

MILLIMETER-WAVE FULL DUPLEX RADIOS: NEW CHALLENGES AND TECHNIQUES

Ian P. Roberts, Jeffrey G. Andrews, Hardik B. Jain, and Sriram Vishwanath

ABSTRACT

Equipping millimeter-wave (mmWave) systems with full duplex capability would accelerate and transform next-generation wireless applications and forge a path for new ones. Full duplex mmWave transceivers could capitalize on the already attractive features of mmWave communication by supplying spectral efficiency gains and latency improvements while also affording future networks deployment solutions in the form of interference management and wireless backhaul. Foreseeable challenges and obstacles in making mmWave full duplex a reality are presented in this article along with noteworthy unknowns warranting further investigation. With these novelties of mmWave full duplex in mind, we lay out potential solutions — beyond active self-interference cancellation — that harness the spatial degrees of freedom bestowed by dense antenna arrays to enable simultaneous transmission and reception in-band.

INTRODUCTION

New communication systems like fifth generation (5G) cellular and IEEE 802.11ad/ay harness the wide bandwidths available at millimeter-wave (mmWave) frequencies (roughly 30–100 GHz) to meet the ever growing demand for high-rate wireless access [1]. Cellular and local area mmWave communication systems rely on high beamforming gains provided by dense antenna arrays — on the order of dozens or hundreds of elements — to overcome the high path loss at mmWave frequencies and achieve sufficient link margin. Hybrid digital/analog beamforming architectures offer an efficient means to control these dense arrays with a reduced number of RF chains, making them ubiquitous in practical mmWave transceivers [2].

Concurrent to recent research on mmWave communication has been the development of in-band full duplex technology — a long sought after capability that allows a device to simultaneously transmit and receive across the same frequencies [3]. Full duplex has come a long way in the past decade, particularly in sub-6 GHz transceivers, largely thanks to novel and effective active self-interference cancellation (SIC) strategies (e.g., analog and digital SIC) that can make a desired receive signal virtually free of self-interference.

Equipping mmWave systems with full duplex would transform what is possible at the physical

layer and in medium access in next-generation networks. Most obviously, the throughput gains provided by full duplex would be magnified by the wideband, high-rate communication that is inherent to mmWave systems. Latency — a driving metric in future mmWave applications and networks — is improved with full duplex since delays associated with half duplexing transmission and reception can be avoided. The deployment of dense mmWave networks can be made more cost effective with full duplex integrated access and backhaul solutions, reducing the required density of fiber connectivity (a key hurdle to mmWave deployments) and preserving precious spectrum. Furthermore, in unlicensed mmWave spectrum such as the 57–64 GHz industrial, scientific, and medical (ISM) band and other lightly used bands, full duplex can introduce new approaches for coexistence between communication, consumer radar, and other incumbents.

Existing solutions for full duplex at lower frequencies do not immediately translate to mmWave systems due to fundamental differences between mmWave and sub-6 GHz transceivers. As noted, mmWave systems utilize many more antennas over much wider bandwidths, and have unique transceiver architectures and system design challenges. As we discuss, analog SIC is not well suited for the dense arrays and wide bandwidths found in mmWave communication. Beyond active cancellation, multiple-input multiple-output (MIMO) precoding and combining strategies were explored in sub-6 GHz full duplex, which aim to mitigate the self-interference by exploiting spatial degrees of freedom [4]. These MIMO-based approaches offer inspiration for mmWave full duplex solutions, but features such as hybrid beamforming, wide bandwidths, high sampling rates, beam alignment, and propagation characteristics will dictate what is possible and practical at mmWave. While passive approaches (e.g., highly directive antennas, polarization separation) have been proposed for mmWave full duplex, this article pertains to transceivers with dense antenna arrays and assumes that the use of passive methods could potentially supplement the discussions herein.

We begin the remainder of this article by highlighting the unique challenges and considerations of mmWave full duplex, chiefly hybrid beamforming and the self-interference channel. Then we present several promising directions for achiev-

ing mmWave full duplex. In particular, we discuss how the spatial domain — which presents some of the key challenges for full duplex at mmWave — can in fact be harnessed to enable beamforming-based approaches to mitigating self-interference. Throughout this article, we aim to inform and inspire readers on the challenges, unknowns, and potential solutions surrounding mmWave full duplex with the hope they identify research problems to pursue.

THE IMPLICATIONS OF HYBRID BEAMFORMING ARCHITECTURES

A plausible full duplex mmWave transceiver is depicted in Fig. 1, where separate, independently controlled arrays are used for transmission and reception. The use of separate arrays appears to be the most practical approach, given that wide-band, small form-factor circulators with sufficient isolation for mmWave full duplex are still out of reach [5]. A full duplex mmWave transceiver aims to transmit to a distant receiver while receiving from a distant transmitter in-band. The self-interference channel between the transmit array and receive array of a full duplex mmWave transceiver is discussed in the next section. In this section, key system-level implications of mmWave transceiver architectures on full duplex are outlined; a more detailed component-wise analysis would be valuable future work.

For mmWave communication, planar arrays on the order of 16–256 antenna elements are typical. To operate these dense antenna arrays with a reduced number of RF chains, mmWave transceivers often employ the combination of analog and digital beamforming, termed “hybrid beamforming,” which provides high beamforming gain while also supporting spatial multiplexing. Digital beamforming takes place at baseband in software or digital logic, whereas analog beamforming is implemented at passband (i.e., RF) in analog as a network of phase shifters and possibly attenuators. The phase shifters and attenuators making up a practical analog beamformer are likely controlled digitally, subjecting them to phase and amplitude quantization, respectively. Fortunately, highly focused beams can be constructed with low-resolution phase shifters and attenuators. The ability to create arbitrary beams, however, is lost with quantized phase and amplitude control, which may restrict beamforming-based full duplex solutions. Furthermore, the mathematical constraints imposed by quantized phase and amplitude control (chiefly their non-convexity) complicate the optimization of beamforming-based solutions for full duplex, although efficient solutions to overcome these complications in half duplex settings may offer inspiration [6, 7]. While most mmWave research ignores amplitude control in analog beamforming for half duplex settings, we expect it will be extremely useful in tailoring beamforming-based solutions for mmWave full duplex.

Given that communication at mmWave harnesses channels on the order of hundreds of megahertz or even gigahertz, frequency selectivity is a serious concern for both beamforming-based and filter-based full duplex solutions. Even with orthogonal frequency-division multiplexing (OFDM), frequency-selective beamforming

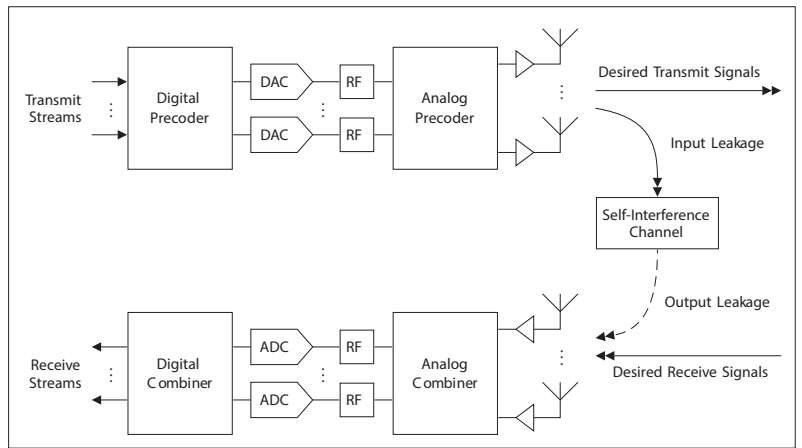


FIGURE 1. A full duplex mmWave transceiver employing hybrid beamforming with separate transmit and receive arrays. Double-ended arrows indicate multi-antenna signals.

is not straightforward with hybrid beamforming. Unlike digital beamforming, which can beamform on a per-subcarrier basis, analog beamforming is (relatively) frequency-flat, treating all subcarriers equally. Filter-based approaches for mmWave full duplex will require several taps to address frequency selectivity and will need to have wide-band support.

Notice that Fig. 1 depicts the power amplifiers (PAs) and low noise amplifiers (LNAs) as placed per-antenna, as is common to avoid losses and noise before and after the antenna, respectively. The digital-to-analog converters (DACs) and analog-to-digital converters (ADCs), along with upconversion and downconversion, are in the RF chains of the transceiver. The placement of these components is of immense importance to full duplex systems needing to address nonlinearity and receiver-side saturation [8]. The several high-rate DACs and ADCs, along with the dozens or hundreds of PAs and LNAs, will play a critical role in the requirements and design of mmWave full duplex solutions. Practical implementations may resort to more affordable, lower-quality components — particularly the numerous PAs and LNAs — which may complicate full duplex. Generally, wideband mmWave communication takes place in low signal-to-noise ratio (SNR) regimes, largely due to propagation losses and high integrated noise power. A raised noise floor actually relaxes full duplex requirements to a degree since more self-interference can be tolerated while remaining noise-limited. Furthermore, low-SNR communication demands fewer bits of quantization, meaning more self-interference can be tolerated at the ADCs (relative to the desired signal) or, alternatively, lower-resolution ADCs can be used, which is particularly attractive given their high sampling rate requirements.

SELF-INTERFERENCE CHANNELS AT MMWAVE: HOW TO MODEL AND ESTIMATE THEM?

When attempting to transmit and receive from a mmWave full duplex transceiver, self-interference is inflicted by each transmit element onto the entire receive array, collectively producing a large MIMO channel (e.g., of size 64×64). This high-dimensional over-the-air self-interfer-

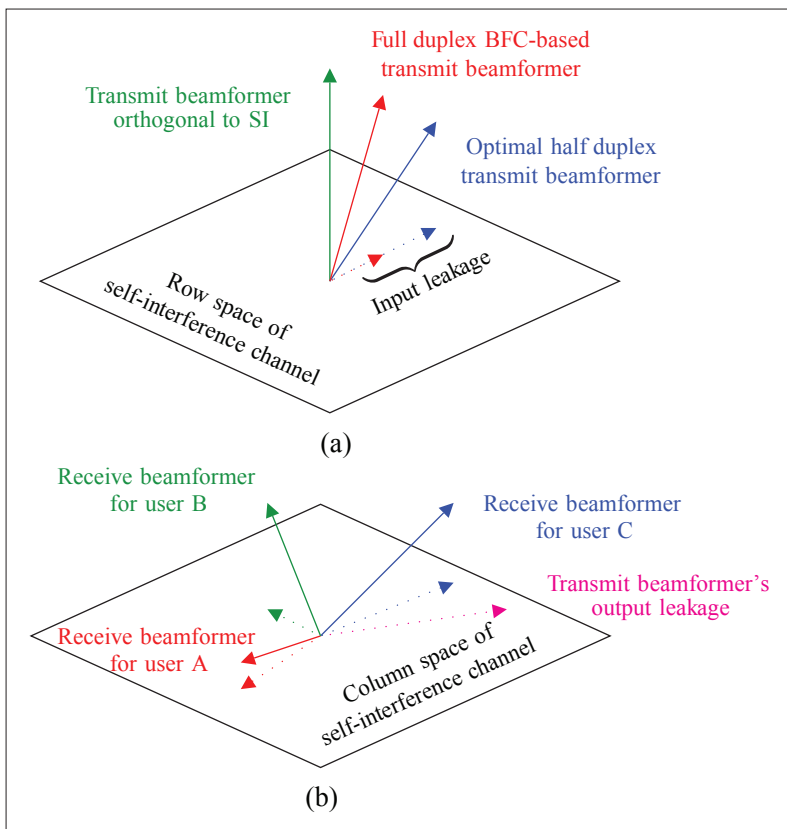


FIGURE 2. Illustrations of beamforming cancellation.

ence channel is an important difference between mmWave and conventional sub-6 GHz full duplex and motivates many points of discussion throughout this article. As shown in Fig. 1, we refer to the portion of the transmit signals that get transformed by the self-interference channel as “input leakage” and the resulting interference striking the receive array as “output leakage.”

At this time, we are not aware of existing work on modeling or measuring the self-interference channel at mmWave frequencies. A reasonable starting point has been presented in [9], which we outline as follows. A consequence of the close-in nature of a full duplex mmWave transceiver’s arrays is that their separation likely does not meet the far-field condition (e.g., $2D^2/\lambda$). Thus, it is reasonable to assume that its arrays will interact in a near-field fashion to some degree. Along with near-field effects, reflections off the environment — presumably in the far-field — will inflict additional self-interference. Combining these two contributors, the MIMO self-interference channel matrix at a given instant can be written in the following manner.

$$\mathbf{H}_{\text{SI}} = G_{\text{SI}} \cdot \left(\underbrace{\sqrt{\frac{\kappa}{\kappa+1}} \mathbf{H}_{\text{SI}}^{\text{NF}}}_{\text{near-field}} + \underbrace{\sqrt{\frac{1}{\kappa+1}} \mathbf{H}_{\text{SI}}^{\text{FF}}}_{\text{far-field}} \right) \quad (1)$$

The component $\mathbf{H}_{\text{SI}}^{\text{NF}}$ captures near-field contributions directly from the transmit array to the receive array (e.g., [10]), whereas $\mathbf{H}_{\text{SI}}^{\text{FF}}$ captures far-field contributions from a reflective environment (e.g., a ray-based model). When $\mathbf{H}_{\text{SI}}^{\text{NF}}$ and $\mathbf{H}_{\text{SI}}^{\text{FF}}$ are normalized to equal energy levels, the Rician factor κ throttles the large-scale power dis-

parity between the two. The large-scale gain G_{SI} captures the RF isolation between the arrays. It is important to note that the high path loss and penetration loss faced at mmWave frequencies is helpful in mitigating self-interference, especially the portion stemming from far-field reflections. While this model has yet to be verified with measurements or electromagnetic simulation software, which are important next steps, it offers a starting point for early research on mmWave full duplex. The delay spread and coherence time of the self-interference channel are difficult to speculate, although they will certainly be pertinent to realizing mmWave full duplex.

Practically, channel estimation is difficult at mmWave for several reasons. Many strategies have been developed to overcome these challenges, often employing compressed sensing to leverage the spatial and temporal sparsity exhibited by point-to-point mmWave channels [2, 6]. This sparsity is known to pertain to point-to-point far-field mmWave channels, such as those between devices in cellular and local area networks, but has not been confirmed to exist in mmWave self-interference channels. Thus, existing channel estimation strategies do not necessarily readily apply to the self-interference channel. An mmWave full duplex solution, therefore, may warrant novel self-interference channel estimation strategies, making it an attractive topic for future work. These new estimation strategies should aim to have low overhead, given the self-interference channel’s size, and perhaps the transmit channel and self-interference channel can be estimated using the same time-frequency resources to further reduce this overhead. However, characterization and modeling of the self-interference channel will be essential first steps before developing means to estimate it. For instance, if the self-interference channel is near-field dominant (i.e., κ is very large), its estimation and how frequently it is estimated will not be at the hand of the dynamics of the far-field environment (e.g., cars, people). In such a case, perhaps reliable near-field channel models and/or proper calibration accounting for nearby infrastructure could accelerate or potentially replace self-interference channel estimation.

CAN WE EXTEND ANALOG AND DIGITAL SELF-INTERFERENCE CANCELLATION TO MMWAVE?

We now turn our attention to potential approaches for enabling mmWave full duplex and begin by considering popular full duplex solutions for sub-6 GHz systems: analog SIC and digital SIC. Digital SIC aims to mitigate residual self-interference — often both linear and significant nonlinear terms — after analog SIC. Fortunately, digital SIC remains a promising candidate for mmWave full duplex since the number of RF chains remains low with hybrid beamforming, although exaggerated impairments in mmWave transceivers may drive up computational costs if not dealt with beforehand. A deeper investigation into digital SIC for mmWave full duplex would likely yield many useful insights, particularly on transceiver nonlinearity and other impairments such as I/Q imbalance, carrier frequency offset, and phase noise, making it a good topic for future work.

Analog SIC traditionally serves two main purposes in full duplex:

- Prevent receiver-side saturation (e.g., at the LNAs and ADCs).
- Capture and cancel nonidealities introduced by the transmit chain.

To consider analog SIC as a potential full duplex solution for mmWave systems, it is important to examine the placement of mmWave transceiver components, chiefly PAs, LNAs, and ADCs. It is often the case that PAs and LNAs are per-antenna, meaning they are after analog precoding and before analog combining, respectively, whereas ADCs are per-RF chain after analog combining. For an analog SIC solution to capture nonlinear terms introduced by the transmit PAs, which have proven to be a bottleneck in sub-6 GHz full duplex systems [8], it will need to grow in one dimension with the number of transmit antennas. Likewise, to prevent LNA saturation, analog SIC will need to grow in its second dimension with the number of receive antennas. Its third dimension, the delay dimension (number of taps), will be based on the impulse response of the self-interference channel. Considering that mmWave systems will operate using dozens or hundreds of antennas over wide bandwidths, analog SIC solutions would presumably be relatively large in all three dimensions at mmWave and, as a result, likely prohibitive in size and complexity. The obstacles faced by extending analog SIC to mmWave systems motivate new approaches that address receiver-side saturation during full duplex operation.

BEAMFORMING CANCELLATION: GIVING MMWAVE FULL DUPLEX SOME SPACE

The dense antenna arrays at mmWave seem to complicate the extension of analog SIC solutions but simultaneously promote the spatial domain as a promising arena for mitigating self-interference. By strategically transmitting and receiving from its numerous antennas, an mmWave full duplex transceiver can potentially mitigate self-interference through various forms of beamforming cancellation [9, 11–13]. Transmit-side beamforming cancellation aims to reduce the output leakage reaching the receive array while still transmitting to a distant receiver. Similarly, receive-side beamforming cancellation aims to reject the output leakage while receiving from a distant transmitter.

The principle of beamforming cancellation is to tailor the analog and digital precoders (\mathbf{F}_{RF} and \mathbf{F}_{BB}) and the analog and digital combiners (\mathbf{W}_{RF} and \mathbf{W}_{BB}) of a full duplex mmWave transceiver to reduce the strength of the following effective self-interference channel.

$$\underbrace{\mathbf{W}_{\text{BB}}^* \mathbf{W}_{\text{RF}}^*}_{\text{combiners}} \mathbf{H}_{\text{SI}} \underbrace{\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}}_{\text{precoders}} \quad (2)$$

By reducing the power of the effective self-interference channel through spatial techniques, beamforming cancellation can reduce the presence of self-interference in the time-frequency domain to facilitate simultaneous transmission and reception in-band. Beamforming cancellation, however, will require more than merely extending existing interference-related MIMO designs due to complications related to hybrid beamforming,

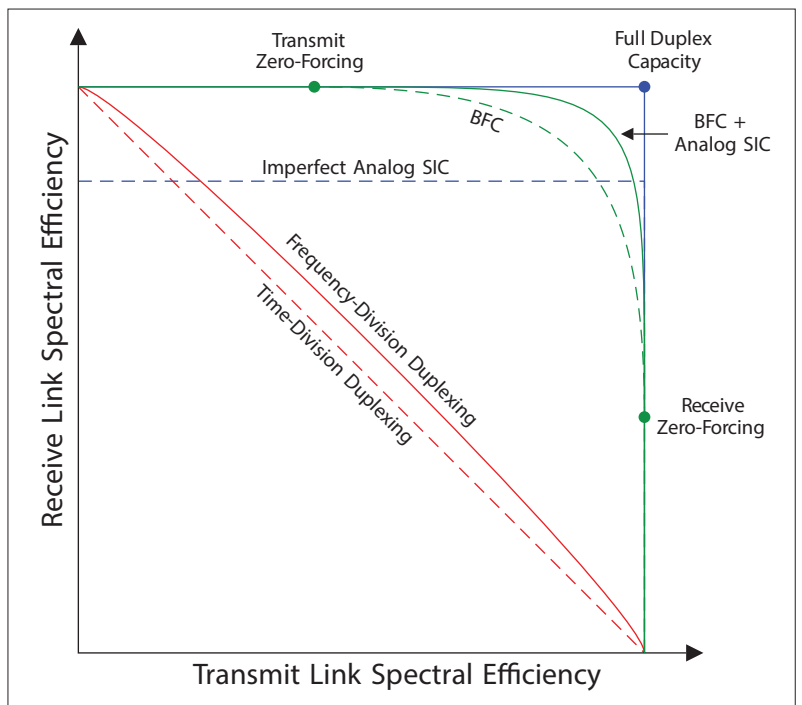


FIGURE 3. A depiction of the spectral efficiency region boundaries for various duplexing strategies employed by an mmWave transceiver (not to scale; BFC: beamforming cancellation). Supplementing any of these strategies with digital SIC can reduce the gap with the full duplex capacity.

the unique intertwining of the self-interference channel, and the need to prevent receiver-side saturation.

We expect that full duplex mmWave transceivers will likely depend on digital SIC to some degree. Thus, acting as a substitute for analog SIC, the primary objective of beamforming cancellation is to prevent LNA and ADC saturation to preserve the quality of the desired receive signal and give digital SIC a fighting chance. Perhaps, however, much more mitigation can be provided beyond merely addressing receiver-side saturation.

Referring to Fig. 1, the effective self-interference channel from the transmitter to per-antenna LNAs is $\mathbf{H}_{\text{SI}} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}$, indicating that the responsibility of preventing LNA saturation lies solely at the transmitter. Note that transmit power control can always ensure that LNA saturation is avoided and would be an attractive tool in conjunction with steering strategies. On the other hand, the effective self-interference channel from the transmitter to per-RF chain ADCs is $\mathbf{W}_{\text{RF}}^* \mathbf{H}_{\text{SI}} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}$, which indicates that the analog combiner at the receiver can aid the transmitter in preventing ADC saturation. This is convenient since ADC saturation requirements are often stricter than those of LNA saturation. The role of the baseband combiner \mathbf{W}_{BB} is somewhat arbitrary since it lives in the digital domain (after the ADCs), meaning that linear interference rejection can be applied and/or more sophisticated digital SIC algorithms.

To illustrate a potential beamforming cancellation strategy resembling conventional zero-forcing and minimum mean square error (MMSE) approaches, let us consider Fig. 2a. Depicted is the row space of the self-interference channel matrix \mathbf{H}_{SI} along with three beamforming vectors. The green vector orthogonal to the row

Beamforming cancellation is an attractive approach toward mmWave full duplex, but it does generally incur some loss in spectral efficiency when compared to the full duplex capacity as it attempts to reduce self-interference spatially. Fortunately, the high spatial degrees of freedom provided by mmWave arrays suggests that this loss is often tolerable.

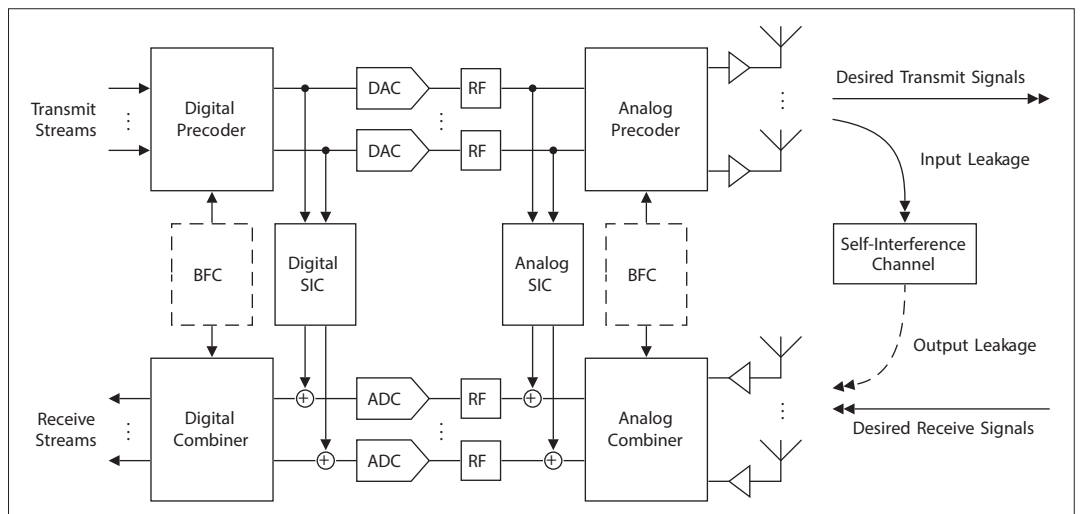


FIGURE 4. A full duplex mmWave transceiver architecture employing beamforming cancellation (BFC), reduced-size analog SIC, and digital SIC.

space represents a transmit beamformer (i.e., a column of \mathbf{F}_{RF} or $\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$) that would *completely* avoid inflicting self-interference onto the receive array of the full duplex transceiver (i.e., a zero-forcing approach), driving Eq. 2 to zero but perhaps transmitting poorly to a distant user. The optimal half duplex transmit beamformer in blue, on the other hand, has a significant component in the row space, potentially leading to a high degree of self-interference onto the receive array. Between these two beamformers, a potential beamforming-cancellation-based transmitter may live, inflicting a reduced amount of self-interference onto the receive array while still transmitting effectively to a distant receiver. Beamforming at the full duplex receiver could operate similarly to further reject self-interference. We remark that this sort of orthogonality-based approach is merely one of countless beamforming cancellation approaches that could be considered — the exact method employed would depend on one’s objective (e.g., weighted sum spectral efficiency, weighted MSE) among other factors. We further remark that the users being served by the full duplex device can (and likely should) adjust their transmit and receive strategies in accordance with beamforming cancellation taking place at the full duplex device.

The degree to which a full duplex device inflicts self-interference depends on not only its transmit beamformer but also its receive beamformer, as evidenced by Eq. 2. This intertwining of transmit and receive beamforming can seriously complicate their optimization. In the illustration of Fig. 2a, we have focused on tackling the input leakage and not the output leakage for this precise reason. While the true significance of transmit beamforming cancellation is actually in tailoring the output leakage, by shrinking the strength (i.e., norm) of the input leakage (and not changing its direction), we directly reduce the strength of the output leakage. Adjusting both the strength and direction of the output leakage would certainly be preferred, but such a strategy demands a joint design of the transmit and receive beamformers. As in Fig. 2b, fixing the transmitter thereby collapsing the effective self-interference channel down to the resulting output leakage can be a

powerful tool since the number of receive antennas is much greater than the number of transmit streams or even transmit RF chains. To summarize, beamforming cancellation is not about orthogonalizing transmission and reception, but rather the output leakage and reception, since transmissions are transformed by the self-interference channel before reaching the receive array.

THE COSTS AND LIMITATIONS OF BEAMFORMING CANCELLATION

Beamforming cancellation is an attractive approach toward mmWave full duplex, but it does generally incur some loss in spectral efficiency when compared to the full duplex capacity¹ as it attempts to reduce self-interference spatially. Fortunately, the high spatial degrees of freedom provided by mmWave arrays suggest that this loss is often tolerable since only a few degrees of freedom are used to communicate information, leaving many to tackle self-interference.

Figure 3 illustrates the spectral efficiency region boundaries of an mmWave full duplex transceiver’s transmit and receive links. Perfect execution of analog SIC cancels self-interference completely, achieving the full duplex capacity since no time-frequency-space resources are consumed to duplex transmission and reception. Imperfect analog SIC plagues the receive link with residual self-interference, degrading its spectral efficiency. Beamforming cancellation can approach the full duplex capacity, although it falls short for two reasons:

- By deviating from optimal half duplex strategies on the transmit and receive links, effectively consuming spatial resources
- By permitting some residual self-interference to more optimally transmit and receive

Applying beamforming cancellation with analog SIC, as we discuss later, can inch the system closer to the full duplex capacity. It is important to note that, unlike relying solely on analog SIC, beamforming cancellation affords the system the ability to trade transmit link performance for receive link performance — a very powerful concept, especially when the transmit and receive links are disparate. While not always possible,

¹ We define the full duplex capacity as the maximum sum spectral efficiency afforded by the time-frequency-space resources of the transmit and receive channels.

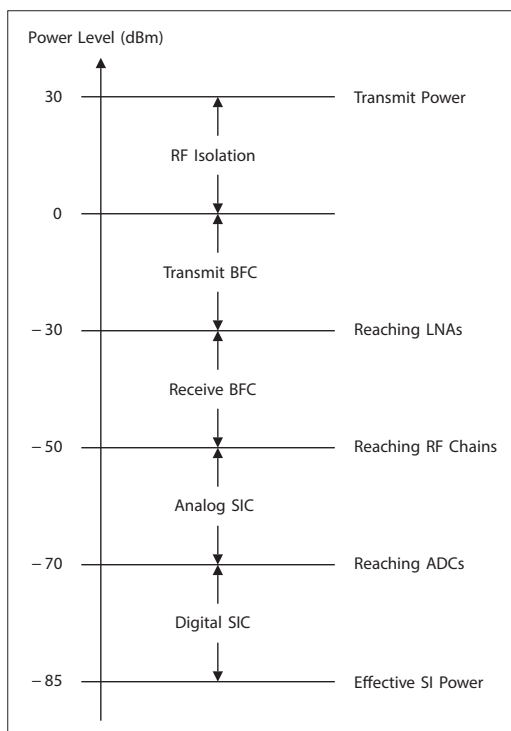


FIGURE 5. Example breakdown of self-interference power levels at various points in a full duplex mmWave transceiver employing beamforming cancellation (BFC), reduced-size analog SIC, and digital SIC.

transmit zero-forcing and receive zero-forcing correspond to completely mitigating self-interference via beamforming cancellation, although this is almost certainly not sum-rate-optimal. Instead, tolerating some self-interference would likely allow a device to more optimally serve users, especially if digital SIC is aiding beamforming cancellation.

Beamforming cancellation designs will certainly need robustness to self-interference channel estimation errors, especially since their effects will be magnified by the relative strength of self-interference. In addition, hybrid beamforming and per-antenna transmit power constraints may bottleneck the practical realization of a desired beamforming cancellation design, especially in frequency-selective settings. Importantly, understanding how beamforming cancellation interacts with nonlinear terms induced by the transmit PAs will shed light on the system design of mmWave full duplex. Investigating the severity of these many factors will be critical to fleshing out beamforming cancellation as a potential mmWave full duplex solution.

LEVERAGING USER SELECTION WITH BEAMFORMING CANCELLATION

Since successful transmission and reception are based on the channels between the full duplex transceiver and the distant users it is serving, the effectiveness of beamforming cancellation may be highly subject to the environment and the users being served. Users whose channels are aligned with the self-interference channel will make beamforming cancellation more costly since significant deviations from optimal half duplex strategies will be required to mitigate self-interference. When

a full duplex transceiver employing beamforming cancellation has the liberty of choosing or scheduling the devices with which it will communicate, the principle of user selection becomes a powerful tool for interference reduction. In essence, some transmit-receive pairings naturally afford more self-interference mitigation than others. However, we make the distinction that the mitigation afforded by two users is not based on *their* relative orthogonality since the full duplex device's transmit beam and receive beam are coupled by the self-interference channel.

We illustrate the power of user selection by referencing Fig. 2b. Fixing the transmit beam, for instance, will inflict some degree of output leakage onto the receive array. Candidates to receive from are transmitting users A, B, and C, which is done by beamforming along their respective solid vector. The shadow cast by each of these vectors represents the portion that will foster self-interference. Choosing the vector whose dotted shadow lies most orthogonal to the output leakage striking the receive array — user B in this case — will reject self-interference most naturally, reducing the spatial resources consumed to achieve full duplex.

Of course, this overly simplistic scenario and approach are not flawless, particularly in regard to fairness, since a greedy approach to maximize the sum rate would be to continuously transmit and receive to the “best” pair of users. Furthermore, asymmetric uplink/downlink demands across users and the dynamics of their channels (e.g., due to mobility) will impact how users are selected for service. Nevertheless, this toy example illustrates how user selection can be leveraged to improve beamforming cancellation, especially as the number of users available for selection grows.

COMBINING BEAMFORMING CANCELLATION AND ANALOG SELF-INTERFERENCE CANCELLATION

While beamforming cancellation could potentially replace the need for analog SIC, we also envision a use for both, considering the potential limitations and performance costs of beamforming cancellation alone. It may be possible that an analog SIC solution growing with the number of RF chains, rather than the number of antennas, could aid in achieving mmWave full duplex by supplementing beamforming cancellation [14]. Suppose an analog SIC filter is placed across the transmit and receive RF chains as shown in Fig. 4. It is important to note that in this case, analog SIC is driven by RF signals before the analog precoder, meaning it will not incorporate nonlinear terms introduced by per-antenna PAs and other nonidealities. The advantage of this architecture, however, is that the responsibility of mitigating self-interference is shared across beamforming cancellation, analog SIC, and digital SIC, which we illustrate in Fig. 5. Transmit beamforming cancellation would remain *solely* responsible for preventing per-antenna LNA saturation, while analog SIC and beamforming cancellation (transmit and receive side) would share the responsibility of addressing ADC saturation. Lastly, digital SIC would aim to cancel residual self-interference.

By taking this approach, analog SIC solutions of reasonable size can aid in mitigating self-interference and preventing receiver-side saturation.

Since successful transmission and reception are based on the channels between the full duplex transceiver and the distant users it is serving, the effectiveness of beamforming cancellation may be highly subject to the environment and the users being served.

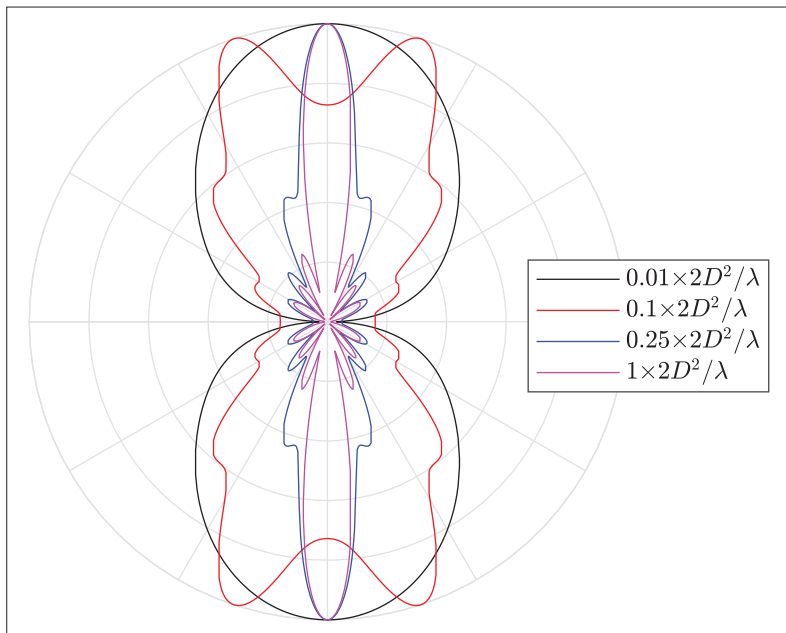


FIGURE 6. The normalized array factor of an eight-element half-wavelength uniform linear array at fractions of the far-field distance rule of thumb.

Furthermore, by reducing the responsibility of beamforming cancellation in mitigating self-interference, the full duplex transceiver can serve users more optimally, as illustrated in Fig. 3.

In frequency-selective settings, we imagine that this staged cancellation approach will be very attractive since it allows both beamforming cancellation and analog SIC to address the selectivity, offering the flexibility to trade frequency-selective beamforming for many-tap analog SIC filters. Finally, note that analog SIC does not need explicit knowledge of the over-the-air channel; it merely needs to know the significantly reduced effective channel from transmit RF chains to receive RF chains (i.e., $\mathbf{W}_{\text{RF}}^* \mathbf{H}_{\text{SI}} \mathbf{F}_{\text{RF}}$). This relatively small channel can be observed digitally with conventional estimation strategies, which are likely more reliable and frequent than estimation of the over-the-air counterpart (i.e., \mathbf{H}_{SI}).

CUSTOM ANALOG BEAMFORMING CODEBOOKS FOR MMWAVE FULL DUPLEX

To address channel estimation and initial access challenges and reduce complexity, practical mmWave networks have turned to codebook-based beam alignment [15], which is often executed as a search through an analog beamforming codebook for beams that work well (e.g., offer high SNR) between two devices. Since analog beamforming supplies the high beamforming gain on which mmWave communication relies, beams in an analog beamforming codebook generally have two properties:²

- Offer high beamforming gain (i.e., are highly directional)
- Collectively provide good spatial coverage to ensure a user's location within the service region does not inhibit it from being served

A codebook with entries satisfying these two properties are generally suitable for half duplex use cases but do not necessarily offer much resilience to self-interference.

Extremely narrow transmit beams, for instance, have the potential to inflict substantial self-interference on the neighboring receive array. This is because an “extremely narrow transmit beam” is extremely narrow in the far-field and not necessarily so in the near-field. Figure 6 shows the array factor of an 8-element half-wavelength uniform linear array at various ranges from the center of the array.³ At the far-field distance rule of thumb, we get the familiar sinc-like shape with a narrow main lobe in the broadside directions. Getting closer and closer to the array, the far-field approximation deteriorates, and we begin to see the effects of the near-field. Those exhibited in Fig. 6 become quite omnidirectional, leading us to see how a near-field self-interference channel can be difficult to avoid with conventional analog beamforming. Since we expect that the self-interference channel at mmWave may contain a significant near-field portion (not to mention reflections from the environment), we again can see that achieving mmWave full duplex is not as simple as merely “aligning the nulls” of the transmit and receive beams.

Thus, we anticipate research on custom beamforming codebooks tailored for mmWave full duplex to be a promising future direction. For these custom codebooks to replace those conventionally used for beam alignment, their beams will need to provide sufficient beamforming gain and coverage. Beams would also ideally reject near-field self-interference as well as any stemming from the far-field, if possible.⁴ While potentially a good starting place, it is unlikely that “off-the-shelf” codebooks would naturally reject self-interference sufficiently, and the design of a custom codebook with these properties would likely be difficult to perfect.

If the ideal codebook could be designed — offering sufficient isolation between all transmit and receive beam pairs while still providing high gain and adequate coverage — an mmWave transceiver could *blindly* operate in a full duplex fashion with little to no need for supplemental analog or digital SIC. Given that the design of such a custom codebook depends on the self-interference channel, it may be updated according to the channel dynamics or created based on the long-term statistics of the channel.

We expect successful custom codebook designs for mmWave full duplex to be a fast track to deployment since they could integrate into existing beam alignment schemes and are much simpler to execute (once designed) than the previously described methods. In addition to beam alignment, custom codebook designs like this could be used to simplify and accelerate general beamforming design and optimization, including that of beamforming cancellation.

CONCLUSION: TOWARD MAKING MMWAVE FULL DUPLEX A REALITY

We conclude by summarizing important research directions that will be required to mature mmWave full duplex from theory to concept and beyond to practice. Reliable characterization and modeling of the self-interference channel will provide a foundation on which future research can build. Following that, self-interference channel estimation strategies can be developed, which

² In hierarchical codebooks, we refer to codewords in the finest tier.

³ This plot is based on ideal near-field behavior (i.e., [10]), free of coupling and other various electromagnetic artifacts, which would further misshape the beams.

⁴ Environmental reflections may be unavoidable with a codebook-based approach given the need for good spatial coverage.

may exploit newfound structure or sparsity. Beamforming cancellation designs subject to the constraints imposed by hybrid beamforming can be explored to better understand what performance guarantees and receiver-side saturation requirements can be met in realistic environments. Investigating how beamforming cancellation can be supplemented by analog SIC will provide rich insights on how the two can jointly tackle self-interference, especially in frequency-selective settings. A thorough analysis of nonlinear self-interference would facilitate system-level studies involving digital SIC, analog SIC, and beamforming cancellation. Network-level analyses will indicate the power of user selection and the effects full duplex has on mmWave access and backhaul. Prototyping full duplex mmWave systems early on will be essential in identifying unexpected obstacles and steering future research. Drafting full-duplex-based protocols that integrate well with existing networks will be critical in its standardization and deployment.

REFERENCES

- [1] T. S. Rappaport *et al.*, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access*, vol. 1, 2013, pp. 335–49.
- [2] R. W. Heath *et al.*, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," *IEEE J. Sel. Topics Signal Processing*, vol. 10, no. 3, Apr. 2016, pp. 436–53.
- [3] A. Sabharwal *et al.*, "In-Band Full-Duplex Wireless: Challenges and Opportunities," *IEEE JSAC*, vol. 32, no. 9, Sept. 2014, pp. 1637–52.
- [4] E. Everett *et al.*, "SoftNull: Many-Antenna Full-Duplex Wireless via Digital Beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, Dec. 2016, pp. 8077–92.
- [5] T. Dinc *et al.*, "Synchronized Conductivity Modulation to Realize Broadband Lossless Magnetic-Free Non-Reciprocity," *Nature Commun.*, vol. 8, no. 11, Oct. 2017, p. 19.
- [6] A. Alkhateeb *et al.*, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems," *IEEE J. Sel. Topics Signal Processing*, vol. 8, no. 5, Oct. 2014, pp. 831–46.
- [7] F. Sohrabi and W. Yu, "Hybrid Beamforming with Finite-Resolution Phase Shifters for Large-Scale MIMO Systems," *Proc. IEEE SPAWC*, June 2015, pp. 136–40.
- [8] D. Korpi *et al.*, "Full-Duplex Transceiver System Calculations: Analysis of ADC and Linearity Challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, July 2014, pp. 3821–36.
- [9] Z. Xiao, P. Xia, and X. Xia, "Full-Duplex Millimeter-Wave Communication," *IEEE Wireless Commun.*, vol. 24, no. 6, Dec. 2017, pp. 136–43.
- [10] J.-S. Jiang and M. A. Ingram, "Spherical-Wave Model for Short-Range MIMO," *IEEE Trans. Commun.*, vol. 53, no. 9, 2005, pp. 1534–41.
- [11] X. Liu *et al.*, "Beamforming Based Full-Duplex for Millimeter-Wave Communication," *Sensors*, vol. 16, no. 7, July 2016, p. 1130.

- [12] K. Satyanarayanan *et al.*, "Hybrid Beamforming Design for Full-Duplex Millimeter Wave Communication," *IEEE Trans. Vehic. Tech.*, vol. 68, no. 2, Feb. 2019, pp. 1394–1404.
- [13] L. Zhu *et al.*, "Millimeter-Wave Full-Duplex UAV Relay: Joint Positioning, Beamforming, and Power Control," *IEEE JSAC*, vol. 38, no. 9, 2020, pp. 2057–73.
- [14] I. P. Roberts, H. B. Jain, and S. Vishwanath, "Equipping Millimeter-Wave Full-Duplex with Analog Self-Interference Cancellation," *Proc. IEEE ICC Wksp.*, Dublin, Ireland, June 2020, pp. 1–6.
- [15] J. Wang *et al.*, "Beam Codebook Based Beamforming Protocol for Multi-Gbps Millimeter-Wave WPAN Systems," *IEEE JSAC*, vol. 27, no. 8, Oct. 2009, pp. 1390–99.

BIOGRAPHIES

IAN P. ROBERTS [S'16] (ipr@utexas.edu) is a graduate 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 GenXComm, Inc., Amazon, Sandia National Laboratories, and Dynetics, Inc. 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 [F'13] is the Cockrell Family Endowed Chair in Engineering at the University of Texas at Austin, where his research focuses on wireless communication systems. He received the 2015 Terman Award, the NSF CAREER Award, and the 2019 IEEE Kiyo Tomiyasu technical field award. He is an ISI Highly Cited Researcher and has been co-recipient of 15 best paper awards. He holds a Ph.D. in electrical engineering from Stanford University.

HARDIK B. JAIN received his B.E. degree in electrical engineering from the Birla Institute of Technology and Science, Pilani, in 2011 and his M.S. degree in electrical engineering from the University of Texas at Austin in 2015. He is a graduate student currently working toward a Ph.D. degree in electrical engineering at the University of Texas at Austin. His research interest includes full duplex wireless transmission, design of wireless communication systems, and coding theory. His work on full duplex radios and self-interference cancellation technology is being commercialized to build a product at GenXComm Inc., where he serves as co-founder and CTO. His former industry experience includes work at Texas Instruments, Bangalore in 2011 and Broadcom India Research, Bangalore from 2011–2013.

SRIRAM VISHWANATH [SM'08] received his B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, India in 1998, his M.S. degree in electrical engineering from California Institute of Technology, Pasadena, California, in 1999, and his Ph.D. degree in electrical engineering from Stanford University, California, in 2003. He 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 AI/machine learning, information theory, networking, blockchain, and large-scale systems. He is the co-founder of multiple startups in the AI, networking, healthcare, and enterprise software domains.

Network-level analyses will indicate the power of user selection and the effects full duplex has on mmWave access and backhaul. Prototyping full duplex mmWave systems early on will be essential in identifying unexpected obstacles and steering future research. Drafting full-duplex-based protocols that integrate well with existing networks will be critical in its standardization and deployment.