Hybrid Beamforming for Millimeter Wave Full-Duplex under Limited Receive Dynamic Range

Ian P. Roberts, *Student Member, IEEE,* Jeffrey G. Andrews, *Fellow, IEEE,*
and Sriram Vishwanath, *Senior Member, IEEE*

***Abstract*—Full-duplex millimeter wave (mmWave) communication has shown increasing promise for self-interference cancellation via hybrid precoding and combining. This paper proposes a novel mmWave multiple-input multiple-output (MIMO) design for configuring the analog and digital beamformers of a full-duplex transceiver. This work is the first to holistically consider the key practical constraints of analog beamforming codebooks, a minimal number of radio frequency (RF) chains, limited channel knowledge, beam alignment, and a limited receive dynamic range. To prevent self-interference from saturating receive components, such as LNAs and ADCs, a design framework is developed that limits the degree of self-interference on a per-antenna and per-RF chain basis. We present a means for constructing analog beamforming candidates from beam alignment measurements to afford our design greater flexibility in its aim to reduce self-interference. Numerical results evaluate the design in a variety of settings and validate the need to prevent receiver-side saturation. These results and corresponding insights serve as useful design references and benchmarks for practical full-duplex mmWave transceivers.**

***Index Terms*—millimeter wave, full-duplex, hybrid beamforming, self-interference, saturation, beam alignment.**

I. Introduction

THE ABILITY for a transceiver to transmit and receive simultaneously in-band introduces an exciting upgrade at the physical layer and in medium access when compared to existing half-duplex schemes such as time-division duplexing (TDD) and frequency-division duplexing (FDD) [1]. The gains supplied by full-duplex capability in millimeter wave (mmWave) systems are particularly attractive [2], [3], beyond the usual gains in spectral efficiency and latency. By full-duplexing access and backhaul, heterogeneous mmWave networks can be deployed with lower latency, higher spectral efficiency, and a reduced number of fiber drops. Key challenges in mmWave systems, such as beam alignment and beam tracking, could be transformed when devices can transmit and receive simultaneously, especially in highly dynamic applications like vehicular communication. The presence of communication, radar, and other incumbents in lightly regulated mmWave spectrum highlights the potential of novel strategies for medium access, in-band coexistence, and interference management via full-duplex.

A. Prior Work and Motivation

The majority of existing research on full-duplex has been in the context of lower carrier frequencies (e.g., sub-6 GHz). While many aspects of this existing work can be extended to mmWave, new approaches are necessary to enable full-duplex at mmWave [2], [3]. Dense antenna arrays and wide bandwidths are two key challenges for analog self-interference cancellation at mmWave in particular [3], [4]. Furthermore, directly translating multiple-input multiple-output (MIMO)-based self-interference mitigation (e.g., [5]–[7]) to mmWave is complicated by hybrid digital/analog beamforming, propagation characteristics at mmWave, and system-level factors like beam alignment. While passive and polarization-based approaches have been proposed for mmWave [8]–[10], they are difficult to generalize to dense mmWave antenna arrays.

A number of recent bodies of work investigate methods where self-interference is mitigated by appropriately configuring the transmit and receive beamformers at a mmWave full-duplex device, sometimes termed *beamforming cancellation* [11]–[20]. Resembling MIMO-based approaches from sub-6 GHz full-duplex literature, existing beamforming cancellation designs suggest that mmWave full-duplex is theoretically possible without the hardware and computational costs associated with analog and digital self-interference cancellation. However, existing beamforming designs for mmWave full-duplex often fail to account for critical practical transceiver-level and system-level considerations. To start, practical systems typically rely on codebook-based analog beamforming and beam alignment [4], meaning there is extremely limited freedom in the choice of analog beamformers. Designs such as those in [11]–[16] do not account for codebook-based analog beamforming and assume the ability to dynamically fine-tune each phase shifter in analog beamforming networks. Moreover, [11]–[16] assume infinite-precision phase shifters; in reality, phase shifters are almost certainly configured digitally, subjecting them to some degree of phase resolution. Almost all designs assume a lack of amplitude control in analog beamforming even though it is not uncommon to have both phase and amplitude control in practice. Those in [11]–[17] do not account for beam alignment and assume full over-

Manuscript received December 21, 2020; revised April 4, 2021; accepted May 21, 2021.

I. P. Roberts and J. G. Andrews are with the Wireless Networking and Communications Group at the University of Texas at Austin, Austin, TX, USA. S. Vishwanath is with GenXComm, Inc., Austin, TX, USA. Corresponding author: I. P. Roberts (ipr@utexas.edu).

I. P. Roberts is supported by the National Science Foundation Graduate Research Fellowship Program (Grant No. DGE-1610403). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

the-air channel knowledge, which is highly unlikely to have in practice.

Several designs [12]–[15] involve analog-only beamforming, meaning they only support single-stream communication, which simplifies the design of beamforming-based self-interference mitigation. This is especially true in [11]–[16] where the designs may be highly dependent on near-field self-interference channel conditions and are not shown to be robust against such. Some designs, such as those in [16]–[19], take advantage of an increased number of radio frequency (RF) chains that enable them to exploit the consequent dimensionality to mitigate self-interference in the digital domain. This is a strong assumption since the minimal number of RF chains necessary in hybrid beamforming is equal to the number of streams; increasing beyond this is undesirable in terms of financial cost and power consumption.

Finally, and perhaps most pertinent to this work, the majority of existing designs neglect the limited dynamic range of practical receivers [21], [22]. This is particularly important for full-duplex transceivers since self-interference—which is likely many orders of magnitude stronger than a desired receive signal—can overwhelm and saturate components of a receive chain if not sufficiently mitigated (hence, a limited dynamic range) [21], [23]. The work in [12]–[16] accounts for analog-to-digital converter (ADC) saturation by *completely* mitigating self-interference beforehand—which is not always possible—but do not account for other sources of saturation such as low noise amplifiers (LNAs). In [18], the need to prevent ADC saturation is mentioned but is assumed to be satisfied without any mathematical basis. In [11], [17], [19], the need to prevent receiver-side saturation is ignored.

B. Contributions

In this work, we formulate mmWave MIMO expressions capturing practical receive dynamic range limitations *per-antenna* and *per-RF chain*. In particular, we motivate this work by the limited dynamic range of LNAs placed per-antenna and of ADCs placed per-RF chain, though it can be generalized to arbitrary dynamic range considerations. Using these formulations, we outline constraints to limit the self-interference power inflicted on each antenna and each RF chain at the receiver of the full-duplex device to prevent it from saturating. We outline conditions where meeting these per-antenna and per-RF chain self-interference power constraints are implicitly met, either by one another or other system factors.

Incorporating these constraints, we present a hybrid beamforming design that enables a mmWave transceiver to operate in a full-duplex fashion, serving two devices simultaneously in-band. Adhering to a multitude of practical considerations beyond a limited receive dynamic range, our design supports beam alignment schemes and codebook-based analog beamforming, rather than impractically assuming full knowledge of over-the-air channels and the ability to dynamically fine-tune phase shifters and attenuators. Furthermore, we limit the number of RF chains to the minimum necessary for multi-stream communication. To provide our design with freedom in the choice of its analog beamformers, we present a methodology for building sets of *candidate* analog beamformers based on measurements from codebook-based beam alignment. Our design is not self-interference channel model-dependent in that it does not exploit any particular structure.

Numerical results indicate scenarios where our design thrives, offering significant spectral efficiency gains over conventional half-duplex operation. These results also outline conditions under which per-antenna and per-RF chain self-interference power constraints restrict what is possible for mmWave full-duplex, providing useful insights to engineers on relationships between system parameters such as transmit power, RF isolation, ADC resolution, and the self-interference power reaching each antenna and each RF chain. Understanding the degree of self-interference mitigation required at specific points in the receiver can drive full-duplex system analyses, including those that may supplement beamforming-based approaches with analog and/or digital self-interference cancellation [18]. The goal of this work is not necessarily to produce a deployment-ready approach to mmWave full-duplex but rather to present a design methodology motivated by a select variety of practical considerations. We hope our approach and numerical results act as a step toward even more practically sound solutions and as a benchmark for systems constrained by a limited receive dynamic range.

II. System Model

This work considers the wireless system in Fig. 1, where a mmWave transceiver i aims to transmit to a device j while receiving from a device k in the same band. Instead of turning to half-duplexing strategies like TDD or FDD to avoid self-interference, this work presents a design that enables in-band full-duplex operation by leveraging the spatial domain to mitigate self-interference. In this work, we consider the case where devices j and k are separate half-duplex devices, though many aspects of our contribution would extend naturally, or even simplify, when j and k are separate full-duplex devices or comprise a single full-duplex device.

Ubiquitous among practical mmWave transceivers is the use of hybrid digital/analog beamforming architectures where transmit precoding and receive combining are implemented by the combination of digital (baseband) and analog (RF) signal processing, as exhibited in Fig. 1. We assume devices i, j, and k all employ hybrid beamforming in a fully-connected fashion where each antenna is connected to each RF chain via an analog beamforming network [4]. As illustrated in Fig. 1, we assume that separate arrays are used at device i for transmission and reception and independent precoding and combining on the two is supported [3].

For devices $(m,n) \in \{(i,j),(k,i)\}$, we use the following notation. Let $N_{\mathrm{t}}^{(m)}$ and $N_{\mathrm{r}}^{(n)}$ be the number of transmit and receive antennas, respectively. Connecting the digital and analog stages, let $L_{\mathrm{t}}^{(m)}$ and $L_{\mathrm{r}}^{(n)}$ be the number of transmit and receive RF chains, respectively. Let $N_{\mathrm{s}}^{(mn)}$ be the number of symbol streams transmitted from device m to device n. Let $\mathbf{F}_{\mathrm{BB}}^{(m)} \in \mathbb{C}^{L_{\mathrm{t}}^{(m)} \times N_{\mathrm{s}}^{(mn)}}$ be the digital precoding matrix and

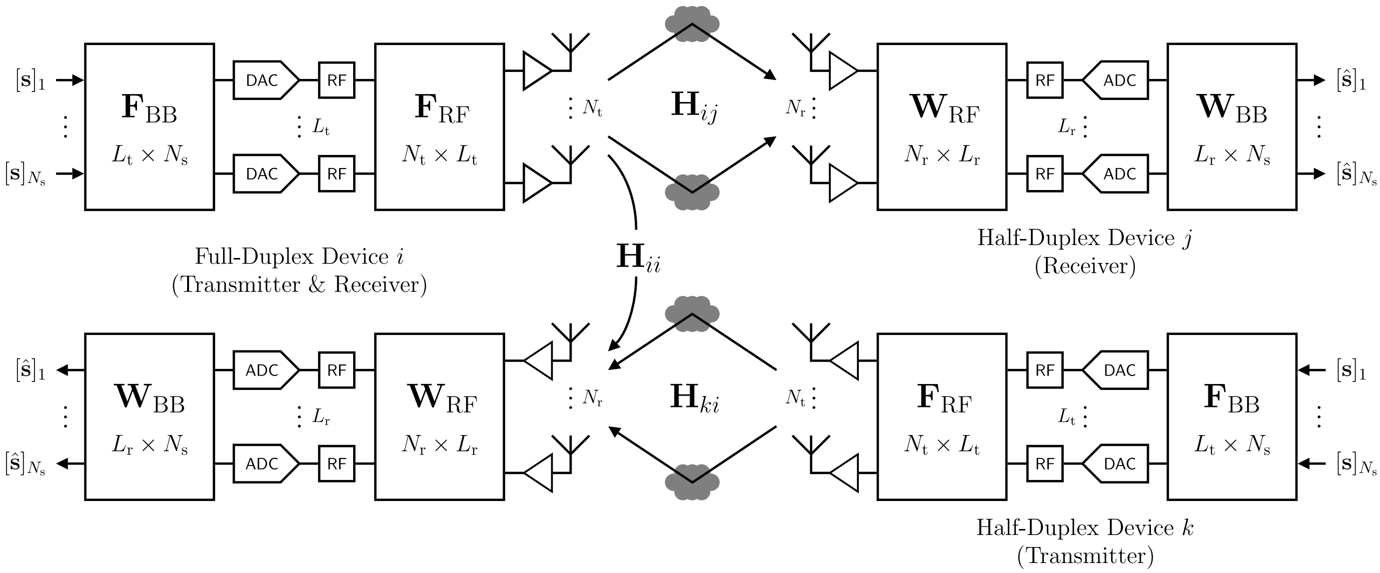


Fig. 1. A full-duplex mmWave device i transmitting to a device j as it receives from a device k in-band.

$\mathbf{F}_{\mathrm{RF}}^{(m)} \in \mathbb{C}^{N_{\mathrm{t}}^{(m)} \times L_{\mathrm{t}}^{(m)}}$ be the analog precoding matrix, responsible for transmitting from m. Let $\mathbf{W}_{\mathrm{BB}}^{(n)} \in \mathbb{C}^{L_{\mathrm{r}}^{(n)} \times N_{\mathrm{s}}^{(mn)}}$ be the digital combining matrix and $\mathbf{W}_{\mathrm{RF}}^{(n)} \in \mathbb{C}^{N_{\mathrm{r}}^{(n)} \times L_{\mathrm{r}}^{(n)}}$ be the analog combining matrix, responsible for receiving at n.

For devices $(m,n) \in \{(i,j),(k,i)\}$, let $\mathbf{s}^{(m)} \in \mathbb{C}^{N_{\mathrm{s}}^{(mn)} \times 1}$ be the symbol vector transmitted by m intended for device n, where the symbol covariance is

$$\mathbb{E}\left[\mathbf{s}^{(m)} \mathbf{s}^{(m)*}\right] = \frac{1}{N_{\mathrm{s}}^{(mn)}} \mathbf{I}. \tag{1}$$

We do not consider a specific signaling constellation, though we will evaluate our work assuming Gaussian signaling is employed. Let $\mathbf{n}^{(n)} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \sigma_{\mathrm{n}}^2 \cdot \mathbf{I}\right)$ be the $N_{\mathrm{r}}^{(n)} \times 1$ additive noise vector incurred at the receive array of n, where σ_{n}^2 represents a per-antenna noise power spectral density in watts/Hz and is assumed common across devices for simplicity. We denote the symbol period as T and the symbol bandwidth $B = T^{-1}$, which we assume to be constant across both links, being a full-duplex system. Let $\tilde{P}_{\mathrm{tx}}^{(m)}$ be the total transmit power of device m in watts and $P_{\mathrm{tx}}^{(m)}$ be the resulting transmit power in joules per symbol. We extend this convention, representing power quantities in watts using a tilde, $\tilde{P}$, and in joules per symbol without a tilde, P, which are linked via $P = \tilde{P} \cdot B^{-1} = \tilde{P} \cdot T$.

We impose the following digital precoding power constraint

$$\left\|\mathbf{F}_{\mathrm{BB}}^{(m)}\right\|_{\mathrm{F}}^2 \leq 1 \tag{2}$$

and normalize the columns of $\mathbf{F}_{\mathrm{RF}}^{(m)}$ to have squared ℓ_2-norm $L_{\mathrm{t}}^{(m)}$ and of $\mathbf{W}_{\mathrm{RF}}^{(n)}$ to have squared ℓ_2-norm $N_{\mathrm{r}}^{(n)}$. Since this work is focused on receiver-side power levels, our analog combining normalization differs from the analog precoding one to ensure consistency with the physical combining taking place at receivers. As is common in practice, we will assume that columns of our analog precoders and combiners will come from analog beamforming codebooks that account for hardware constraints such as phase shifter resolution and amplitude control. That is, for $(m,n) \in \{(i,j),(k,i)\}$, we have

$$\left[\mathbf{F}_{\mathrm{RF}}^{(m)}\right]_{:,\ell} \in \mathcal{F}_{\mathrm{RF}}^{(m)}, \ \ell = 1, \ldots, L_{\mathrm{t}}^{(m)} \tag{3}$$

$$\left[\mathbf{W}_{\mathrm{RF}}^{(n)}\right]_{:,\ell} \in \mathcal{W}_{\mathrm{RF}}^{(n)}, \ \ell = 1, \ldots, L_{\mathrm{r}}^{(n)} \tag{4}$$

where $\mathcal{F}_{\mathrm{RF}}^{(m)}$ and $\mathcal{W}_{\mathrm{RF}}^{(n)}$ denote analog precoding and combining codebooks, respectively.

Now, let us consider $(m,n) \in \{(i,j),(k,i),(i,i)\}$. We assume that the large-scale power gain between devices m and n is given by G_{mn}^2. The $N_{\mathrm{r}}^{(n)} \times N_{\mathrm{t}}^{(m)}$ channel matrix between a transmitter m and receiver n is denoted $\mathbf{H}_{mn} \in \mathbb{C}^{N_{\mathrm{r}}^{(n)} \times N_{\mathrm{t}}^{(m)}}$. In this work, we consider the more straightforward case of frequency-flat MIMO channels and will address frequency-selective ones in future work. Taking the perspective of our full-duplex device i, we term $\mathbf{H}_{ij}$ the *transmit* channel, $\mathbf{H}_{ki}$ the *receive* channel, and $\mathbf{H}_{ii}$ the *self-interference* channel. Note that we have not considered an inter-user interference channel between devices k and j since we assume that, with sufficient separation and/or user scheduling, the interference between the two to be negligible given the high path loss at mmWave and highly directional steering of energy that is typical; investigating the impacts of inter-user interference would be interesting future work.

We assume devices i and j as well as devices k and i are separated in a far-field fashion. As such, we model the transmit and receive channels as the composition of discrete rays with the Saleh-Valenzuela-based model [4]. Explicitly, channels $\mathbf{H}_{ij}$ and $\mathbf{H}_{ki}$ are of the form

$$\mathbf{H}_{mn} = \sqrt{\frac{1}{N_{\mathrm{rays}}^{(mn)}}} \sum_{u=1}^{N_{\mathrm{rays}}^{(mn)}} \beta_u \ \mathbf{a}_{\mathrm{rx}}^{(n)}(\mathrm{AoA}_u) \ \mathbf{a}_{\mathrm{tx}}^{(m)}(\mathrm{AoD}_u)^*. \tag{5}$$

where $(m,n) \in \{(i,j),(k,i)\}$. In each channel, $N_{\mathrm{rays}}^{(mn)}$ is a random variable dictating the number of rays in the channel.

The complex gain of ray u is given as $\beta_u \sim \mathcal{N}_{\mathbb{C}}(0,1)$. The u-th ray's angle of departure (AoD) and angle of arrival (AoA) are given as AoD_u and AoA_u, respectively. The transmit and receive array response vectors at these angles are given as $\mathbf{a}_{\mathrm{tx}}^{(m)}(\mathrm{AoD}_u)$ and $\mathbf{a}_{\mathrm{rx}}^{(n)}(\mathrm{AoA}_u)$, which have squared ℓ_2-norm $N_{\mathrm{t}}^{(m)}$ and $N_{\mathrm{r}}^{(n)}$, respectively. The coefficient in front of the summations handles a channel power normalization to ensure $\mathbb{E}\left[\|\mathbf{H}_{mn}\|_{\mathrm{F}}^2\right] = N_{\mathrm{t}}^{(m)} N_{\mathrm{r}}^{(n)}$.

A lack of measurements and characterizations of mmWave self-interference prevents us from confidently assuming a particular channel model for $\mathbf{H}_{ii}$ [3]. As such, our contribution herein does not rely on the self-interference channel's structure or properties. However, to evaluate our design, we employ a model that aims to capture the near-field nature of the transmit and receive arrays at i along with reflections that may stem from the environment [2], [11], which we explicitly state in Section VI. The large-scale power gain of the self-interference channel is represented by G_{ii}^2, which captures the RF isolation between the transmit and receive arrays at i. We define the signal-to-noise ratio (SNR) between two devices $(m,n) \in \{(i,j),(k,i)\}$ as

$$\mathrm{SNR}_{mn} \triangleq \frac{P_{\mathrm{tx}}^{(m)} G_{mn}^2}{\sigma_{\mathrm{n}}^2} = \frac{\tilde{P}_{\mathrm{tx}}^{(m)} G_{mn}^2}{\sigma_{\mathrm{n}}^2 \cdot B} \tag{6}$$

which captures the received power (without beamforming gains) versus the noise power.

III. Problem Formulation

This work is motivated by the fact that a receive chain of a full-duplex device—which practically has a limited dynamic range—is susceptible to saturation due to the overwhelming strength of self-interference [21], [22]. To highlight this, we consider two sources of limited receive dynamic range in this work: LNAs and ADCs. Like other amplifiers, LNAs begin saturating and introduce significant nonlinearities beyond a certain input power level, meaning only signals below some input power threshold see an approximately linear amplifier. The limited resolution of an ADC is a classical example of limited dynamic range; since the combination of a desired receive signal and self-interference enters the ADC, self-interference can drive up quantization noise and degrade the quality of the desired receive signal, even when digital cancellation is used [21]. If the LNAs, ADCs, and the rest of the receive chain had an infinite dynamic range and were fully linear, self-interference could be removed solely in the digital domain. This is not the case practically, meaning a portion of self-interference should be mitigated by our system *via beamforming* before it can saturate the receiver and destroy any sense of linearity. If beamforming sufficiently mitigates self-interference to levels that prevent saturation, residual self-interference may be cancelled digitally thanks to a well-preserved receive chain.

Consider Fig. 1, where LNAs are placed per-antenna and ADCs are placed per-RF chain. To avoid saturating the LNAs at the full-duplex device, the self-interference power reaching each antenna must be mitigated to below some threshold. Similarly, to avoid saturating the ADCs, the self-interference reaching each RF chain must also be limited. This work investigates relying solely on beamforming to achieve this. Note that we are only concerned with preventing saturation at the receiver of the *full-duplex* device i—not that of the *half-duplex* device j—since the saturation we are considering stems from self-interference. There may exist other motivations (beyond LNAs and ADCs) for restricting the self-interference power at each antenna and at each RF chain. With this in mind, the design we present is not strictly for LNAs and ADCs but rather for meeting arbitrary per-antenna and per-RF chain self-interference power constraints; LNAs and ADCs are an important special case.

Using the previously defined system model, we can begin analyzing the signals that reach the LNAs and ADCs of our full-duplex device. The symbol vector at the input of the LNAs of i is

$$\mathbf{y}_{\mathrm{LNA}} = \mathbf{y}_{\mathrm{des,LNA}} + \mathbf{y}_{\mathrm{int,LNA}} + \mathbf{y}_{\mathrm{noise,LNA}} \in \mathbb{C}^{N_{\mathrm{r}}^{(i)} \times 1} \tag{7}$$

where the desired term is $\mathbf{y}_{\mathrm{des,LNA}} = \sqrt{P_{\mathrm{tx}}^{(k)}} G_{ki} \mathbf{H}_{ki} \mathbf{F}_{\mathrm{RF}}^{(k)} \mathbf{F}_{\mathrm{BB}}^{(k)} \mathbf{s}^{(k)}$, the self-interference term is $\mathbf{y}_{\mathrm{int,LNA}} = \sqrt{P_{\mathrm{tx}}^{(i)}} G_{ii} \mathbf{H}_{ii} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \mathbf{s}^{(i)}$, and the noise term is $\mathbf{y}_{\mathrm{noise,LNA}} = \mathbf{n}^{(i)}$. The signal at each antenna passes through its respective LNA, which we represent as $\mathcal{L}(\cdot)$ and model as $\mathcal{L}(x) = G_{\mathrm{rx}} \cdot x$ for $|x|^2 \leq P_{\mathrm{LNA}}^{\max}$, where the per-antenna symbol x undergoes a power gain of $G_{\mathrm{rx}}^2 > 0$ if its average power (over the symbol period) is below some threshold $P_{\mathrm{LNA}}^{\max}$. Otherwise, for $|x|^2 > P_{\mathrm{LNA}}^{\max}$, the LNA is saturated and x undergoes some nonlinear operation, often characterized by an established amplifier signal model (e.g., $P_{\mathrm{LNA}}^{\max}$ may be approximated by the P1dB point).

It is difficult to translate the effects of LNA saturation to the symbol level (the scope of this work) since it is inflicted instantaneously on time-domain signals. For this reason, we make no attempt to characterize LNA saturation with the understanding that linear LNA operation can be ensured by restricting the power of x up to $P_{\mathrm{LNA}}^{\max}$. We would like to point out that for a properly chosen $P_{\mathrm{LNA}}^{\max}$—which will likely include appropriate backoffs for the signal distribution and pulse shape—the *time-domain signal* will undergo linear amplification and, thus, so will the symbols. Since the gain of the LNA acts on signal-plus-noise, for conciseness, it can be abstracted out henceforth as $G_{\mathrm{rx}} = 1$ with appropriate scaling of the noise variance σ_{n}^2, which can simultaneously capture the noise figure of the LNA.

We overload the LNA transfer function $\mathcal{L}(\cdot)$ function to support vector input by the simple element-wise extension $[\mathcal{L}(\mathbf{x})]_\ell = \mathcal{L}(x_\ell)$, where $\mathbf{x} = [x_1, x_2, \ldots]^{\mathrm{T}}$. Following per-antenna LNAs, the signals are combined as

$$\mathbf{y}_{\mathrm{ADC}} = \mathbf{W}_{\mathrm{RF}}^{(i)*} \times \mathcal{L}(\mathbf{y}_{\mathrm{LNA}}) \in \mathbb{C}^{L_{\mathrm{r}}^{(i)} \times 1} \tag{8}$$

where $\mathbf{y}_{\mathrm{ADC}}$ is the vector of per-RF chain symbols at the input of the ADCs. Under linear LNA operation, (8) can be written as

$$\mathbf{y}_{\mathrm{ADC}} \stackrel{\mathrm{lin}}{=} \mathbf{W}_{\mathrm{RF}}^{(i)*} \times \mathbf{y}_{\mathrm{LNA}} \tag{9}$$

$$= \mathbf{y}_{\mathrm{des,ADC}} + \mathbf{y}_{\mathrm{int,ADC}} + \mathbf{y}_{\mathrm{noise,ADC}} \tag{10}$$

where $\mathbf{y}_{(\cdot),\mathrm{ADC}} = \mathbf{W}_{\mathrm{RF}}^{(i)*} \times \mathbf{y}_{(\cdot),\mathrm{LNA}}$. We use $\mathcal{Q}(\cdot)$ to represent a b-bit ADC, modeled as

$$\mathcal{Q}(x) = x + e_{\mathrm{quant}} \tag{11}$$

where x is the symbol reaching the ADC and e_{quant} is the error in perfectly digitizing x due to quantization noise, whose power under b-bit, uniform quantization can be approximated as [22]

$$\mathbb{E}\left[\left|e_{\mathrm{quant}}\right|^2\right] \approx \frac{|x|^2}{1.5 \cdot 2^{2b}}. \tag{12}$$

As the number of bits b increases, the magnitude of e_{quant} decreases. Similarly, as the magnitude of x increases, the quantization noise power increases. From this, one can see why mitigating self-interference before the ADC input is so important: increased self-interference can plague a desired receive signal with increased quantization noise. We overload $\mathcal{Q}(\cdot)$ to vectors as $[\mathcal{Q}(\mathbf{x})]_\ell = \mathcal{Q}(x_\ell)$, where $\mathbf{x} = [x_1, x_2, \dots]^{\mathrm{T}}$. The symbol vector out of the ADCs is

$$\mathbf{y}_{\mathrm{dig}} = \mathcal{Q}(\mathbf{y}_{\mathrm{ADC}}) \in \mathbb{C}^{L_{\mathrm{r}}^{(i)} \times 1} \tag{13}$$

$$= \mathbf{y}_{\mathrm{des,ADC}} + \mathbf{y}_{\mathrm{int,ADC}} + \mathbf{y}_{\mathrm{noise,ADC}} + \mathbf{e}_{\mathrm{quant}}. \tag{14}$$

From this discussion, we can see that limiting the power into each of the LNAs and each of the ADCs is critical in preserving the linearity of the receive chain and reducing the effects of quantization. With these models and formulations in hand, we begin laying out our contribution, which aims to mitigate self-interference per-antenna and per-RF chain.

IV. Beam Candidate Acquisition

Practical mmWave systems employ *beam alignment* schemes that aim to identify transmit-receive analog beamforming pairs that afford link margin sufficient for communication. Typically, these analog beamformers are chosen from a predetermined codebook of beams, which reduces complexity and offers robustness. Given that our full-duplex system entertains two independent links, beam alignment needs to be executed on the transmit link and on the receive link. In this section, we present a methodology for creating a set of beam *candidates* for transmission and reception on each link, offering our design in the next section greater freedom in its aim to reduce self-interference. This is motivated by the fact that some beam selections naturally afford more isolation at the full-duplex device than other pairs, meaning it may be preferable to use them for full-duplex, even if they are sub-optimal in a half-duplex sense.

We have not assumed knowledge of the over-the-air channels $\mathbf{H}_{ij}$ and $\mathbf{H}_{ki}$. Rather, we will be inspecting them using beam alignment, which we conduct in a half-duplex fashion on separate time-frequency resources on each link, meaning we avoid self-interference and any concern for saturation. For our design, we propose that candidate analog beamforming matrices be created as follows, which can be accommodated by existing beam alignment schemes. Let $\mathbf{F}_{\mathrm{tr}}^{(i)} \in \mathbb{C}^{N_{\mathrm{t}}^{(i)} \times M_{\mathrm{t}}^{(i)}}$ be a matrix whose $M_{\mathrm{t}}^{(i)}$ columns are *training* analog precoders used by i during candidate beam acquisition as it illuminates $\mathbf{H}_{ij}$. To observe these illuminations, let $\mathbf{W}_{\mathrm{tr}}^{(j)} \in \mathbb{C}^{N_{\mathrm{r}}^{(j)} \times M_{\mathrm{r}}^{(j)}}$ be a matrix whose $M_{\mathrm{r}}^{(j)}$ columns are *training* analog combiners used by j. We assume, for simplicity, that each of the $M_{\mathrm{t}}^{(i)}$ training precoders is observed by all $M_{\mathrm{r}}^{(j)}$ training combiners, though the ideas herein could be adapted when this is not the case. Analogously, let $\mathbf{F}_{\mathrm{tr}}^{(k)} \in \mathbb{C}^{N_{\mathrm{t}}^{(k)} \times M_{\mathrm{t}}^{(k)}}$ and $\mathbf{W}_{\mathrm{tr}}^{(i)} \in \mathbb{C}^{N_{\mathrm{r}}^{(i)} \times M_{\mathrm{r}}^{(i)}}$ be the training analog precoders and analog combiners used by k and i, respectively, used to inspect $\mathbf{H}_{ki}$. On both links, we assume the analog beamformers used during training come from their respective codebooks according to (3)–(4), though we do not suggest that it is necessary to sweep *all* beams (more beams will certainly help) nor do we specify particular beam codebooks. Instead, we present a methodology that can be tailored and applied to a variety of beam alignment schemes and codebooks. The collection of measurements for each link can be written in matrix form as

$$\mathbf{M}_{ij} = \sqrt{P_{\mathrm{tx}}^{(i)}} G_{ij} \mathbf{W}_{\mathrm{tr}}^{(j)*} \mathbf{H}_{ij} \mathbf{F}_{\mathrm{tr}}^{(i)} \mathbf{I}_{M_{\mathrm{t}}^{(i)}} \in \mathbb{C}^{M_{\mathrm{r}}^{(j)} \times M_{\mathrm{t}}^{(i)}} \tag{15}$$

$$\mathbf{M}_{ki} = \sqrt{P_{\mathrm{tx}}^{(k)}} G_{ki} \mathbf{W}_{\mathrm{tr}}^{(i)*} \mathbf{H}_{ki} \mathbf{F}_{\mathrm{tr}}^{(k)} \mathbf{I}_{M_{\mathrm{t}}^{(k)}} \in \mathbb{C}^{M_{\mathrm{r}}^{(i)} \times M_{\mathrm{t}}^{(k)}}. \tag{16}$$

Under channel reciprocity, $\mathbf{M}_{ij}$ may be measured in reverse from j to i to avoid feedback overhead. For consistency, we maintain notation as if measurements take place from i to j.

Given that the training precoders and combiners come from their respective codebooks, the strength of the measurements in $\mathbf{M}_{ij}$ and $\mathbf{M}_{ki}$ directly indicates which analog precoders (columns) and analog combiners (rows) are promising candidates on each link. Having multiple (say $L \geq 1$) RF chains, building a set of K candidates requires choosing K tuples of L strong beam pairs from the measurement matrix $\mathbf{M}$. In the particular case when $L = 1$ (i.e., analog-only beamforming), this process amounts to simply choosing the K strongest beam pairs from $\mathbf{M}$.

Let $\mathcal{T}_{ij}$ be a set of K_{ij} *candidate* analog precoding-combining pairs for communicating from i to j. Similarly, let $\mathcal{T}_{ki}$ be a set of K_{ki} *candidate* analog precoding-combining pairs for communicating from k to i.

$$\mathcal{T}_{ij} = \left\{ \left(\mathbf{F}_{\mathrm{RF}}^{(i)}, \mathbf{W}_{\mathrm{RF}}^{(j)} \right)_n : n = 1, \dots, K_{ij} \right\} \tag{17}$$

$$\mathcal{T}_{ki} = \left\{ \left(\mathbf{F}_{\mathrm{RF}}^{(k)}, \mathbf{W}_{\mathrm{RF}}^{(i)} \right)_n : n = 1, \dots, K_{ki} \right\} \tag{18}$$

Our goal is to form the candidate sets $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ with promising beamforming pairs for each link. We describe our method for doing so—which can be replicated for $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ independently—by using generic notation (e.g., $\mathcal{T}$, K, $\mathbf{M}$) according to the summary shown in Algorithm 1.

We begin by finding the indices of the training precoders $\mathcal{M}_{\mathrm{tx}}$ and the training combiners $\mathcal{M}_{\mathrm{rx}}$ that revealed the top K strongest measurements in $\mathbf{M}$, sorted according to descending strength. We initialize $\mathcal{T}$ to an empty set. The first beam (column) of the n-th candidate is steered along the n-th strongest entry in $\mathbf{M}$. To ensure that we do not transmit or receive along directions more than once, we keep track of each beam's transmit and receive indices (i.e., t and r) in sets $\mathcal{J}_{\mathrm{tx}}$ and $\mathcal{J}_{\mathrm{rx}}$, respectively. Then, to choose the next column, we locate the strongest entry in $\mathbf{M}$ whose transmit or receive beam has not already been selected for this n-th candidate. When the

Algorithm 1 Beam candidate acquisition algorithm.

Input: $\mathbf{M}$, $\mathbf{F}_{\rm tr}$, $\mathbf{W}_{\rm tr}$, L, K
  $\mathcal{T} = \{\varnothing\}$
  $\mathcal{M}_{\rm rx}, \mathcal{M}_{\rm tx} = \arg\max{\rm k}\left(\mathrm{abs}\left(\mathbf{M}\right), K\right)$
  for $n = 1 : K$ **do**
    $t = \left[\mathcal{M}_{\rm tx}\right]_n$
    $r = \left[\mathcal{M}_{\rm rx}\right]_n$
    $\mathbf{F}_{\rm RF} = \left[\mathbf{F}_{\rm tr}\right]_{:,t}$
    $\mathbf{W}_{\rm RF} = \left[\mathbf{W}_{\rm tr}\right]_{:,r}$
    $\mathcal{J}_{\rm tx} = \{t\}$
    $\mathcal{J}_{\rm rx} = \{r\}$
    for $\ell = 1 : L-1$ **do**
      $r, t = \arg\max_{r,t} \left|\left[\mathbf{M}\right]_{r,t}\right| \ \text{s.t.}\ t \notin \mathcal{J}_{\rm tx},\ r \notin \mathcal{J}_{\rm rx}$
      $\mathbf{F}_{\rm RF} = \left[\mathbf{F}_{\rm RF} \quad \left[\mathbf{F}_{\rm tr}\right]_{:,t}\right]$
      $\mathbf{W}_{\rm RF} = \left[\mathbf{W}_{\rm RF} \quad \left[\mathbf{W}_{\rm tr}\right]_{:,r}\right]$
      $\mathcal{J}_{\rm tx} = \mathcal{J}_{\rm tx} \cup t$
      $\mathcal{J}_{\rm rx} = \mathcal{J}_{\rm rx} \cup r$
    end for
    $\mathcal{T} = \mathcal{T} \cup \left(\mathbf{F}_{\rm RF}, \mathbf{W}_{\rm RF}\right)$
  end for
Output: $\mathcal{T}$

columns of $\mathbf{F}_{\rm tr}$ and $\mathbf{W}_{\rm tr}$ are linearly independent, this ensures that each analog precoding candidate and analog combining candidate are rank-L, which is necessary for multiplexing up to L streams. Note that, in this method, we have assumed that $L = L_{\rm t} = L_{\rm r} = N_{\rm s}$ for a given link, which is more practical than supplying devices with more RF chains than streams. This process is repeated until all L columns are populated, and then the candidate pair is appended to our candidate set $\mathcal{T}$. Once all K candidates have been generated, the set $\mathcal{T}$ is returned.

To potentially reduce overhead, modifications to our proposed method can accommodate creative approaches to beam alignment, such as hierarchical beam search, compressed sensing, and partial codebook search. With slight modifications, our method could also accommodate cases where training measurements map to codebook candidates rather than measuring the candidates themselves. Additionally, we point out that L^2 beam pairs on a given link may be measured simultaneously with L RF chains. In general, to build promising candidates, it is critical that the number of training beams ($M_{\rm t}$ and $M_{\rm r}$) be sufficiently large to locate at least N strong rays in each channel such that $K \leq \binom{N}{L}$.

Following the construction of $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ using this method, we now suggest that the following channel estimates be made, having not assumed knowledge of $\mathbf{H}_{ij}$ or $\mathbf{H}_{ki}$. To provide the design presented in the next section with channel information, we assume a set $\mathcal{H}_{ij}$ has been populated with estimates of the effective channel seen by each of the candidates in $\mathcal{T}_{ij}$ described as

$$\mathcal{H}_{ij} = \left\{\mathbf{W}_{\rm RF}^{(j)*}\mathbf{H}_{ij}\mathbf{F}_{\rm RF}^{(i)} : \left(\mathbf{F}_{\rm RF}^{(i)}, \mathbf{W}_{\rm RF}^{(j)}\right) \in \mathcal{T}_{ij}\right\} \tag{19}$$

where $\left|\mathcal{H}_{ij}\right| = K_{ij}$. Note that each effective channel in $\mathcal{H}_{ij}$ is merely $L_{\rm r}^{(j)} \times L_{\rm t}^{(i)}$, a very small size relative to $\mathbf{H}_{ij}$, and is observed digitally. Also, under under channel reciprocity, measurements in $\mathcal{H}_{ij}$ can be conducted from j to i to reduce overhead since only one of its entries needs to be fed back to j following the design, which takes place at i. By these accounts, it is our hope that the overhead associated with collecting the measurements in $\mathcal{H}_{ij}$ not be prohibitive. Furthermore, we would like to point out that our design does not require executing this estimation on the link from k to i. This concludes beam candidate acquisition, having populated $\mathcal{T}_{ij}$, $\mathcal{T}_{ki}$, and $\mathcal{H}_{ij}$, which will enable the design presented in the next section.

V. Hybrid Beamforming Design for mmWave Full-Duplex

In this section, we present a hybrid beamforming design to enable mmWave full-duplex while accounting for per-antenna and per-RF chain power constraints at the receiver of a full-duplex device. The goal of our design is to achieve a high spectral efficiency on both links while ensuring that the power of self-interference reaching the full-duplex device's receiver is below some thresholds. To do so, our design leverages the analog beamforming candidate sets $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ found in the previous section to configure the analog beamformers at each device. The design that follows holds for general per-antenna and per-RF chain power constraints, even though we are considering LNA and ADC power constraints as particular motivators.

Our design supports spatial multiplexing multiple streams and importantly does not require more RF chains than necessary, where $L_{\rm t}^{(i)} = L_{\rm r}^{(j)} = N_{\rm s}^{(ij)}$ and $L_{\rm t}^{(k)} = L_{\rm r}^{(i)} = N_{\rm s}^{(ki)}$. Furthermore, we support the important case where $N_{\rm s}^{(ij)} = N_{\rm s}^{(ki)}$, implying $L_{\rm t}^{(i)} = L_{\rm r}^{(i)}$. This, along with the fact that we have not assumed anything about $\mathbf{H}_{ii}$, means our design's formulation does not rely on *completely* avoiding (i.e., zero-forcing) the over-the-air self-interference channel or even the *effective* self-interference channel, unlike several existing approaches. We assume we have perfect channel knowledge of $\mathbf{H}_{ii}$ but do not assume knowledge of $\mathbf{H}_{ij}$ or $\mathbf{H}_{ki}$. We motivate this assumption by presuming that the self-interference channel can be more reliably estimated given its strength and can be done so possibly through regular calibration. This work aims to provide an approximate "best-case" design under select practical considerations, chiefly a limited receive dynamic range, motivating us to ignore the impact of channel estimation error in this work; it would be valuable future work to investigate its impact. We assume large-scale quantities (e.g., transmit powers, large-scale channel gains, SNRs) are known.

A. Expressing Per-Antenna and Per-RF Chain Received Power Constraints

Let $\tilde{P}_{\rm SI,LNA}^{\rm max}$ and $\tilde{P}_{\rm SI,ADC}^{\rm max}$ be the maximum average self-interference power (in watts) over the symbol period allowed at each LNA and each ADC of the receiver of the full-duplex device i, respectively. Referring to the terms presented in our system model and problem formulation, let us form our LNA and ADC constraints using $\tilde{P}_{\rm SI,LNA}^{\rm max}$ and $\tilde{P}_{\rm SI,ADC}^{\rm max}$. A constraint

$$\mathcal{R}_{ij} = \log_2 \left| \mathbf{I} + \frac{\mathrm{SNR}_{ij}}{N_\mathrm{s}^{(ij)}} \mathbf{W}_\mathrm{BB}^{(j)*} \mathbf{W}_\mathrm{RF}^{(j)*} \mathbf{H}_{ij} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)*} \mathbf{F}_\mathrm{RF}^{(i)*} \mathbf{H}_{ij}^{*} \mathbf{W}_\mathrm{RF}^{(j)} \mathbf{W}_\mathrm{BB}^{(j)} \left(\mathbf{Q}_\mathrm{n}^{(j)} \right)^{-1} \right| \tag{27}$$

$$\mathcal{R}_{ki} = \log_2 \left| \mathbf{I} + \frac{\mathrm{SNR}_{ki}}{N_\mathrm{s}^{(ki)}} \mathbf{W}_\mathrm{BB}^{(i)*} \mathbf{W}_\mathrm{RF}^{(i)*} \mathbf{H}_{ki} \mathbf{F}_\mathrm{RF}^{(k)} \mathbf{F}_\mathrm{BB}^{(k)} \mathbf{F}_\mathrm{BB}^{(k)*} \mathbf{F}_\mathrm{RF}^{(k)*} \mathbf{H}_{ki}^{*} \mathbf{W}_\mathrm{RF}^{(i)} \mathbf{W}_\mathrm{BB}^{(i)} \left(\mathbf{Q}_\mathrm{n}^{(i)} + \mathbf{Q}_\mathrm{int}^{(i)} \right)^{-1} \right| \tag{28}$$

bounding the symbol power (in watts) at each LNA can be written as

$$\left| \sqrt{\tilde{P}_\mathrm{tx}^{(i)}} G_{ii} \left[\mathbf{H}_{ii} \right]_{\ell,:} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \mathbf{s}^{(i)} \right|^2 \leq \tilde{P}_\mathrm{SI,LNA}^\mathrm{max} \tag{20}$$

for all $\ell = 1, \ldots, N_\mathrm{r}^{(i)}$. Collecting these $N_\mathrm{r}^{(i)}$ constraints together, we can write

$$\mathrm{diag}\left(\tilde{P}_\mathrm{tx}^{(i)} G_{ii}^2 \mathbf{H}_{ii} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \mathbf{s}^{(i)} \mathbf{s}^{(i)*} \mathbf{F}_\mathrm{BB}^{(i)*} \mathbf{F}_\mathrm{RF}^{(i)*} \mathbf{H}_{ii}^{*} \right) \leq \tilde{P}_\mathrm{SI,LNA}^\mathrm{max} \cdot \mathbf{1}_{N_\mathrm{r}^{(i)}} \tag{21}$$

where $\mathbf{a} \leq \mathbf{b}$ denotes element-wise inequality. For a given channel realization, it is impractical to attempt to satisfy (21) on a per-symbol basis. Furthermore, it may be computationally expensive, severely sub-optimal, or potentially impossible to ensure (21) is met for all symbol vectors in a constellation. This motivates us to satisfy our constraint in expectation over $\mathbf{s}^{(i)}$ and, noting our defined symbol covariance (1), results in the constraint

$$\frac{\tilde{P}_\mathrm{tx}^{(i)} G_{ii}^2}{N_\mathrm{s}^{(ij)}} \cdot \mathrm{diag}\left(\mathbf{H}_{ii} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)*} \mathbf{F}_\mathrm{RF}^{(i)*} \mathbf{H}_{ii}^{*} \right) \leq \tilde{P}_\mathrm{SI,LNA}^\mathrm{max} \cdot \mathbf{1}_{N_\mathrm{r}^{(i)}}. \tag{22}$$

In a similar fashion, by incorporating the analog combiner $\mathbf{W}_\mathrm{RF}^{(i)}$, we can express our per-RF chain received self-interference power constraint as

$$\frac{\tilde{P}_\mathrm{tx}^{(i)} G_{ii}^2}{N_\mathrm{s}^{(ij)}} \cdot \mathrm{diag}\left(\mathbf{W}_\mathrm{RF}^{(i)*} \mathbf{H}_{ii} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)*} \mathbf{F}_\mathrm{RF}^{(i)*} \mathbf{H}_{ii}^{*} \mathbf{W}_\mathrm{RF}^{(i)} \right) \leq \tilde{P}_\mathrm{SI,ADC}^\mathrm{max} \cdot \mathbf{1}_{L_\mathrm{r}^{(i)}} \tag{23}$$

which captures all $L_\mathrm{r}^{(i)}$ ADCs in a single expression.

To abstract out the impact $\tilde{P}_\mathrm{tx}^{(i)}$ and G_{ii}^2 have on meeting our per-antenna power constraint $\tilde{P}_\mathrm{SI,LNA}^\mathrm{max}$ and per-RF chain power constraint $\tilde{P}_\mathrm{SI,ADC}^\mathrm{max}$, we introduce the unitless variables

$$\eta_\mathrm{LNA} \triangleq \frac{\tilde{P}_\mathrm{SI,LNA}^\mathrm{max}}{\tilde{P}_\mathrm{tx}^{(i)} \cdot G_{ii}^2}, \quad \eta_\mathrm{ADC} \triangleq \frac{\tilde{P}_\mathrm{SI,ADC}^\mathrm{max}}{\tilde{P}_\mathrm{tx}^{(i)} \cdot G_{ii}^2} \tag{24}$$

which will provide more generalized analysis across combinations of $\tilde{P}_\mathrm{SI,LNA}^\mathrm{max}$, $\tilde{P}_\mathrm{SI,ADC}^\mathrm{max}$, $\tilde{P}_\mathrm{tx}^{(i)}$, and G_{ii}^2. Stricter constraints are when η_LNA and η_ADC are low (we permit little self-interference at the receiver), while relaxed constraints are when η_LNA and η_ADC are high (we permit high self-interference). Using (24), the constraints in (22) and (23) can be equivalently expressed as

$$\frac{1}{N_\mathrm{s}^{(ij)}} \cdot \sigma_\mathrm{max}^2 \left(\mathbf{H}_{ii} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \right) \leq \eta_\mathrm{LNA} \tag{25}$$

$$\frac{1}{N_\mathrm{s}^{(ij)}} \cdot \sigma_\mathrm{max}^2 \left(\mathbf{W}_\mathrm{RF}^{(i)*} \mathbf{H}_{ii} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \right) \leq \eta_\mathrm{ADC} \tag{26}$$

respectively, where $\sigma_\mathrm{max}(\mathbf{A})$ denotes the maximum singular value of $\mathbf{A}$.

B. Satisfying our Constraints

With our LNA and ADC constraints in hand, we turn our attention to producing a hybrid beamforming design that satisfies (25) and (26) while achieving an appreciable spectral efficiency on our two links. The mutual information (under Gaussian signaling), or spectral efficiency (in bps/Hz), of the link from i to j is referred to as $\mathcal{R}_{ij}$ and takes the familiar form in (27) [24]. Treating the effects of self-interference as noise, the mutual information of the link from k to i is referred to as $\mathcal{R}_{ki}$ and expressed in (28). We let $\mathbf{Q}_\mathrm{n}^{(n)}$ be the covariance of noise at the detector of $n \in \{i, j\}$ normalized to the noise power and let $\mathbf{Q}_\mathrm{int}^{(i)}$ be the covariance of self-interference at the detector of i normalized to the noise power. Since noise and the transmitted symbols from i to j are uncorrelated, the self-interference-plus-noise covariance can be written as $\mathbf{Q}_\mathrm{n}^{(i)} + \mathbf{Q}_\mathrm{int}^{(i)}$.

We design our full-duplex system with the objective to maximize the sum spectral efficiency $\mathcal{R}_{ij} + \mathcal{R}_{ki}$ subject to our per-antenna and per-RF chain self-interference power constraints as well as the precoding power constraint in (2), described below in problem (29).

$$\max_{\substack{\left(\mathbf{F}_\mathrm{RF}^{(i)}, \mathbf{W}_\mathrm{RF}^{(j)}\right) \in \mathcal{T}_{ij} \\ \left(\mathbf{F}_\mathrm{RF}^{(k)}, \mathbf{W}_\mathrm{RF}^{(i)}\right) \in \mathcal{T}_{ki}}} \ \max_{\substack{\mathbf{F}_\mathrm{BB}^{(i)}, \mathbf{W}_\mathrm{BB}^{(j)} \\ \mathbf{F}_\mathrm{BB}^{(k)}, \mathbf{W}_\mathrm{BB}^{(i)}}} \ \mathcal{R}_{ij} + \mathcal{R}_{ki} \tag{29a}$$

$$\text{s.t.} \quad \left\| \mathbf{F}_\mathrm{BB}^{(i)} \right\|_\mathrm{F}^2 \leq 1, \ \left\| \mathbf{F}_\mathrm{BB}^{(k)} \right\|_\mathrm{F}^2 \leq 1 \tag{29b}$$

$$\frac{1}{N_\mathrm{s}^{(ij)}} \cdot \sigma_\mathrm{max}^2 \left(\mathbf{H}_{ii} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \right) \leq \eta_\mathrm{LNA} \tag{29c}$$

$$\frac{1}{N_\mathrm{s}^{(ij)}} \cdot \sigma_\mathrm{max}^2 \left(\mathbf{W}_\mathrm{RF}^{(i)*} \mathbf{H}_{ii} \mathbf{F}_\mathrm{RF}^{(i)} \mathbf{F}_\mathrm{BB}^{(i)} \right) \leq \eta_\mathrm{ADC} \tag{29d}$$

Note that we have restricted our choices for analog beamforming to the sets $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ supplied from beam candidate acquisition in Section IV. Solving problem (29) is difficult for a variety of reasons, chiefly the non-convexity arising from the interplay of the precoder at i in both transmit link performance and self-interference, meaning it impacts both $\mathcal{R}_{ij}$ and $\mathcal{R}_{ki}$. This motivates us to split our design into two stages. The first stage will be to configure a portion of our system subject to our constraints. Then, with the constraints met by the first stage, the second stage of our design will configure the remaining precoders and combiners.

The responsibility of satisfying the per-antenna self-interference power constraint lay solely in the transmitter of i as evidenced by (29c). In satisfying the per-RF chain constraint

$$\mathcal{I}_{ij} = \log_2 \left| \mathbf{I} + \frac{\mathrm{SNR}_{ij}}{N_{\mathrm{s}}^{(ij)}} \mathbf{W}_{\mathrm{RF}}^{(j)*} \mathbf{H}_{ij} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)*} \mathbf{F}_{\mathrm{RF}}^{(i)*} \mathbf{H}_{ij}^{*} \mathbf{W}_{\mathrm{RF}}^{(j)} \left(\mathbf{W}_{\mathrm{RF}}^{(j)*} \mathbf{W}_{\mathrm{RF}}^{(j)} \right)^{-1} \right| \tag{30}$$

(29d), the analog combiner $\mathbf{W}_{\mathrm{RF}}^{(i)}$ can assist, though—like the analog precoder $\mathbf{F}_{\mathrm{RF}}^{(i)}$—it may be very limited in its ability to do so. For both constraints, it is expected that $\mathbf{F}_{\mathrm{BB}}^{(i)}$ will often carry most of the weight in meeting these two constraints, given its digital nature. Since the price of meeting the per-antenna constraint is paid solely by the transmitter in $\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}$ and since we expect the digital precoder $\mathbf{F}_{\mathrm{BB}}^{(i)}$ to incur even more sacrifice as it meets the per-RF chain constraint, we are motivated to prioritize the transmit link during the first stage of our design *while* meeting these constraints. Note that by meeting these constraints, we will be *inherently* improving receive link performance by reducing self-interference.

Suppose during beam candidate acquisition, we let $K_{ij} = K_{ki} = 1$ (we consider $K_{ij}, K_{ki} \geq 1$ shortly), leading to the analog beamformers $\mathbf{F}_{\mathrm{RF}}^{(i)}$, $\mathbf{W}_{\mathrm{RF}}^{(j)}$, $\mathbf{F}_{\mathrm{RF}}^{(k)}$, and $\mathbf{W}_{\mathrm{RF}}^{(i)}$ being fixed as the first and only candidates in $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$. In such a case, we formulate problem (31), where we aim to maximize the transmit link mutual information $\mathcal{I}_{ij}$ in (30) subject to our per-antenna and per-RF chain constraints and the aforementioned precoding power constraint in (2).

$$\max_{\mathbf{F}_{\mathrm{BB}}^{(i)}} \quad \mathcal{I}_{ij} \tag{31a}$$

$$\text{s.t.} \quad \left\| \mathbf{F}_{\mathrm{BB}}^{(i)} \right\|_{\mathrm{F}}^{2} \leq 1 \tag{31b}$$

$$\frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^{2} \left(\mathbf{H}_{ii} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \right) \leq \eta_{\mathrm{LNA}} \tag{31c}$$

$$\frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^{2} \left(\mathbf{W}_{\mathrm{RF}}^{(i)*} \mathbf{H}_{ii} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \right) \leq \eta_{\mathrm{ADC}} \tag{31d}$$

Solving problem (31) would ensure that transmission from i to j is prioritized while preventing receiver-side saturation. Problem (31) is technically not convex but can be easily recast as such.

Theorem 1. *Problem* (31) *can be recast as a convex problem.*

Proof. Noting that problem (31) is non-convex in $\mathbf{F}_{\mathrm{BB}}^{(i)}$ but is convex in the product $\mathbf{F}_{\mathrm{BB}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)*}$, we rewrite problem (31) using the positive semidefinite substitution $\mathbf{X}_{\mathrm{BB}}^{(i)} = \mathbf{F}_{\mathrm{BB}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)*} \succeq \mathbf{0}$ as

$$\max_{\mathbf{X}_{\mathrm{BB}}^{(i)} \succeq \mathbf{0}} \quad \mathcal{I}_{ij} \tag{32a}$$

$$\text{s.t.} \quad \left\| \mathbf{F}_{\mathrm{BB}}^{(i)} \right\|_{\mathrm{F}}^{2} \leq 1 \tag{32b}$$

$$\frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^{2} \left(\mathbf{H}_{ii} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \right) \leq \eta_{\mathrm{LNA}} \tag{32c}$$

$$\frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^{2} \left(\mathbf{W}_{\mathrm{RF}}^{(i)*} \mathbf{H}_{ii} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \right) \leq \eta_{\mathrm{ADC}} \tag{32d}$$

$$\mathbf{F}_{\mathrm{BB}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)*} = \mathbf{X}_{\mathrm{BB}}^{(i)} \tag{32e}$$

With this simple restructuring, problem (32) is convex and can be solved efficiently using a convex solver (e.g., we used CVX [25]). Once solved, $\mathbf{X}_{\mathrm{BB}}^{(i)}$ can be factored to recover $\mathbf{F}_{\mathrm{BB}}^{(i)}$. □

Remark 1. The solution to problem (32) is unique but the factorization $\mathbf{X}_{\mathrm{BB}}^{(i)} = \mathbf{F}_{\mathrm{BB}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)*}$ is not. As such, our problem is a function of the digital precoder covariance and not of the digital precoder itself. In other words, the solution $\mathbf{X}_{\mathrm{BB}}^{(i)}$ to problem (32) can be factored arbitrarily, since $L_{\mathrm{t}}^{(i)} = N_{\mathrm{s}}^{(ij)}$, to retrieve a globally optimal $\mathbf{F}_{\mathrm{BB}}^{(i)}$.

Remark 2. Since $\mathbf{X}_{\mathrm{BB}}^{(i)} = \mathbf{F}_{\mathrm{BB}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)*} = \mathbf{0}$ is always a solution to problem (32), it is feasible. Intuitively, this can be attributed to the fact that it is always possible for the transmitter at i to shut off completely to ensure the LNAs and ADCs do not saturate.

Remark 3. At least one of the constraints (32b)–(32d) will be tight (have equality) under the optimal solution to problem (32). Intuitively, this can be attributed to the fact that additional power should be supplied to the digital precoder $\mathbf{F}_{\mathrm{BB}}^{(i)}$—which will increase the mutual information $\mathcal{I}_{ij}$—until it exceeds its power budget or self-interference power is too high at an LNA or ADC.

We now incorporate our sets of candidate beams $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ when $K_{ij}, K_{ki} \geq 1$, which offers the system more freedom in its design. By doing so, the diversity between candidate beams will generally allow our full-duplex transceiver i to better transmit to j while meeting the constraints. This is courtesy of the fact that some candidate beam pairs will naturally afford more isolation at the full-duplex device, reducing the costs paid by $\mathbf{F}_{\mathrm{BB}}^{(i)}$ as it attempts to reduce self-interference. This motivates us to wrap problem (31) with an outer maximization over $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$, resulting in the following optimization problem.

$$\max_{\substack{\left(\mathbf{F}_{\mathrm{RF}}^{(i)}, \mathbf{W}_{\mathrm{RF}}^{(j)}\right) \in \mathcal{T}_{ij} \\ \left(\mathbf{F}_{\mathrm{RF}}^{(k)}, \mathbf{W}_{\mathrm{RF}}^{(i)}\right) \in \mathcal{T}_{ki}}} \ \max_{\mathbf{F}_{\mathrm{BB}}^{(i)}} \ \mathcal{I}_{ij} \tag{33a}$$

$$\text{s.t.} \quad \left\| \mathbf{F}_{\mathrm{BB}}^{(i)} \right\|_{\mathrm{F}}^{2} \leq 1 \tag{33b}$$

$$\frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^{2} \left(\mathbf{H}_{ii} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \right) \leq \eta_{\mathrm{LNA}} \tag{33c}$$

$$\frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^{2} \left(\mathbf{W}_{\mathrm{RF}}^{(i)*} \mathbf{H}_{ii} \mathbf{F}_{\mathrm{RF}}^{(i)} \mathbf{F}_{\mathrm{BB}}^{(i)} \right) \leq \eta_{\mathrm{ADC}} \tag{33d}$$

Having shown that the inner maximization can be solved via the convex reformulation in (32), we can solve problem (33) exhaustively over all possible candidate combinations in $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$. Solving (33) will prioritize the transmit link while meeting the self-interference power constraints and will leverage the candidates from $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ to do so. As a result, the receive link may incur costs if the candidates chosen from

$\mathcal{T}_{ki}$ are not the optimal ones in a half-duplex sense. Choosing a smaller K_{ki} would constrain this receive link cost but would simultaneously increase transmit link cost. Increasing K_{ij} can only help transmit link performance but may incur additional complexity during beam candidate acquisition and/or solving problem (33).

Note that, as evidenced in (30), $\mathcal{I}_{ij}$ contains $\mathbf{H}_{ij}$, which we have not assumed explicit knowledge of. Instead, using $\mathcal{H}_{ij}$ from (19), we do have knowledge of the effective channel $\mathbf{W}_{\mathrm{RF}}^{(j)*}\mathbf{H}_{ij}\mathbf{F}_{\mathrm{RF}}^{(i)}$ for all $(\mathbf{F}_{\mathrm{RF}}^{(i)}, \mathbf{W}_{\mathrm{RF}}^{(i)}) \in \mathcal{T}_{ij}$, which can be used when solving problem (33). Furthermore, note that we do not require knowledge of $\mathbf{H}_{ki}$ or $\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ki}\mathbf{F}_{\mathrm{RF}}^{(k)}$ to solve (33). We do require knowledge of $\mathbf{H}_{ii}$ to construct constraints (33c) and (33d), which we have assumed knowledge of. Notice, however, that perhaps an estimate of $\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}$ can be used to compute the per-RF chain constraint (33d) since it may be estimated more reliably and frequently than that of $\mathbf{H}_{ii}$, given its relatively small size and fully-digital nature.

Solving problem (33) yields the design for five of the eight precoding and combining matrices: $\mathbf{W}_{\mathrm{BB}}^{(j)}$, $\mathbf{F}_{\mathrm{BB}}^{(k)}$, and $\mathbf{W}_{\mathrm{BB}}^{(i)}$ remain to be designed. Designing these will take place in the next stage of our design. Having prevented receiver-side components from saturating with appropriately chosen η_{LNA} and η_{ADC}, the receive chain at device i is approximately linear, allowing for a much more straightforward design of the receive link and preventing severe degradation of a desired receive signal.

C. Constraint Redundancy Conditions

We now pause from our design to more closely examine the interplay between our precoding power constraint, per-antenna self-interference power constraint, and per-RF chain self-interference power constraint. In doing so, we will see that under appropriate conditions the latter two constraints may be inherently met by the precoding power constraint and under other conditions, the per-RF chain self-interference power constraint may be inherently met by the per-antenna self-interference power constraint. Knowledge of these conditions can potentially accelerate solving problem (33) in a variety of ways by relieving us of one or more constraints. As such, the complexity associated with solving problem (33)—which depends heavily on one's choice of algorithm—is also dictated by the interplay between its three constraints. This interplay, as we will see, is characterized by a variety of factors including the realized self-interference channel $\mathbf{H}_{ii}$, η_{LNA}, η_{ADC}, and choice of analog beamformers/codebooks.

Theorem 2. *When condition* (34) *holds, the per-antenna constraint* (33c) *is implicitly met by the precoding power constraint* (33b).

$$\eta_{\mathrm{LNA}} \geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\right) \tag{34}$$

Proof. Since the precoding power constraint (31b) is satisfied, we note that

$$\sigma_{\max}^2\left(\mathbf{F}_{\mathrm{BB}}^{(i)}\right) \leq \left\|\mathbf{F}_{\mathrm{BB}}^{(i)}\right\|_{\mathrm{F}}^2 \leq 1 \tag{35}$$

which allows us to see that

$$\begin{aligned}\eta_{\mathrm{LNA}} &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\right) && (36)\\ &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\right) \cdot \sigma_{\max}^2\left(\mathbf{F}_{\mathrm{BB}}^{(i)}\right) && (37)\\ &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}\right) && (38)\end{aligned}$$

where we have used $\sigma_{\max}^2(\mathbf{AB}) \leq \sigma_{\max}^2(\mathbf{A}) \cdot \sigma_{\max}^2(\mathbf{B})$. □

Theorem 3. *When condition* (39) *holds, the per-RF chain constraint* (33d) *is implicitly met by the precoding power constraint* (33b).

$$\eta_{\mathrm{ADC}} \geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\right) \tag{39}$$

Proof. Using the fact from (35) and assuming (39) to be true, we can write

$$\begin{aligned}\eta_{\mathrm{ADC}} &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\right) && (40)\\ &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\right) \cdot \sigma_{\max}^2\left(\mathbf{F}_{\mathrm{BB}}^{(i)}\right) && (41)\\ &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}\right) && (42)\end{aligned}$$

□

Remark 4. When (34) and (39) hold, the (half-duplex) capacity-achieving strategy on the transmit link is the solution to problem (31) since the per-antenna and per-RF chain constraints vanish, leaving only the precoding power constraint.

Theorem 4. *When the per-antenna constraint* (33c) *is satisfied and condition* (43) *holds, the per-RF chain constraint* (33d) *is satisfied.*

$$\eta_{\mathrm{ADC}} \geq \eta_{\mathrm{LNA}} \cdot \sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\right) \tag{43}$$

Proof. Starting with our assumption that the per-antenna constraint (33c) is satisfied, we have

$$\frac{1}{N_{\mathrm{s}}^{(ij)}}\sigma_{\max}^2\left(\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}\right) \leq \eta_{\mathrm{LNA}} \leq \frac{\eta_{\mathrm{ADC}}}{\sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\right)} \tag{44}$$

which yields

$$\begin{aligned}\eta_{\mathrm{ADC}} &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\right) \cdot \sigma_{\max}^2\left(\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}\right) && (45)\\ &\geq \frac{1}{N_{\mathrm{s}}^{(ij)}} \cdot \sigma_{\max}^2\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}\right) && (46)\end{aligned}$$

which indicates directly that (33d) is satisfied. □

Corollary 4.1. *When $\eta_{\mathrm{LNA}} = 0$ and $\eta_{\mathrm{ADC}} > 0$, satisfying the per-antenna constraint* (33c) *also satisfies the per-RF chain constraint* (33d).

Remark 5. A condition analogous to Theorem 4 stating that satisfying the per-antenna constraint (33c) based solely on meeting the per-RF chain constraint (33d) is not possible.

$$\mathbb{E}\left[\mathbf{y}_{\mathrm{ADC}}\mathbf{y}_{\mathrm{ADC}}^{*}\right] = \frac{P_{\mathrm{tx}}^{(k)} G_{ki}^{2}}{N_{\mathrm{s}}^{(ki)}} \cdot \mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ki}\mathbf{F}_{\mathrm{RF}}^{(k)}\mathbf{F}_{\mathrm{BB}}^{(k)}\mathbf{F}_{\mathrm{BB}}^{(k)*}\mathbf{F}_{\mathrm{RF}}^{(k)*}\mathbf{H}_{ki}^{*}\mathbf{W}_{\mathrm{RF}}^{(i)} + \frac{P_{\mathrm{tx}}^{(i)} G_{ii}^{2}}{N_{\mathrm{s}}^{(ij)}} \cdot \mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)*}\mathbf{F}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}^{*}\mathbf{W}_{\mathrm{RF}}^{(i)} + \sigma_{\mathrm{n}}^{2} \cdot \mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{W}_{\mathrm{RF}}^{(i)} \tag{54}$$

$$\mathbf{W}_{\mathrm{BB}}^{(i)} = \frac{1}{\sqrt{P_{\mathrm{tx}}^{(k)}}G_{ki}}\left(\tilde{\mathbf{H}}_{ki}\tilde{\mathbf{H}}_{ki}^{*} + \frac{N_{\mathrm{s}}^{(ki)}}{\mathrm{SNR}_{ki}}\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{W}_{\mathrm{RF}}^{(i)} + \frac{N_{\mathrm{s}}^{(ki)}}{P_{\mathrm{tx}}^{(k)}G_{ki}^{2}}\mathbf{R}_{\mathrm{quant}}\right)^{-1}\tilde{\mathbf{H}}_{ki} \tag{57}$$

D. Remainder of the Design

We now complete our mmWave MIMO design. Solving (33) yields selections for $\mathbf{F}_{\mathrm{BB}}^{(i)}$, $\mathbf{F}_{\mathrm{RF}}^{(i)}$, $\mathbf{W}_{\mathrm{RF}}^{(j)}$, $\mathbf{F}_{\mathrm{RF}}^{(k)}$, and $\mathbf{W}_{\mathrm{RF}}^{(i)}$. Configuring $\mathbf{W}_{\mathrm{BB}}^{(j)}$, $\mathbf{F}_{\mathrm{BB}}^{(k)}$, and $\mathbf{W}_{\mathrm{BB}}^{(i)}$ remains, which we execute as follows. Having the rest of the transmit link configured, the optimal linear baseband combiner at j can be designed in a linear minimum mean square error (LMMSE) fashion as follows. Let $\tilde{\mathbf{H}}_{ij} \triangleq \mathbf{W}_{\mathrm{RF}}^{(j)*}\mathbf{H}_{ij}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)} \in \mathbb{C}^{L_{\mathrm{r}}^{(j)} \times N_{\mathrm{s}}^{(ij)}}$ be the effective transmit channel after solving (33). Note that this can computed from the product of $\mathbf{W}_{\mathrm{RF}}^{(j)*}\mathbf{H}_{ij}\mathbf{F}_{\mathrm{RF}}^{(i)}$, which is referenced from $\mathcal{H}_{ij}$, and $\mathbf{F}_{\mathrm{BB}}^{(i)}$. Then, the LMMSE baseband combiner at j can be constructed as

$$\mathbf{W}_{\mathrm{BB}}^{(j)} = \frac{1}{\sqrt{P_{\mathrm{tx}}^{(i)}}G_{ij}}\left(\tilde{\mathbf{H}}_{ij}\tilde{\mathbf{H}}_{ij}^{*} + \frac{N_{\mathrm{s}}^{(ij)}}{\mathrm{SNR}_{ij}}\mathbf{W}_{\mathrm{RF}}^{(j)*}\mathbf{W}_{\mathrm{RF}}^{(j)}\right)^{-1}\tilde{\mathbf{H}}_{ij}. \tag{47}$$

Since $\mathbf{F}_{\mathrm{BB}}^{(i)}$ will not generally diagonalize the effective channel, an LMMSE combiner at j will aim to reduce inter-stream interference and reject noise. This concludes configuration of the transmit link, having set $\mathbf{F}_{\mathrm{BB}}^{(i)}$, $\mathbf{F}_{\mathrm{RF}}^{(i)}$, $\mathbf{W}_{\mathrm{RF}}^{(j)}$, and $\mathbf{W}_{\mathrm{BB}}^{(j)}$.

We now turn our attention to the receive link, where we need to configure $\mathbf{F}_{\mathrm{BB}}^{(k)}$ and $\mathbf{W}_{\mathrm{BB}}^{(i)}$. Now that we have made our selections of $\mathbf{F}_{\mathrm{RF}}^{(k)}$ and $\mathbf{W}_{\mathrm{RF}}^{(i)}$, we begin by estimating the relatively small channel $\tilde{\mathbf{H}}_{ki} \triangleq \mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ki}\mathbf{F}_{\mathrm{RF}}^{(k)} \in \mathbb{C}^{L_{\mathrm{r}}^{(i)} \times L_{\mathrm{t}}^{(k)}}$, which can be observed digitally and we assume is error-free. Taking the singular value decomposition (SVD) of this effective channel from k to i and accounting for noise coloring, we get $\mathbf{U}_{ki}\mathbf{\Sigma}_{ki}\mathbf{V}_{ki}^{*} = \mathrm{SVD}\left(\left(\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{W}_{\mathrm{RF}}^{(i)}\right)^{-1/2}\tilde{\mathbf{H}}_{ki}\right)$. We then build the precoder that maximizes the mutual information offered to the RF chains of i as

$$\mathbf{F}_{\mathrm{BB}}^{(k)} = \left[\mathbf{V}_{ki}\right]_{:,1:N_{\mathrm{s}}^{(ki)}} \times \mathbf{P}^{(k)} \tag{48}$$

where $\mathbf{P}^{(k)}$ is a diagonal water-filling power allocation matrix [24].

Finally, we are left to configure $\mathbf{W}_{\mathrm{BB}}^{(i)}$. Before doing so, however, we notice that $\mathbf{W}_{\mathrm{BB}}^{(i)}$ and the received symbols it acts on exist in the digital domain, residing after the ADCs. Having knowledge of $\mathbf{H}_{ii}$, $\mathbf{s}^{(i)}$, and all other beamformers, we can synthesize the received self-interference and subtract it before applying our combiner $\mathbf{W}_{\mathrm{BB}}^{(i)}$. Recall from (14) that the symbol vector after the ADCs is $\mathbf{y}_{\mathrm{dig}} = \mathbf{y}_{\mathrm{des,ADC}} + \mathbf{y}_{\mathrm{int,ADC}} + \mathbf{y}_{\mathrm{noise,ADC}} + \mathbf{e}_{\mathrm{quant}}$. Since $\mathbf{y}_{\mathrm{dig}}$ is in the digital domain, we can compute $\mathbf{y}_{\mathrm{int,ADC}}$ as

$$\mathbf{y}_{\mathrm{int,ADC}} = \sqrt{P_{\mathrm{tx}}^{(i)}}G_{ii}\mathbf{W}_{\mathrm{RF}}^{(i)*}\mathbf{H}_{ii}\mathbf{F}_{\mathrm{RF}}^{(i)}\mathbf{F}_{\mathrm{BB}}^{(i)}\mathbf{s}^{(i)} \tag{49}$$

and can subtract it from $\mathbf{y}_{\mathrm{dig}}$ before applying our combiner, which yields

$$\mathbf{y}^{(i)} = \mathbf{y}_{\mathrm{dig}} - \mathbf{y}_{\mathrm{int,ADC}} \tag{50}$$

$$= \mathbf{y}_{\mathrm{des,ADC}} + \mathbf{y}_{\mathrm{noise,ADC}} + \mathbf{e}_{\mathrm{quant}}. \tag{51}$$

We can estimate the symbols intended for i from k by applying a combiner $\mathbf{W}_{\mathrm{BB}}^{(i)}$ to $\mathbf{y}^{(i)}$ as

$$\hat{\mathbf{s}}^{(k)} = \mathbf{W}_{\mathrm{BB}}^{(i)*}\mathbf{y}^{(i)} \tag{52}$$

$$= \mathbf{W}_{\mathrm{BB}}^{(i)*}\left(\mathbf{y}_{\mathrm{des,ADC}} + \mathbf{y}_{\mathrm{noise,ADC}} + \mathbf{e}_{\mathrm{quant}}\right) \tag{53}$$

which will be corrupted by additive noise and the effects of quantization. At first glance, it appears that self-interference does not play a role our symbol estimate $\hat{\mathbf{s}}^{(k)}$. However, given that our ADCs have limited resolution, the power of the quantization error term $\mathbf{e}_{\mathrm{quant}}$ will increase with increased self-interference power. Recall that this is precisely the motivation for our per-RF chain self-interference power constraint.

As evidenced by (53), we can see that the linear combiner $\mathbf{W}_{\mathrm{BB}}^{(i)}$ acts on a desired signal plus two noise terms. Before proceeding, let us find the covariance of $\mathbf{e}_{\mathrm{quant}}$. Finding the covariance of the symbols reaching the ADCs, we get (54). Applying (12) per-ADC allows us to write the covariance matrix of $\mathbf{e}_{\mathrm{quant}}$ as

$$\mathbf{R}_{\mathrm{quant}} = \mathbb{E}\left[\mathbf{e}_{\mathrm{quant}}\mathbf{e}_{\mathrm{quant}}^{*}\right] \tag{55}$$

$$= \frac{1}{1.5 \cdot 2^{2b}} \cdot \mathbf{I} \odot \mathbb{E}\left[\mathbf{y}_{\mathrm{ADC}}\mathbf{y}_{\mathrm{ADC}}^{*}\right] \tag{56}$$

where $\odot$ denotes the Hadamard (element-wise) product. The linear design of $\mathbf{W}_{\mathrm{BB}}^{(i)}$ that minimizes the mean square error (MSE) of $\hat{\mathbf{s}}^{(k)}$ is that of (57).

Now that our design is complete, we characterize the covariance of self-interference and of noise at the detector, which can be used to evaluate $\mathcal{R}_{ij}$ and $\mathcal{R}_{ki}$ in (27) and (28), respectively. Let $\mathbf{Q}_{\mathrm{int}}^{(n)}$—the covariance of the received quantization noise due to self-interference at i, normalized to the noise power—be $\mathbf{Q}_{\mathrm{int}}^{(i)} = \frac{1}{\sigma_{\mathrm{n}}^{2}}\mathbf{W}_{\mathrm{BB}}^{(i)*}\mathbf{R}_{\mathrm{quant}}\mathbf{W}_{\mathrm{BB}}^{(i)}$. Let $\mathbf{Q}_{\mathrm{n}}^{(n)}$—the covariance of the received noise at $n \in \{i,j\}$, normalized to the noise power—be written as $\mathbf{Q}_{\mathrm{n}}^{(n)} = \mathbf{W}_{\mathrm{BB}}^{(n)*}\mathbf{W}_{\mathrm{RF}}^{(n)*}\mathbf{W}_{\mathrm{RF}}^{(n)}\mathbf{W}_{\mathrm{BB}}^{(n)}$. This concludes our design, which we evaluate in the following section.

$$\mathcal{I}_{ki} = \log_2 \left| \mathbf{I} + \frac{\mathrm{SNR}_{ki}}{N_{\mathrm{s}}^{(ki)}} \mathbf{W}_{\mathrm{RF}}^{(i)*} \mathbf{H}_{ki} \mathbf{F}_{\mathrm{RF}}^{(k)} \mathbf{F}_{\mathrm{BB}}^{(k)} \mathbf{F}_{\mathrm{BB}}^{(k)*} \mathbf{F}_{\mathrm{RF}}^{(k)*} \mathbf{H}_{ki}^{*} \mathbf{W}_{\mathrm{RF}}^{(i)} \left(\mathbf{W}_{\mathrm{RF}}^{(i)*} \mathbf{W}_{\mathrm{RF}}^{(i)} \right)^{-1} \right| \tag{61}$$

VI. Numerical Results

We have simulated our system in a Monte Carlo fashion with the following parameters. For simplicity, we use 32-element, half-wavelength uniform linear arrays with isotropic elements at all devices, where the horizontal transmit and receive array at i are separated vertically by 10 wavelengths. Each transmitter and receiver is equipped with 2 RF chains and accordingly multiplexes 2 streams. The transmit power at each device is 30 dBm, and the noise power is -85 dBm. To model the channel between the transmit and receive arrays of device i, we use the following summation [2], [11],

$$\mathbf{H}_{ii} = \sqrt{\frac{\kappa}{\kappa+1}} \mathbf{H}_{ii}^{\mathrm{NF}} + \sqrt{\frac{1}{\kappa+1}} \mathbf{H}_{ii}^{\mathrm{FF}} \tag{58}$$

where the Rician factor κ captures the amount of power in the near-field portion relative to the far-field portion. The near-field component is modeled using a spherical-wave model [26] as $\left[\mathbf{H}_{ii}^{\mathrm{NF}}\right]_{v,u} = \frac{\gamma}{r_{u,v}} \exp\left(-\mathrm{j}2\pi \frac{r_{u,v}}{\lambda}\right)$, where $r_{u,v}$ is the distance between the u-th transmit antenna and the v-th receive antenna, λ is the carrier wavelength, and γ ensures that the channel is normalized such that $\mathbb{E}\left[\left\|\mathbf{H}_{ii}\right\|_{\mathrm{F}}^2\right] = N_{\mathrm{t}}^{(i)} N_{\mathrm{r}}^{(i)}$. Note that this near-field model is deterministic for a given relative array geometry at i. The far-field component captures reflections from the environment and is modeled using (5) with $N_{\mathrm{rays}} \sim \mathrm{Unif}\,(1, 15)$. The transmit and receive channels are modeled with $N_{\mathrm{rays}} \sim \mathrm{Unif}\,(4, 15)$. For both channel models, each ray's AoD and AoA are drawn from $\mathrm{Unif}\,(-\pi/2, \pi/2)$. During beam candidate acquisition, we assume properly normalized discrete Fourier transform (DFT) codebooks are used. Furthermore, we assume the number of measurements taken during beam candidate acquisition is sufficiently large on each link such that $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$ are built using the strongest rays in their respective channels.

Let us define the transmit link capacity $\mathcal{C}_{ij}$ as the maximum spectral efficiency possible on the link from i to j when drawing the analog precoder $\mathbf{F}_{\mathrm{RF}}^{(i)}$ and analog combiner $\mathbf{W}_{\mathrm{RF}}^{(j)}$ from the beam candidate set $\mathcal{T}_{ij}$ as

$$\mathcal{C}_{ij} = \max_{\mathbf{F}_{\mathrm{BB}}^{(i)}, \left(\mathbf{F}_{\mathrm{RF}}^{(i)}, \mathbf{W}_{\mathrm{RF}}^{(j)}\right) \in \mathcal{T}_{ij}} \mathcal{I}_{ij} \quad \text{s.t.} \quad \left\| \mathbf{F}_{\mathrm{BB}}^{(i)} \right\|_{\mathrm{F}}^2 \leq 1 \tag{59}$$

which can be achieved using the well known method of water-filled eigenbeamforming. Let us define the receive link capacity $\mathcal{C}_{ki}$ as the maximum spectral efficiency possible on the link from k to i when drawing the analog precoder $\mathbf{F}_{\mathrm{RF}}^{(k)}$ and the analog combiner $\mathbf{W}_{\mathrm{RF}}^{(i)}$ from the candidate set $\mathcal{T}_{ki}$ as

$$\mathcal{C}_{ki} = \max_{\mathbf{F}_{\mathrm{BB}}^{(k)}, \left(\mathbf{F}_{\mathrm{RF}}^{(k)}, \mathbf{W}_{\mathrm{RF}}^{(i)}\right) \in \mathcal{T}_{ki}} \mathcal{I}_{ki} \quad \text{s.t.} \quad \left\| \mathbf{F}_{\mathrm{BB}}^{(k)} \right\|_{\mathrm{F}}^2 \leq 1 \tag{60}$$

where $\mathcal{I}_{ki}$ is defined in (61).

While $\mathcal{C}_{ij}$ and $\mathcal{C}_{ki}$ are not the true channel capacities of $\mathbf{H}_{ij}$ and $\mathbf{H}_{ki}$, it is more meaningful when evaluating our results to use our codebook-based analog beamforming approach to accurately interpret the spectral efficiency gains (and costs) associated with our design versus a half-duplex system that is offered the same freedom in analog beamforming contained in $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$. Our design will, therefore, hope to achieve a sum spectral efficiency $\mathcal{R}_{ij} + \mathcal{R}_{ki} \geq \max\left\{\mathcal{C}_{ij}, \mathcal{C}_{ki}\right\}$ to justify operating in a full-duplex fashion rather than a half-duplex one.

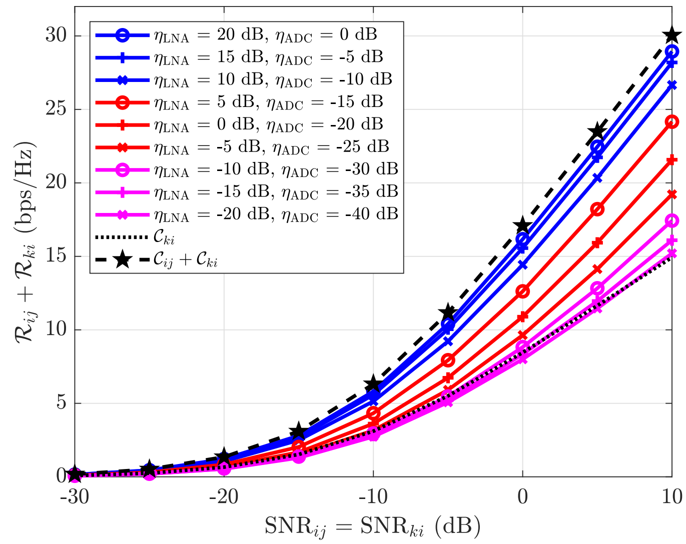
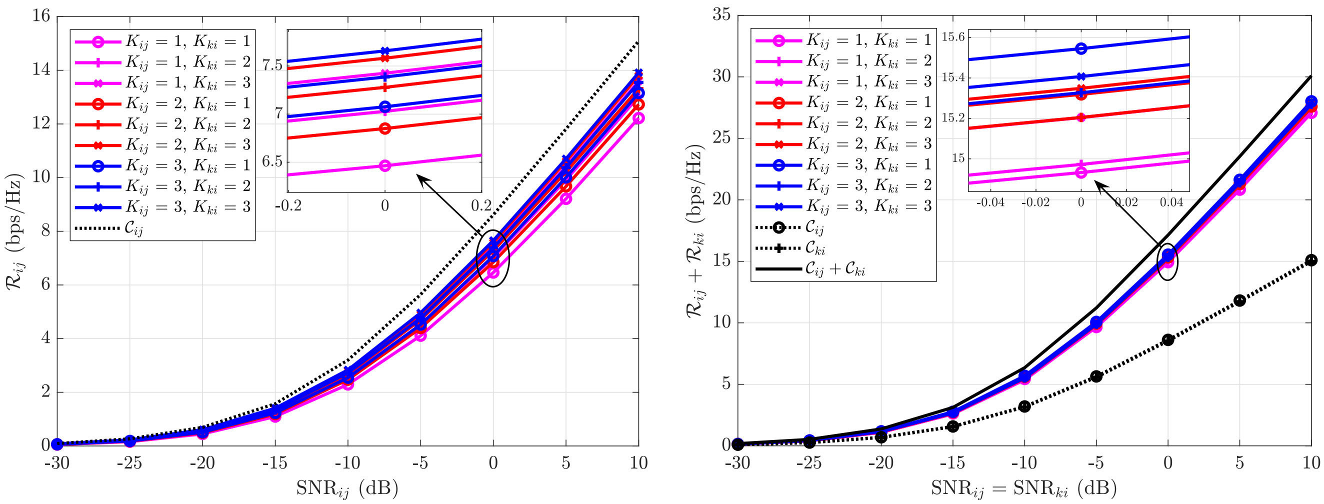


Fig. 2. Sum spectral efficiency as a function of SNR for various $\eta_{\mathrm{LNA}} = \eta_{\mathrm{ADC}} + 20$ dB, where $\kappa = 10$ dB, $b = 12$ bits, and $K_{ij} = K_{ki} = 3$. While not obvious from this figure, $\mathcal{R}_{ki}$ is fairly robust to self-interference with $b = 12$ bits. Stricter choices of $\eta_{\mathrm{LNA}} = \eta_{\mathrm{ADC}} + 20$ dB degrade $\mathcal{R}_{ij}$ and, thus, the sum $\mathcal{R}_{ij} + \mathcal{R}_{ki}$.

As intuition suggests, **the stricter the LNA and ADC constraints, the greater the sacrifice made on the transmit link's spectral efficiency to meet these constraints**. Fig. 2 confirms this, where the sum spectral efficiency $\mathcal{R}_{ij} + \mathcal{R}_{ki}$ as a function of SNR is evaluated at various LNA and ADC power constraints. While not explicitly shown, the loss in sum spectral efficiency due to decreasing $\eta_{\mathrm{LNA}} = \eta_{\mathrm{ADC}} + 20$ dB is due to loss in $\mathcal{R}_{ij}$ in its attempt to prevent receiver-side saturation. Given that we are primarily preserving the receive link by limiting the self-interference power reaching it, the receive link sees little sacrifice as a function of $\eta_{\mathrm{LNA}} = \eta_{\mathrm{ADC}} + 20$ dB, having assumed the resolution of the ADC is $b = 12$ bits, which is fairly robust to these levels of $\eta_{\mathrm{LNA}} = \eta_{\mathrm{ADC}} + 20$ dB (more on this later). Therefore, the lower bound on $\mathcal{R}_{ij} + \mathcal{R}_{ki}$, as one would hope, is approximately the half-duplex receive capacity $\mathcal{C}_{ki}$. The small gap when $\mathcal{R}_{ij} + \mathcal{R}_{ki} < \mathcal{C}_{ki}$ at very low $\eta_{\mathrm{LNA}} = \eta_{\mathrm{ADC}} + 20$ dB in Fig. 2 can be attributed to a small degree of quantization noise and, more significantly, the fact that the analog beamforming candidate chosen from $\mathcal{T}_{ki}$ is not always the $\mathcal{C}_{ki}$-achieving one, given that we have $K_{ki} = 3$. In other words, by design, the candidate from $\mathcal{T}_{ki}$ that maximizes



(a) Transmit link spectral efficiency.

(b) Sum spectral efficiency.

Fig. 3. Spectral efficiency as a function of SNR for various K_{ij} and K_{ki}, where $\kappa = 10$ dB, $\eta_{\mathrm{LNA}} = 15$ dB, $\eta_{\mathrm{ADC}} = -5$ dB, and $b = 12$ bits. Increasing K_{ij} and K_{ki} can only improve $\mathcal{R}_{ij}$, whereas increasing K_{ki} may degrade $\mathcal{R}_{ki}$. The choice of K_{ij} and K_{ki} that maximizes the $\mathcal{R}_{ij} + \mathcal{R}_{ki}$ varies with each realization.

transmit link performance subject to our constraints may not be the one that maximizes receive link performance.

Fig. 3 exhibits the gains in spectral efficiency afforded by increasing the number of analog beamforming candidates to our design. **By increasing K_{ij} and K_{ki}, the system can improve its performance on the *transmit* link while meeting the per-antenna and per-RF chain constraints.** Referring to Fig. 3a, when K_{ij} and K_{ki} increases from having only one candidate ($K_{ij} = K_{ki} = 1$) to having three candidates on each link ($K_{ij} = K_{ki} = 3$), we see a gain of approximately 1.25 bps/Hz in $\mathcal{R}_{ij}$ on average at $\mathrm{SNR}_{ij} = 0$ dB. This can be attributed to the fact that widening the search space will yield greater flexibility in meeting the constraints while maximizing performance on the transmit link. That is, rather than our optimization problem taking place over only $\mathbf{F}_{\mathrm{BB}}^{(i)}$, it also takes place over the candidates in $\mathcal{T}_{ij}$ and $\mathcal{T}_{ki}$. Interestingly, we can see that, on average, supplying our design with increased K_{ki} has a relatively greater impact than K_{ij}; this can be seen by the fact that $K_{ij} = 1, K_{ki} = 3$ outperforms $K_{ij} = 3, K_{ki} = 1$ and even $K_{ij} = 3, K_{ki} = 2$ in terms of $\mathcal{R}_{ij} + \mathcal{R}_{ki}$. With increased K_{ki}, some sacrifices may be made on the receive link by choosing a candidate from $\mathcal{T}_{ki}$ that is not the $\mathcal{C}_{ki}$-achieving one. In general, K_{ij} and K_{ki} can be chosen to throttle performance between the transmit link and receive link. Choosing a small K_{ki}, for example, preserves the receive link but reduces the full-duplex device's flexibility in avoiding self-interference. For a fixed K_{ki}, choosing a large K_{ij} can generally only help the system by increasing $\mathcal{R}_{ij}$, though, increasing the number of candidates (K_{ij} or K_{ki}) adds to the overhead associated with our design. It is difficult to draw generalized conclusions on which beam pairs work well or even precisely why since this depends so heavily on the transmit and receive channels, the self-interference channel, the beam codebooks, η_{LNA}, η_{ADC}, and a number of other factors; such would be interesting future work.

In Fig. 4, we evaluate the sum spectral efficiency $\mathcal{R}_{ij} + \mathcal{R}_{ki}$ for various selections of η_{LNA} and η_{ADC}. To better illustrate this, we let $K_{ij} = K_{ki} = 1$, placing the sole responsibility of preventing saturation on $\mathbf{F}_{\mathrm{BB}}^{(i)}$ and making $\mathcal{R}_{ki}$ approximately constant (≈ 3.2 bps/Hz) across all η_{LNA}, η_{ADC} (since $b = 12$ bits). Intuitively, as η_{LNA} and η_{ADC} increase, optimizing transmission from i becomes more relaxed, allowing for higher $\mathcal{R}_{ij}$. For a given choice of η_{LNA}, we can see that at increasing η_{ADC} beyond some point has little to no effect on $\mathcal{R}_{ij}$. This can be attributed to the fact that **the LNA constraint begins to supersede the ADC constraint** at these points (recall Theorem 4). As η_{LNA} increases, the point at which η_{ADC} plays no role also increases. Furthermore, beyond a certain η_{LNA} (e.g., $\eta_{\mathrm{LNA}} \geq 10$ dB), we can see that the LNA constraint becomes immaterial, suggesting that the precoding

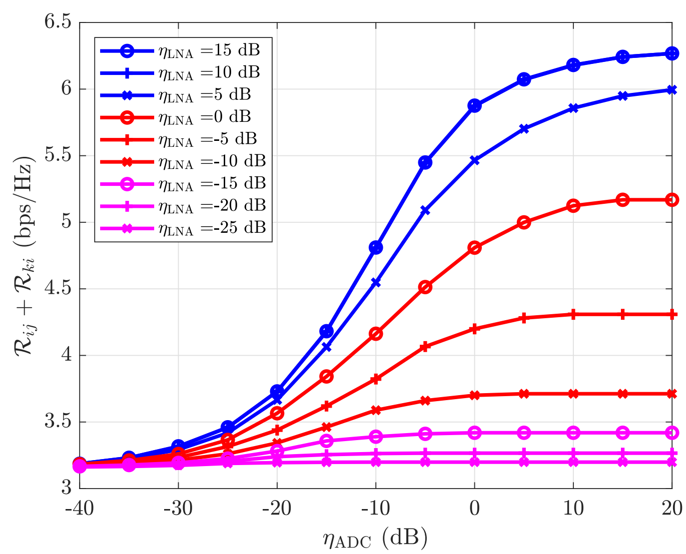
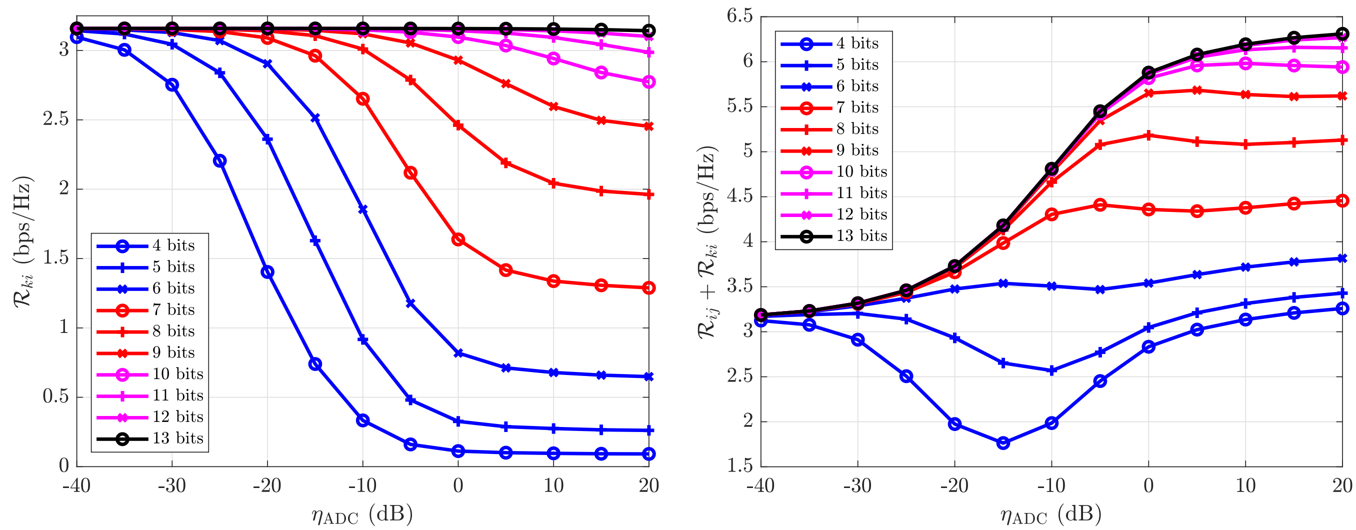


Fig. 4. Sum spectral efficiency as a function of η_{ADC} for various η_{LNA}, where $\kappa = 10$ dB, $K_{ij} = K_{ki} = 1$, $\mathrm{SNR}_{ij} = \mathrm{SNR}_{ki} = -10$ dB, and $b = 12$ bits. As η_{LNA} and η_{ADC} are relaxed, the system can achieve a greater sum spectral efficiency since transmit performance is less constrained by a limited receive dynamic range.



(a) Receive link spectral efficiency.

(b) Sum spectral efficiency.

Fig. 5. Spectral efficiency as a function of η_{ADC} for various ADC resolutions, where $\kappa = 10$ dB, $K_{ij} = K_{ki} = 1$, $\mathrm{SNR}_{ij} = \mathrm{SNR}_{ki} = -10$ dB, and $\eta_{\mathrm{LNA}} = 20$ dB. Higher-resolution ADCs are more robust to self-interference whereas lower-resolution ADCs saturate if η_{ADC} is not properly chosen. The tradeoff associated with constraining transmit link performance and preventing ADC saturation is not so obvious in terms of sum spectral efficiency.

power constraint implicitly satisfies the LNA constraint (recall Theorem 2).

Now, we examine the importance of η_{ADC} for various ADC resolutions. A key motivator for this work is the fact that self-interference can increase quantization noise, degrading the effective SNR of a desired signal out of the ADC. ADCs having greater resolution have a higher dynamic range, allowing them to quantize signal-plus-interference-plus-noise without suffering from ADC saturation as severely as lower-resolution ADCs. This can be seen in Fig. 5a, where we have evaluated $\mathcal{R}_{ki}$ as a function of η_{ADC} for various ADC resolutions, taking $\eta_{\mathrm{LNA}} = 20$ dB and $K_{ij} = K_{ki} = 1$ to reduce their impacts on interpreting these results. Higher-resolution ADCs are practically invariant across η_{ADC}, allowing them to achieve approximately the same $\mathcal{R}_{ki}$ regardless of the relative self-interference power at the ADCs. As the resolution decreases, we can see that quantization noise begins to take its toll on the spectral efficiency $\mathcal{R}_{ki}$, where it eventually plateaus beyond a certain η_{ADC} as the other constraints take effect. This highlights that an appropriate choice of η_{ADC} is intimately connected with the resolution of the ADCs and further justifies the motivation for this work: **under limited ADC resolution, the need to limit the self-interference power reaching the ADCs is critical**.

Interesting things happen in terms of the sum spectral efficiency $\mathcal{R}_{ij}+\mathcal{R}_{ki}$, with varying η_{ADC} and ADC resolutions, as depicted in Fig. 5b. As discussed, high-resolution ADCs are relatively invariant to η_{ADC}, and therefore, changes in sum spectral efficiency can be attributed almost exclusively to changes in $\mathcal{R}_{ij}$ as a function of η_{ADC}. With low-resolution ADCs (e.g., $b = 4, 5$ bits), we see that $\mathcal{R}_{ij} + \mathcal{R}_{ki}$ initially decreases sharply as η_{ADC} increases. This is due to the falloff that we saw in Fig. 5a. At very low (strict) η_{ADC}, $\mathcal{R}_{ij}$ is also very low, meaning the sharp falloff in $\mathcal{R}_{ki}$ drastically degrades the sum spectral efficiency. As η_{ADC} is increased (e.g., $\eta_{\mathrm{ADC}} = -10$ dB), $\mathcal{R}_{ij}$ also increases while $\mathcal{R}_{ki}$ begins to plateau. Recall that $\mathcal{R}_{ij}$ is invariant to the ADC resolution at i. As η_{ADC} further increases (e.g., $\eta_{\mathrm{ADC}} = 0$ dB), $\mathcal{R}_{ij}$ further increases and more rapidly so as its ADC saturation requirements become even more relaxed whereas $\mathcal{R}_{ki}$ further plateaus. Finally, as η_{ADC} increases further (e.g., $\eta_{\mathrm{ADC}} \geq 10$ dB), $\mathcal{R}_{ki}$ remains plateaued and $\mathcal{R}_{ij}$ begins to also plateau as it sees less gain in $\mathcal{R}_{ij}$ as changes in η_{ADC} hold less meaning, given the presence of $\eta_{\mathrm{LNA}} = 20$ dB and the precoding power constraint. For ADCs falling between very high resolutions and very low resolutions, the behavior can be explained in a similar fashion by this intertwining of $\mathcal{R}_{ij}$ and $\mathcal{R}_{ki}$, both of which begin to saturate beyond a certain η_{ADC}.

Fig. 5b highlights an important fact: **an appropriate design is necessary to make full-duplex operation worthwhile over half-duplex**. This is evidenced by the fact that low-resolution ADCs (e.g., $b = 4, 5$ bits) demand such significant self-interference mitigation that the sacrifice made on the transmit link is not worthwhile. Furthermore, we can see that the $(\mathcal{R}_{ij} + \mathcal{R}_{ki})$-optimal degree of self-interference power permitted at the ADCs varies with resolution. At these optimal choices for η_{ADC}, introducing a modest amount of quantization noise to improve transmit link performance balances $\mathcal{R}_{ij}$ and $\mathcal{R}_{ki}$ such that their sum is maximized.

Finally, for the sake of completeness and to better understand the role κ (the self-interference channel's Rician factor) from (58) plays in our design, we have included Fig. 6. We have evaluated the sum spectral efficiency for various pairs of η_{LNA} and η_{ADC} as a function of κ, fixing all other variables. Under relaxed conditions (i.e., high η_{LNA}, η_{ADC}), avoiding the self-interference channel becomes less of a priority, rendering κ less impactful. Under stringent conditions (i.e., low η_{LNA}, η_{ADC}), we can also see that κ does not play much of a role. This can be attributed to the fact that satisfying these strict LNA and ADC constraints is done so largely by power control of $\mathbf{F}_{\mathrm{BB}}^{(i)}$ rather than through steering strategies. This renders the underlying structure of $\mathbf{H}_{ii}$, and thus κ, less of a factor.

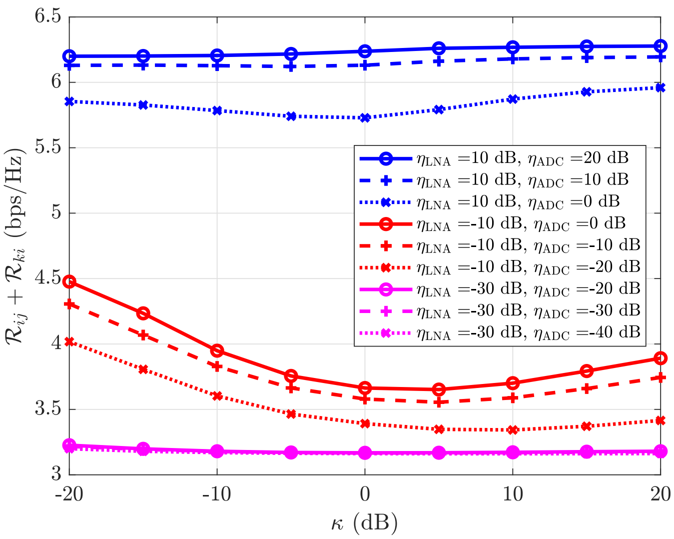


Fig. 6. Sum spectral efficiency as a function of κ for various $\eta_{\rm LNA}$ and $\eta_{\rm ADC}$, where $K_{ij} = K_{ki} = 1$, $\text{SNR}_{ij} = \text{SNR}_{ki} = -10$ dB, and $b = 12$ bits.

In between, however, we see that modest choices of $\eta_{\rm LNA}$ and $\eta_{\rm ADC}$ makes the role of κ more significant. When κ is very low, the self-interference channel is comprised primarily of far-field reflections. This leads to a self-interference channel that is spatially sparse, meaning avoiding $\mathbf{H}_{ii}$ becomes easier with highly directional DFT beams. This can be similarly stated that the inherent isolation between the rays of $\mathbf{H}_{ij}$ and $\mathbf{H}_{ii}$ and of $\mathbf{H}_{ki}$ and $\mathbf{H}_{ii}$ leads to more easily avoiding pushing self-interference onto the receiver of our full-duplex device. Similar behavior happens when κ is high, though the spatial sparsity of the self-interference channel stems from the near-field channel structure produced by our vertically-separated horizontal uniform linear arrays at the full-duplex device. Between, when κ approaches zero, the two spatially sparse channel components—the far-field portion and the near-field portion—mix relatively evenly, which leads to a self-interference channel that is less spatially sparse, making it more difficult to avoid pushing energy into with highly directional DFT beams. This leads to a relatively lower $\mathcal{R}_{ij}$ and, thus, lower sum spectral efficiency.

VII. Conclusion

We have presented a hybrid beamforming design for mmWave full-duplex that holistically covers a number of practical considerations including codebook-based analog beamforming and beam alignment, a desirably low number of RF chains, and the need to prevent receiver-side saturation at the full-duplex device. Core to our design is its focus on limiting the self-interference power reaching each antenna and each RF chain to prevent saturating LNAs and ADCs. Our design utilizes sets of candidate analog beamformers to improve its flexibility in mitigating self-interference while maintaining service and to accommodate codebook-based analog beamforming. Numerical results highlight the costs and limitations associated with preventing receiver-side saturation, which can be used in system analyses when developing a mmWave full-duplex transceiver and determining what levels of self-interference mitigation should be aimed for. Potential future work includes prototyping mmWave full-duplex, characterization of mmWave self-interference, and integrating full-duplex into mmWave cellular standards.

References

[1] A. Sabharwal *et al.*, "In-band full-duplex wireless: Challenges and opportunities," *IEEE JSAC*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[2] Z. Xiao, P. Xia, and X. Xia, "Full-duplex millimeter-wave communication," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 136–143, Dec. 2017.

[3] I. P. Roberts, J. G. Andrews, H. B. Jain, and S. Vishwanath, "Millimeter-wave full duplex radios: New challenges and techniques," *IEEE Wireless Commun.*, pp. 36–43, Feb. 2021.

[4] R. W. Heath *et al.*, "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] 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.

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

[7] G. C. Alexandropoulos and M. Duarte, "Joint design of multi-tap analog cancellation and digital beamforming for reduced complexity full duplex MIMO systems," in *Proc. IEEE Intl. Conf. Commun.*, May 2017, pp. 1–7.

[8] T. Dinc, A. Chakrabarti, and H. Krishnaswamy, "A 60 GHz CMOS full-duplex transceiver and link with polarization-based antenna and RF cancellation," *IEEE J. Solid-State Circuits*, vol. 51, no. 5, pp. 1125–1140, May 2016.

[9] T. Dinc *et al.*, "Synchronized conductivity modulation to realize broadband lossless magnetic-free non-reciprocity," *Nature Commun.*, vol. 8, no. 11, pp. 1–9, Oct. 2017.

[10] V. Singh *et al.*, "Millimeter-wave full duplex radios," in *Proc. ACM Intl. Conf. Mob. Comput. and Netw. (MobiCom)*, Apr. 2020.

[11] K. Satyanarayana *et al.*, "Hybrid beamforming design for full-duplex millimeter wave communication," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1394–1404, Feb. 2019.

[12] X. Liu *et al.*, "Beamforming based full-duplex for millimeter-wave communication," *Sensors*, vol. 16, no. 7, p. 1130, Jul. 2016.

[13] R. López-Valcarce and N. González-Prelcic, "Analog beamforming for full-duplex millimeter wave communication," in *Proc. Intl. Symp. Wireless Commun. Syst.*, Aug. 2019, pp. 687–691.

[14] ——, "Beamformer design for full-duplex amplify-and-forward millimeter wave relays," in *Proc. Intl. Symp. Wireless Commun. Syst.*, Aug. 2019, pp. 86–90.

[15] L. Zhu *et al.*, "Millimeter-wave full-duplex UAV relay: Joint positioning, beamforming, and power control," *IEEE JSAC*, vol. 38, no. 9, pp. 2057–2073, Sep. 2020.

[16] J. Palacios, J. Rodríguez-Fernández, and N. González-Prelcic, "Hybrid precoding and combining for full-duplex millimeter wave communication," in *Proc. IEEE Global Commun. Conf.*, Dec. 2019, pp. 1–6.

[17] I. P. Roberts and S. Vishwanath, "Beamforming cancellation design for millimeter-wave full-duplex," in *Proc. IEEE Global Commun. Conf.*, Dec. 2019, pp. 1–6.

[18] I. P. Roberts, H. B. Jain, and S. Vishwanath, "Equipping millimeter-wave full-duplex with analog self-interference cancellation," in *Proc. IEEE Intl. Conf. Commun. Wkshp.*, Jun. 2020.

[19] ——, "Frequency-selective beamforming cancellation design for millimeter-wave full-duplex," in *Proc. IEEE Intl. Conf. Commun.*, Jun. 2020, pp. 1–6.

[20] Y. Cai *et al.*, "Robust joint hybrid transceiver design for millimeter wave full-duplex MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1199–1215, Feb. 2019.

[21] D. Korpi *et al.*, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.

[22] Q. Gu, *RF System Design of Transceivers for Wireless Communications*. Springer, 2005.

[23] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE JSAC*, vol. 30, no. 8, pp. 1541–1553, Sep 2012.

[24] R. W. Heath Jr. and A. Lozano, *Foundations of MIMO Communication*. Cambridge University Press, 2018.
[25] M. Grant and S. Boyd, "CVX: Software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.
[26] J.-S. Jiang and M. A. Ingram, "Spherical-wave model for short-range MIMO," *IEEE Trans. Commun.*, vol. 53, no. 9, pp. 1534–1541, Sep. 2005.

BIOGRAPHIES

Ian P. Roberts is a Ph.D. student in the Department of Electrical and Computer Engineering at the University of Texas at Austin, where he is part of the Wireless Networking and Communications Group. He received his B.S. in electrical engineering from Missouri University of Science and Technology in 2018. He has industry experience developing and prototyping wireless technologies at AT&T Labs, GenXComm, Amazon, Sandia National Laboratories, and Dynetics. His research interests are in the theory and implementation of millimeter wave communication, in-band full-duplex, and communication system optimization. He is a National Science Foundation Graduate Research Fellow.

Jeffrey G. Andrews (S'98, M'02, SM'06, F'13) received the B.S. in Engineering with High Distinction from Harvey Mudd College, and the M.S. and Ph.D. in Electrical Engineering from Stanford University. He is the Cockrell Family Endowed Chair in Engineering at the University of Texas at Austin. He developed CDMA systems at Qualcomm, and has served as a consultant to Samsung, Nokia, Qualcomm, Apple, Verizon, AT&T, Intel, Microsoft, Sprint, and NASA. He is co-author of the books *Fundamentals of WiMAX* (Prentice-Hall, 2007) and *Fundamentals of LTE* (Prentice-Hall, 2010). He was the Editor-in-Chief of the IEEE Transactions on Wireless Communications from 2014-2016, and is the founding Chair of the Steering Committee for the IEEE Journal on Selected Areas in Information Theory, and the Chair of the IEEE Communication Theory Technical Committee (2021-22).

Dr. Andrews is an IEEE Fellow and ISI Highly Cited Researcher and has been co-recipient of 15 best paper awards including the 2016 IEEE Communications Society & Information Theory Society Joint Paper Award, the 2014 IEEE Stephen O. Rice Prize, the 2014 and 2018 IEEE Leonard G. Abraham Prize, the 2011 and 2016 IEEE Heinrich Hertz Prize, and the 2010 IEEE ComSoc Best Tutorial Paper Award. He received the 2015 Terman Award, the NSF CAREER Award, the 2021 Gordon Lepley Memorial Teaching Award, and the 2019 IEEE Kiyo Tomiyasu technical field award.

Sriram Vishwanath received the B. Tech. degree in Electrical Engineering from the Indian Institute of Technology (IIT), Madras, India in 1998, the M.S. degree in Electrical Engineering from California Institute of Technology (Caltech), Pasadena USA in 1999, and the Ph.D. degree in Electrical Engineering from Stanford University, Stanford, CA USA in 2003. He is the co-founder of GenXComm and a professor in the Department of Electrical & Computer Engineering at The University of Texas, Austin, USA.

Sriram received the NSF CAREER award in 2005 and the ARO Young Investigator Award in 2008. He has published over 250 refereed research papers, across the domains of Wireless Systems, AI/Machine Learning, Information Theory, Networking, Blockchain and Large-Scale Systems.