has a relative time delay τ_{rl}. Now, we write the $(m, n)-$the element of the channel in the subcarrier n as,

$$\mathbf{H}[n] = \sum_{d=1}^{D} \mathbf{H}_d e^{-j2\pi \frac{nd}{N_s}}. \quad (4)$$

In a more compact form, this can be represented as,

$$\begin{aligned} &\mathbf{H}[n] = \mathbf{H}_r \sum_{d=1}^{D} \mathbf{A}_d[n]\mathbf{H}_t^H, \text{ where} \\ &\mathbf{H}_r = [\mathbf{h}_r(\theta_{1,1}), ..., \mathbf{h}_r(\theta_{L_s,L_r})], \\ &\mathbf{H}_t = [\mathbf{h}_t(\phi_{1,1}), ..., \mathbf{h}_t(\phi_{L_s,L_r})], \\ &\mathbf{A}_d[n] = \text{ diag}\,(\alpha_{1,1} p(dTs - \tau_1 - \tau_{r1}), \\ &..., \alpha_{L_s,L_r} p(dTs - \tau_{L_s} - \tau_{rL_r}))e^{-j2\pi \frac{nd}{N_s}}. \end{aligned} \quad (5)$$

Note that our HBF design which follows, is applicable for general MIMO channel models and the channel model outlined here is utilized for the simulations in Section VI. Further considering the SI channel, as the distance between the transmit and receive arrays doesn't satisfy the far-field range condition, we need to employ the near-field model which has spherical wavefront. In such a case, the SI channel coefficients highly depend on the placement of the transmit and receive arrays and can be written as,

$$(\mathbf{H}_{i,i})_{m,n} = \frac{\rho}{r_{m,n}} \exp(-j2\pi \frac{r_{m,n}}{\lambda}), \quad (6)$$

where $r_{m,n}$ is the distance between $m-$th element of the receive array and $n-$the element of the transmit array and ρ being the SI channel power normalization factor. Note that, (6) is a simple model which doesn't take into account the mutual antenna coupling or signal reflections in the SI channel.

III. WSR MAXIMIZATION THROUGH WSMSE

Consider the optimization of the two-stage BF/hybrid combiner design using WSR maximization of the Multi-cell MU-MIMO system:

$$\begin{aligned} [\mathbf{V}\ \mathbf{G}\ \mathbf{F}_{RF}\ \mathbf{F}_{BB}] &= \arg \max_{\substack{\mathbf{V},\mathbf{G}, \\ \mathbf{F}_{RF},\mathbf{F}_{BB}}} \quad WSR(\mathbf{G}, \mathbf{V}, \mathbf{F}_{RF}, \mathbf{F}_{BB}) \\ &= \arg \max_{\mathbf{V},\mathbf{G}} \sum_{i=1}^{2} \sum_{n=1}^{N_s} u_i \ln\det(\mathbf{R}_{\bar{i}}^{-1}[n]\mathbf{R}_i[n]), \end{aligned} \quad (7)$$

where the u_i are the rate weights, $\mathbf{G}$ represents the collection of digital BFs $\mathbf{G}_i[n]$, $\mathbf{V}$ the collection of analog BFs $\mathbf{V}^i$. At the receiver, we apply a hybrid combiner with analog BF denoted by $\mathbf{F}_{RF,i}$ of size $M_r^i \times N_r^i$, where M_r^i represents the number of RF chains at the Rx side. $\mathbf{F}_{BB,i}[n]$ represent the baseband digital combiner of size $d_j \times M_r^i$. For notational convenience, we define the received signal covariance matrices $\mathbf{\Theta}_{i,j}[n] = \mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n], \mathbf{\Phi}_{i,j}[n] = \mathbf{H}_{i,j}[n] \text{ diag}\,(\mathbf{Q}_j[n])\mathbf{H}_{i,j}^H[n]$. Similarly the self interence parts $\mathbf{\Theta}_{i,i}[n], \mathbf{\Phi}_{i,i}[n]$ are also defined. The covariance matrix of the effective noise part at the output of the RF chains, $\mathbf{R}_{\bar{i}}[n]$ can be approximated under $\alpha_i \ll 1, \beta_i \ll 1$ as follows [15],

$$\begin{aligned} &\mathbf{R}_{\bar{i}}[n] = \mathbf{F}_{RF,i}(\alpha_j \mathbf{\Phi}_{i,j}[n] + \alpha_i \mathbf{\Phi}_{i,i}[n])\mathbf{F}_{RF,i}^H + \\ &\beta_i \text{ diag}\,(\mathbf{F}_{RF,i}(\mathbf{\Theta}_{i,j}[n] + \mathbf{\Theta}_{i,i}[n])\mathbf{F}_{RF,i}^H) \end{aligned} \quad (8)$$

Also define, $\mathbf{R}_i[n] = \mathbf{R}_{\bar{i}}[n] + \mathbf{F}_{RF,i}\mathbf{\Theta}_{i,j}[n]\mathbf{F}_{RF,i}^H$, where $\mathbf{R}_i[n]$ is the signal plus interference plus noise covariance matrix. Further after the receive combining, we obtain $\mathbf{\Sigma}_{\bar{i}}[n] = \mathbf{F}_{BB,i}[n]\mathbf{R}_{\bar{i}}[n]\mathbf{F}_{BB,i}^H[n]$ and $\mathbf{\Sigma}_i[n] = \mathbf{F}_{BB,i}[n]\mathbf{R}_i[n]\mathbf{F}_{BB,i}^H[n]$. Direct maximization of (7), however, requires a joint optimization over the four matrix variables $(\mathbf{V}, \mathbf{G}, \mathbf{F}_{RF}, \mathbf{F}_{BB})$. Unfortunately, finding a global optimum solution for similar constrained optimization is found to be intractable. So we decouple the joint transmitter-receiver optimization and focus on the design of the Rx combiners first. We assume that the node i applies the hybrid combiner $\mathbf{F}_i[n] = \mathbf{F}_{BB,i}[n]\mathbf{F}_{RF,i}$ to estimate the signal transmitted from node j. The analog combiner $\mathbf{F}_{RF,i}$ serves to reduce the SI component from the received signal, while the digital combiner $\mathbf{F}_{BB,i}$ decouples the streams $(\mathbf{d}_j)$ intended for user i from j. The estimated signal $\widehat{\mathbf{d}}_j[n]$ can be written as,

$$\widehat{\mathbf{d}}_j[n] = \mathbf{F}_i[n]\mathbf{H}_{i,j}[n]\mathbf{x}_j[n] + \mathbf{F}_{BB,i}[n]\mathbf{v}_i[n]. \quad (9)$$

At the Rx side, maximizing the WSR is equivalent to minimizing the weighted MSE with the MSE weights being chosen as $\mathbf{W}_i[n] = u_i \mathbf{R}_{\widetilde{\mathbf{d}}_j\widetilde{\mathbf{d}}_j}[n]^{-1}$ [7], [16]. Further we can obtain the error covariance matrix for the detection of $\mathbf{d}_j$ at node i as,

$$\begin{aligned} &\mathbf{R}_{\widetilde{\mathbf{d}}_j\widetilde{\mathbf{d}}_j}[n] = \mathrm{E}\{(\widehat{\mathbf{d}}_j[n] - \mathbf{d}_j[n])(\widehat{\mathbf{d}}_j[n] - \mathbf{d}_j[n])^H\} = \\ &(\mathbf{F}_i[n]\mathbf{H}_{i,j}[n]\mathbf{V}^j\mathbf{G}_j[n] - \mathbf{I})(\mathbf{F}_i[n]\mathbf{H}_{i,j}[n]\mathbf{V}^j\mathbf{G}_j[n] - \mathbf{I})^H + \\ &\mathbf{F}_{BB,i}\mathbf{R}_{\bar{i}}[n]\mathbf{F}_{BB,i}[n]^H. \end{aligned} \quad (10)$$

The MMSE Rx combiner can be alternatively optimized as follows,

$$\begin{aligned} &[\mathbf{F}_{RF,i},\ \mathbf{F}_{BB,i}[n], \forall n] = \arg \min_{\mathbf{F}_{RF,i},\mathbf{F}_{BB,i}[n]} \sum_{n=1}^{N_s} \mathrm{tr}\{\mathbf{R}_{\widetilde{\mathbf{d}}_j\widetilde{\mathbf{d}}_j}[n]\}, \\ &\mathbf{F}_{BB,i}[n] = \mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{F}_{RF,i}^H\mathbf{R}_i[n]^{-1} \end{aligned} \quad (11)$$

Optimization of the digital BF in (11) can be done independently across different subcarriers, as it is evident. We define $\mathbf{F}_{BB,i}^H[n]\mathbf{F}_{BB,i}[n] = \mathbf{P}_{B,i}[n]$, $\sum_{n=1}^{N_s}[(\mathbf{\Theta}_{i,j}[n])^T \otimes \mathbf{P}_{B,i}[n] + ((\alpha_j\mathbf{\Phi}_{i,j}[n] + \alpha_i\mathbf{\Phi}_{i,i}[n])^T \otimes \mathbf{P}_{B,i}[n]) + (\beta_i(\mathbf{\Theta}_{i,j}[n] + \mathbf{\Theta}_{i,i}[n])^T \otimes \text{ diag}\,(\mathbf{P}_{B,i}[n])) = \mathbf{B}_i$. To derive the unconstrained analog BF matrix, we take the gradient of (10) w.r.t $\mathbf{F}_{RF,i}$,

$$\begin{aligned} &\sum_{n=1}^{N_s} \mathbf{P}_{B,i}[n]\mathbf{F}_{RF,i}\mathbf{\Theta}_{i,j}[n] - \mathbf{F}_{BB,i}^H[n]\mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n] + \\ &\mathbf{P}_{B,i}[n]\mathbf{F}_{RF,i}(\alpha_j\mathbf{\Phi}_{i,j}[n] + \alpha_i\mathbf{\Phi}_{i,i}[n]) + \\ &\beta_i \text{ diag}\,(\mathbf{P}_{B,i}[n])\mathbf{F}_{RF,i}(\mathbf{\Theta}_{i,j}[n] + \mathbf{\Theta}_{i,i}[n]) = 0, \\ &\mathbf{B}_i vec(\mathbf{F}_{RF,i}) \overset{(a)}{=} vec(\sum_{n=1}^{N_s} \mathbf{F}_{BB,i}^H[n]\mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]). \end{aligned} \quad (12)$$

In (a), we use the result $vec(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})vec(\mathbf{X})$ [17]. Further we obtain the expression for the analog combiner as,

$$vec(\mathbf{F}_{RF,i}) = \mathbf{B}_i^\dagger vec(\sum_{n=1}^{N_s} \mathbf{F}_{BB,i}^H[n]\mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]), \quad (13)$$

where $\dagger$ represents the pseudoinverse.

A. Two stage transmit BF design

In this section, we consider the design of two stage Tx BFs $\mathbf{V}^j, \mathbf{G}_j[n]$ under a sum power constraint at the Tx. To faciliate the gradients, we use the result $\frac{\partial \mathrm{tr}\{\mathbf{A}\,\text{diag}\,(\mathbf{CXD})\mathbf{B}\}}{\partial \mathbf{X}} = [\mathbf{D}\,\text{diag}\,(\mathbf{BA})\mathbf{C}]^T$. However, due to space limitations, we skip the derivations for this gradient result. We propose to design the Tx BFs using weighted sum MSE and can be formulated as follows,

$$\begin{aligned}
&\min_{\substack{\mathbf{V}^i,\mathbf{G}_i[n],\\ \mathbf{V}^j,\mathbf{G}_j[n]}} \sum_{n=1}^{N_s} \text{tr}\{\mathbf{W}_i[n]\,\text{E}(\widehat{\mathbf{d}}_j[n]-\mathbf{d}_j[n])(\widehat{\mathbf{d}}_j[n]-\mathbf{d}_j[n])^H\}+\\
&\text{tr}\{\mathbf{W}_j[n]\,\text{E}\{(\widehat{\mathbf{d}}_i[n]-\mathbf{d}_i[n])(\widehat{\mathbf{d}}_i[n]-\mathbf{d}_i[n])^H\},\\
&\text{s.t.} \sum_{n=1}^{N_s} \text{tr}\{\mathbf{Q}_i[n]\} <= P_i, \forall i.
\end{aligned} \tag{14}$$

Here $\mathbf{W}_i[n]$ represents the weight matrix of size $d_i \times d_i$. Augmenting the power constraints, the Lagrangian function can be written as,

$$\begin{aligned}
\mathcal{L}=&\sum_{n=1}^{N_s}\sum_{i=1}^{2}\sum_{j=1,j\neq i}^{2} \text{tr}\{\mathbf{W}_i[n](\mathbf{I}-\mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\\
&-\mathbf{F}_i[n]\mathbf{H}_{i,j}[n]\mathbf{V}^j\mathbf{G}_j[n]+\mathbf{F}_i[n]\mathbf{H}_{i,j}[n]\mathbf{Q}_j\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]+\\
&\mathbf{F}_{BB,i}[n]\mathbf{R}_{\bar{i}}[n]\mathbf{F}_{BB,i}^H[n])\}+(\sum_{i=1}^{2}\lambda_i(\sum_{n=1}^{N_s}\text{tr}\{\mathbf{Q}_i[n]\})-P_i),
\end{aligned} \tag{15}$$

For convenience of the analysis we define $\mathbf{A}_j[n] = \mathbf{F}_j^H[n]\mathbf{W}_j[n]\mathbf{F}_j[n], \widehat{\mathbf{A}}_j[n] = \mathbf{F}_{BB,j}^H[n]\mathbf{W}_j[n]\mathbf{F}_{BB,j}[n]$. Taking the partial derivative of (15) with respect to the inner BF $\mathbf{G}_j[n]$, we obtain,

$$\begin{aligned}
&-\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n]+\\
&\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{A}_i[n]\mathbf{H}_{i,j}[n]\mathbf{V}^j\mathbf{G}_j[n]\\
&+\frac{\text{tr}\{\mathbf{F}_{BB,i}[n]^H\mathbf{F}_{BB,i}[n]\partial\mathbf{R}_{\bar{i}}[n]\}}{\partial\mathbf{G}_j[n]}+\frac{\text{tr}\{\mathbf{F}_{BB,j}[n]^H\mathbf{F}_{BB,j}[n]\partial\mathbf{R}_{\bar{j}}[n]\}}{\partial\mathbf{G}_j[n]}\\
&+\lambda_j\mathbf{V}^{j\,H}\mathbf{V}^j\mathbf{G}_j[n]=\mathbf{0},\text{ where, } i\neq j.
\end{aligned} \tag{16}$$

Using the expression for $\mathbf{R}_{\bar{i}}[n]$ in (8), we can write,

$$\begin{aligned}
&\frac{\text{tr}\{\mathbf{F}_{BB,i}[n]^H\mathbf{F}_{BB,i}[n]\partial\mathbf{R}_{\bar{i}}[n]\}}{\partial\mathbf{G}_j[n]}=\\
&\alpha_j\mathbf{V}^{j\,H}\,\text{diag}(\mathbf{H}_{i,j}^H[n]\mathbf{A}_i[n]\mathbf{H}_{i,j}[n])\mathbf{V}^j\mathbf{G}_j[n]\\
&+\beta_i\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{F}_{RF,i}^H\,\text{diag}(\widehat{\mathbf{A}}_i[n])\mathbf{F}_{RF,i}\mathbf{H}_{i,j}[n]\mathbf{V}^j\mathbf{G}_j[n],\\
&\frac{\text{tr}\{\mathbf{F}_{BB,j}[n]^H\mathbf{F}_{BB,j}[n]\partial\mathbf{R}_{\bar{j}}[n]\}}{\partial\mathbf{G}_j[n]}=\\
&\alpha_j\mathbf{V}^{j\,H}\,\text{diag}(\mathbf{H}_{j,j}^H[n]\mathbf{A}_j[n]\mathbf{H}_{j,j})\mathbf{V}^j\mathbf{G}_j[n]\\
&+\beta_j\mathbf{V}^{j\,H}\mathbf{H}_{j,j}^H[n]\mathbf{F}_{RF,i}^H\,\text{diag}(\widehat{\mathbf{A}}_j[n])\mathbf{F}_{RF,i}\mathbf{H}_{j,j}[n]\mathbf{V}^j\mathbf{G}_j[n],
\end{aligned} \tag{17}$$

By substituting (17) in (16), we obtain the optimal $\mathbf{G}_j[n]$ as,

$$\mathbf{G}_j[n]=(\mathbf{S}_j[n]+\lambda_j\mathbf{V}^{j\,H}\mathbf{V}^j)^{-1}\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n], \tag{18}$$

where $\mathbf{S}_j[n]$ can be interpreted as the signal plus interference power seen by the digital BF at the Tx side and is expressed as,

$$\begin{aligned}
\mathbf{S}_j[n]=&\,\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{A}_i[n]\mathbf{H}_{i,j}[n]\mathbf{V}^j+\\
&\alpha_j\mathbf{V}^{j\,H}\,\text{diag}\,(\mathbf{H}_{i,j}^H[n]\mathbf{A}_i[n]\mathbf{H}_{i,j}[n])\mathbf{V}^j\\
&+\beta_i\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{F}_{RF,i}^H\,\text{diag}\,(\widehat{\mathbf{A}}_i[n])\mathbf{F}_{RF,i}\mathbf{H}_{i,j}[n]\mathbf{V}^j+\\
&\alpha_j\mathbf{V}^{j\,H}\,\text{diag}\,(\mathbf{H}_{j,j}^H[n]\mathbf{A}_j[n]\mathbf{H}_{j,j}[n])\mathbf{V}^j\\
&+\beta_j\mathbf{V}^{j\,H}\mathbf{H}_{j,j}^H[n]\mathbf{F}_{RF,j}^H\,\text{diag}\,(\widehat{\mathbf{A}}_j[n])\mathbf{F}_{RF,j}\mathbf{H}_{j,j}[n]\mathbf{V}^j
\end{aligned} \tag{19}$$

The values of the Lagrangian multipliers $\lambda_j \geq 0, \forall j$ are chosen such that the respective power constraint is satisfied (14). To compute this, we follow a similar approach as in [18], but extended to two-stage BF here. Considering the eigen decomposition of $\mathbf{S}_j[n]=\mathbf{U}_j\mathbf{\Lambda}_j[n]\mathbf{U}_j^H, \mathbf{V}^{j\,H}\mathbf{V}^j=\mathbf{U}_j\mathbf{\Delta}_j\mathbf{U}_j^H$ and let $\mathbf{\Phi}[n]=\mathbf{U}_j^H\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n]\mathbf{W}_i[n]^H\mathbf{F}_i[n]\mathbf{H}_{i,j}[n]\mathbf{V}^j\mathbf{U}_j$ and expanding the power constraint $\sum_{n=1}^{N_s}\text{tr}\{\mathbf{V}^j\mathbf{G}_j[n](\lambda_j)\mathbf{G}_j^H[n](\lambda_j)\mathbf{V}^{j\,H}\} = P_j$, we get the simplified expression,

$$\sum_{n=1}^{N_s}\sum_{k=1}^{M_t^j}\frac{\mathbf{\Phi}[n]_{k,k}(\mathbf{\Delta}_j)_{k,k}}{((\mathbf{\Lambda}_j[n])_{k,k}+\lambda_j(\mathbf{\Delta}_j)_{k,k})^2}=P_j. \tag{20}$$

Here $\mathbf{X}_{k,k}$ represents the k^{th} diagonal element of the matrix $\mathbf{X}$. Note that the $\lambda_j \geq 0$ and the left hand side of (20) is a decreasing function of λ_j for $\lambda_j > 0$. Hence we can compute the values of λ_j using one dimensional linear search techniques such as bisection. Further we consider the optimization of the outer BF at the Tx side, $\mathbf{V}^j$. Given the inner BFs, we update the outer beamformers $\mathbf{V}^j$. Taking the partial derivative of (15) with respect to the inner BF $\mathbf{V}^j$, we obtain,

$$\begin{aligned}
&-\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n]\mathbf{G}_j^H[n]+\\
&\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n]\mathbf{F}_i[n]\mathbf{H}_{i,j}[n]\mathbf{V}^j\mathbf{G}_j[n]\mathbf{G}_j^H[n]\\
&+\frac{\text{tr}\{\mathbf{F}_{BB,i}[n]^H\mathbf{F}_{BB,i}[n]\partial\mathbf{R}_{\bar{i}}[n]\}}{\partial\mathbf{V}^j[n]}+\frac{\text{tr}\{\mathbf{F}_{BB,j}[n]^H\mathbf{F}_{BB,j}[n]\partial\mathbf{R}_{\bar{j}}[n]\}}{\partial\mathbf{V}^j[n]}+\\
&\lambda_j\mathbf{V}^j\mathbf{G}_j[n]\mathbf{G}_j^H[n]=0,\text{where, } i\neq j.
\end{aligned} \tag{21}$$

For notational convenience, we define $\mathbf{P}_{G,j}[n] = \mathbf{G}_j[n]\mathbf{G}_j^H[n]$. Using the expression for $\mathbf{R}_{\bar{i}}[n]$ in (8), we can write,

$$\begin{aligned}
&\frac{\text{tr}\{\mathbf{F}_{BB,i}[n]^H\mathbf{F}_{BB,i}[n]\partial\mathbf{R}_{\bar{i}}[n]\}}{\partial\mathbf{V}^j[n]}=\alpha_j\,\text{diag}\,(\mathbf{H}_{i,j}^H[n]\mathbf{A}_i[n]\mathbf{H}_{i,j}[n])\\
&\mathbf{V}^j\mathbf{P}_{G,j}[n]+\beta_i\mathbf{H}_{i,j}^H[n]\mathbf{F}_{RF,i}^H\,\text{diag}\,(\widehat{\mathbf{A}}_i[n])\mathbf{F}_{RF,i}\mathbf{H}_{i,j}[n]\\
&\mathbf{V}^j\mathbf{P}_{G,j}[n],\frac{\text{tr}\{\mathbf{F}_{BB,j}[n]^H\mathbf{F}_{BB,j}[n]\partial\mathbf{R}_{\bar{j}}[n]\}}{\partial\mathbf{V}^j[n]}=\\
&\alpha_j\,\text{diag}\,(\mathbf{H}_{j,j}^H[n]\mathbf{A}_j[n]\mathbf{H}_{j,j}[n])\mathbf{V}^j\mathbf{P}_{G,j}[n]\\
&+\beta_j\mathbf{H}_{j,j}^H[n]\mathbf{F}_{RF,j}^H\,\text{diag}\,(\widehat{\mathbf{A}}_j[n])\mathbf{F}_{RF,j}\mathbf{H}_{j,j}[n]\mathbf{V}^j\mathbf{P}_{G,j}[n]],
\end{aligned} \tag{22}$$

By substituting (22) in (21) and using the result $vec(\mathbf{AXB}) = (\mathbf{B}^T\otimes\mathbf{A})vec(\mathbf{X})$, we obtain the optimal $\mathbf{V}^j$ as,

$$\begin{aligned}
&vec(\mathbf{V}^j)=\mathbf{B}_j^\dagger\sum_{n=1}^{N_s}\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n]\mathbf{G}_j^H[n],\text{ where}\\
&\mathbf{B}_j=\sum_{n=1}^{N_s}(\mathbf{P}_{G,j}[n]\otimes\mathbf{H}_{i,j}^H[n]\mathbf{A}_i[n]\mathbf{H}_{i,j}[n])+\\
&\alpha_j\mathbf{P}_{G,j}[n]\otimes\,\text{diag}\,(\mathbf{H}_{i,j}^H[n]\mathbf{A}_i[n]\mathbf{H}_{i,j}[n])\\
&+\beta_i\mathbf{P}_{G,j}[n]\otimes\mathbf{H}_{i,j}^H[n]\mathbf{F}_{RF,i}^H\,\text{diag}\,(\widehat{\mathbf{A}}_i[n])\mathbf{F}_{RF,i}\mathbf{H}_{i,j}[n]+\\
&\alpha_j\mathbf{P}_{G,j}[n]\otimes\,\text{diag}\,(\mathbf{H}_{j,j}^H\mathbf{A}_j\mathbf{H}_{j,j}[n])\\
&+\beta_j(\mathbf{P}_{G,j}[n]\otimes(\mathbf{H}_{j,j}^H[n]\mathbf{F}_{RF,j}^H\,\text{diag}\,(\widehat{\mathbf{A}}_j[n])\mathbf{F}_{RF,j}\mathbf{H}_{j,j}[n])).
\end{aligned} \tag{23}$$

Alternating WSR maximization between digital and analog BF or the two stage BFs at Tx/Rx now leads to Algorithm 1. We remark that we propose to either use a two-stage BF at Tx or hybrid combiner at the Rx to null the SI power and both stages

---

Algorithm 1 LDR Multi Stage BF Design via WSMSE

Given: $P_i, \mathbf{H}_{i,j}, \mathbf{H}_{i,i}, u_i\,\forall i, j$.

Initialization: $\mathbf{F}_{RF,i} = e^{j\angle\mathbf{V}_{1:M^i}(\mathbf{H}_{t,i,j})}$, The $\mathbf{G}_i$ are taken as the ZF precoders for the effective channels $\mathbf{V}^i\mathbf{H}_{j,i}$ with uniform powers.

Iteration (t) :

1) Update the Rx side HBF, i.e $\mathbf{F}_{BB,i}^{(t)}[n], \mathbf{F}_{RF,i}^{(t)}\forall i$ using (11), (13) respectively.
2) Update $\mathbf{G'}_i^{(t)}[n]$, $\forall i$, from (18).
3) Update $\mathbf{V}^{i\,(t)}, \forall i$ from (23) and λ_i using bisection method from (20).
4) Check for convergence of the WSR: if not go to step 1).

---

are not required if the antenna or BF/combiner dimensions are sufficient as discussed in Section IV.

Directly optimizing the phasor values of the analog combiner alternatively using the WSR cost function which is a non-convex function results in lot of local optima depending on the initialization [19]. So we utilize here one approach called deterministic annealing (DA) to avoid the problem of local optima and it is discussed in detail in our papers [20, Algorithm 3], [21].

IV. Hybrid Combiner/Two-Stage BF Capabilities for SI Power Reduction

In this section we analyze to what extent a hybrid combiner can achieve the same performance as a fully digital BF and reduce the LDR noise originating from both the direct and SI channels. In particular we shall see that this is possible for a sufficient number of RF chains and with the arbitrary antenna array responses. Consider a specular or pathwise channel model with say L_d multi-paths per link for the direct channel and L_I for the SI channel. For notational simplicity we shall consider a uniform L_d, L_I and $N_k = N_t^i = N_r^i, \forall i$.

Theorem 1. *For a bidirectional full-duplex MIMO system with the number of Rx RF chains $M_r^i \geq L_d$ or the number of Tx RF chains $M_t^i >= L_d$ and arbitrary antenna responses for the direct channel, to achieve optimal all-digital precoding performance at high SNR and mitigation of LDR noise, the unconstrained analog combiner or the time domain Tx BF can be chosen as matched filtered to the direct link channel projected on the orthogonal complement of the low rank SI channel.*

Proof: From [16] or [22, eq. (13)], the optimal all-digital beamformer is of the form

$$\begin{aligned}&\mathbf{F}_i[n] =\\ &\mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{i,j}^H[n](\mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n]+\mathbf{R}_{\bar{i}}[n])^{-1}\\ &=\mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{t,i,j}\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n]^H\mathbf{H}_{r,i,j}^H\\ &(\mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n]+\mathbf{R}_{\bar{i}}[n])^{-1},\end{aligned} \tag{24}$$

where $\mathbf{R}_{\bar{i}}[n]$ is the interference plus noise power received. For the mmWave channel model (5), when $N_k \to \infty$, the terms of the form $\mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n]$ can be simplified as,

$$\begin{aligned}&\mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n] =\\ &\mathbf{H}_{r,i,j}(\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n])\mathbf{H}_{t,i,j}^H\mathbf{Q}_j[n]\mathbf{H}_{t,i,j}(\sum_{d=1}^{D}\mathbf{A}_{d,i,j}^H[n])\mathbf{H}_{r,i,j}^H\\ &\overset{(a)}{=}\frac{1}{N_t^j}\mathbf{H}_{r,i,j}(\sum_{d=1}^{D}\mathbf{A}_{d,i,j}^2)\text{tr}\{\mathbf{Q}_j[n]\}\mathbf{H}_{r,i,j}^H,\end{aligned} \tag{25}$$

where $\sum_{d=1}^{D}\mathbf{A}_{d,i,j}^2 = \sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n]\mathbf{A}_{d,i,j}^H[n]$ is independent of the subcarrier index. In (a), we made the assumption that the Tx array response becomes asymptotically orthogonal. Further assuming that at high SNR the power transmitted across each subcarrier becomes same, then $\text{tr}\{\mathbf{Q}_j[n]\} = \frac{P_j}{N_s}$ and thus $\mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n]$ becomes independent of the frequency. Similarly $\mathbf{R}_{\bar{i}}[n]$ also becomes independent of the frequency since the terms in $\mathbf{R}_{\bar{i}}[n]$ are also of similar form as $\mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n]$. We denote $\mathbf{R}_i = \mathbf{H}_{i,j}[n]\mathbf{Q}_j[n]\mathbf{H}_{i,j}^H[n] + \mathbf{R}_{\bar{i}}[n]$. Thus we can separate the BFs as

$$\begin{aligned}&\mathbf{F}_{RF,i} = \mathbf{H}_{r,i,j}^H\mathbf{R}_i^{-1},\\ &\mathbf{F}_{BB,i}[n] = \mathbf{G}_j^H[n]\mathbf{V}^{j\,H}\mathbf{H}_{t,i,j}\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n].\end{aligned} \tag{26}$$

Similarly considering the Tx side BF design, the optimal fully digital BF can be written as (18),

$$\mathbf{G}_j[n] = (\mathbf{S}_j[n] + \lambda_j\mathbf{I})^{-1}\mathbf{H}_{i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n], \tag{27}$$

As $N_r^i \to \infty$ and substituting the pathwise model for the direct channels similar to the discussions above, we can observe that the quadratic term $(\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n]^H)\mathbf{H}_{r,i,j}^H\mathbf{A}_i\mathbf{H}_{r,i,j}(\sum_{d=1}^{D}\mathbf{A}_d[n]) = \mathbf{P}_r^i[n]$, where $\mathbf{P}_r^i[n]$ can be interpreted as the effective received power in subcarrier n, $\mathbf{P}_r^i[n] = \frac{1}{N_r^i}(\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n]^2)\text{tr}\mathbf{A}_i$. $(\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n]^2)$ is independent of subcarrier index (5) and hence effective received power in all the subcarrier becomes the same in the large antenna limit. Further, substituting the $(\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n]^H)\mathbf{H}_{r,i,j}^H\mathbf{A}_i\mathbf{H}_{r,i,j}(\sum_{d=1}^{D}\mathbf{A}_d[n])$ in (19), we can see that $\mathbf{S}_j[n]$ is independent of the subcarrier index. Further defining $\widehat{\mathbf{S}}_j = \mathbf{S}_j[n] + \lambda_j\mathbf{I}$, we obtain, $\mathbf{V}^j = \widehat{\mathbf{S}}_j^{-1}\mathbf{H}_{t,i,j}$ and $\mathbf{G}_j = (\sum_{d=1}^{D}\mathbf{A}_{d,i,j}[n]^H)\mathbf{H}_{r,i,j}^H[n]\mathbf{F}_i^H[n]\mathbf{W}_i[n]$. Hence we can conclude that $\mathbf{V}^j, \mathbf{F}_{RF,j}$ depends only on the Tx/Rx antenna array responses. □

Note that whereas the digital BF or combiner $\mathbf{G}, \mathbf{F}_{BB}$ in (26) is a function of the instantaneous CSIT, the analog combiner $\mathbf{F}_{RF}$ or the outer precoder $\mathbf{V}$ is only a function of antenna array responses of the direct and SI channels, hence only of the slow fading channel components. Hence analog BF can be the same across all the subcarriers. We also remark that at high SNR, $\widehat{\mathbf{R}}_i$ or $\widehat{\mathbf{S}}_j$ converges to the projection matrix for the null space of the SI channel's Rx or Tx antenna array response matrix respectively. We remark that the main advantage of adding an analog BF stage is to suppress the SI before reaching the ADC and still preserving the signal dimensions by choosing sufficient number of RF chains. Also note that the analytical analog BF solution discussed here is unconstrained and further it requires DA method to reach a phasor BF solution.

V. Simulation Results

Extensive Monte-Carlo simulations are conducted to validate the performance of the proposed hybrid BF algorithms are presented for a bidirectional FD system under LDR noise model. We follow the pathwise channel model $\mathbf{H}_{i,j}$ as in Section II.A, where the complex path gains are assumed to be Gaussian with variance distributed according to an exponential profile. For the SI channel, we ignore the near field effect of amplitude variation with distance and the near field effects in the phase variation. In the Uniform Linear Array (ULA), the

AoD or AoA ϕ, θ are assumed to be uniformly distributed in the interval $[0^o, 30^o]$.

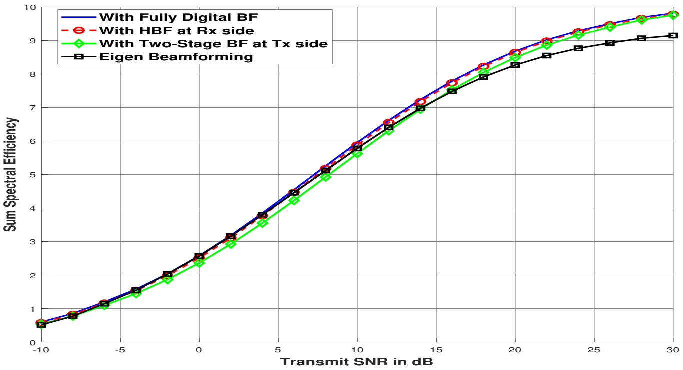

Fig. 2. Sum Rate comparisons for, Single Carrier, $N_t^i = N_r^i = 8, M_t^i = M_r^i = 4, d_i = 1, \forall i, L = 4$ paths.

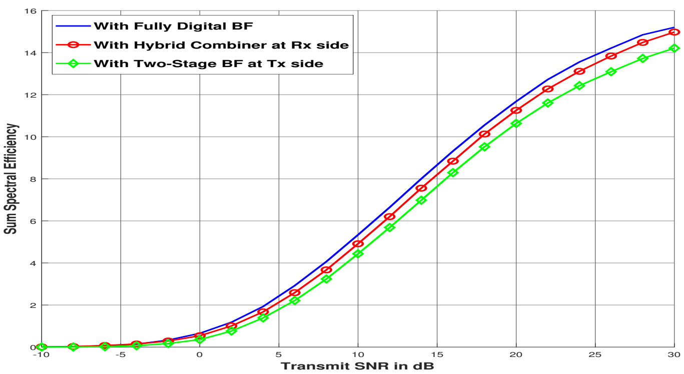

Fig. 3. Sum Rate comparisons for, OFDM, $N_s = 4, N_t^i = N_r^i = 8, M_t^i = M_r^i = 4, d_i = 1, \forall i, L = 4$ paths.

The dimensions of the two-stage BF and hybrid BF are such that the zero forcing capabilities at both sides are comparable. However, the number of LDR noises is the number of antennas at the Tx side, whereas for the analog Rx stage, the number of LDR noises is the number of analog BF outputs, which is less. We conjecture that the analog BF reduces the LDR noise to a significant level and this would explain the better performance of the analog stage at Rx (in both figures) compared to the two-stage architecture at Tx. In Figure 2, we compare against the eigen beamforming (where the left and right singular vectors of the corresponding channels are used as the Combiner/BF and fully digital) and shows that its performance is inferior compared to our proposed design.

VI. Conclusion

In this paper, we looked at beamforming solutions to null the SI power under a more practical noise model called as limited dynamic range. We proposed a multi-stage beamforming design (whose performance is validated through simulations), with a frequency flat analog or time domain combiner/BF stage and a frequency dependent baseband precoder/combiner. We decoupled the beamforming design for the Tx and Rx side. An iterative algorithm is obtained which jointly optimizes both analog/time domain and digital beamformers at the Tx/Rx side. We also discussed the dimensions of the BFs or combiners designed (e.g. the minimum number of RF chains required) such that the SI power can be mitigated fully at high SNR.

Acknowledgements

EURECOM's research is partially supported by its industrial members: ORANGE, BMW, Symantec, SAP, Monaco Telecom, iABG, and by the projects MASS-START (French FUI), DUPLEX (French ANR) and EU ITN project SPOTLIGHT.

References

[1] S. Li and R. D. Murch, "Full-duplex wireless communication using transmitter output based echo cancellation," in *IEEE GLOBECOM*, 2011.

[2] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, 2013.

[3] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. on Wire. Commun.*, vol. 11, no. 12, 2012.

[4] T. Riihonen and R. Wichman, "Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion," in *46th IEEE asilomar conference on signals, systems and computers (ASILOMAR)*, 2012.

[5] 3GPP Technical Report V14.1.0, "Study on scenarios and requirements for next generation access technologies," 2017.

[6] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. on Wire. Commun.*, vol. 14, no. 4, 2014.

[7] A. C. Cirik, R. Wang, , Y. Hua, and M. Latva-aho, "Weighted sum-rate maximization for full-fuplex MIMO interference channels," *IEEE Trans. on Commun.*, vol. 63, no. 3, Mar. 2015.

[8] P. Aquilina, A. C. Cirik, and T. Ratnarajah, "Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks," *IEEE Trans. on Commun.*, vol. 65, no. 4, Apr. 2017.

[9] X. Zhang, A. Molisch, and S. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. on Sig. Process.*, vol. 53, no. 11, 2005.

[10] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 3, March 2014.

[11] A. Alkhateeb, O. E. Ayach, G. Leus, and R. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. for Sel. Topics in Sig Process.*, vol. 8, no. 5, October 2014.

[12] K. Satyanarayana, M. El-Hajjar, , P.-H. Kuo, A. Mourad, and L. Hanzo, "Hybrid beamforming design for full-duplex millimeter wave communication," *IEEE Trans. on Veh. Techn.*, vol. 68, no. 2, Feb 2019.

[13] O. Taghizadeh, V. Radhakrishnan, A. C. Cirik, R. Mathar, and L. Lampe, "Hardware impairments aware transceiver design for bidirectional full-fuplex MIMO OFDM systems," *IEEE Trans. on Vehic. Tech.*, vol. 67, no. 8, Aug. 2018.

[14] A. Alkhateeb and R. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. on Commun.*, vol. 64, no. 5, May 2016.

[15] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-Duplex bidirectional MIMO: achievable rates under limited dynamic range," *IEEE Trans. on Sig. Process.*, vol. 60, no. 7, July 2012.

[16] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. on Wireless Commun.*, December 2008.

[17] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," in *URL http://www2.imm.dtu.dk/pubdb/p.php?3274.*, November 2011.

[18] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, "An iteratively weighted mmse approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. on Sig. Process.*, vol. 59, no. 9, September 2011.

[19] C. K. Thomas and D. Slock, "Mixed time scale weighted sum rate maximization for hybrid beamforming in multi-cell MU-MIMO systems," in *Globecom Wkshps.*, Singapore, December 2017.

[20] ——, "Hybrid beamforming design in multi-cell mu-mimo systems with per-rf or per-antenna power constraints," in *IEEE VTC Fall*, Chicago, USA, Aug. 2018.

[21] ——, "Deterministic annealing for hybrid beamforming design in multi-cell mu-mimo systems," in *Proc. IEEE SPAWC*, Kalamata, Greece, 2018.

[22] F. Negro, I. Ghauri, and D. T. M. Slock, "Deterministic annealing design and analysis of the noisy MIMO interference channel," in *Proc. IEEE Inf. Theo. and Applic. Workshop (ITA), San Diego, CA, USA*, 2011.