RESEARCH ARTICLE

Augmented Envelope Neural Networks on RF-SoC for Digital Self-Interference Cancellation

UDARA DE SILVA, (Graduate Student Member, IEEE),
HIRUNI SILVA, (Graduate Student Member, IEEE),
AND ARJUNA MADANAYAKE, (Member, IEEE)
Electrical and Computer Engineering Department, Florida International University, Miami, FL 33199, USA

Corresponding author: Udara De Silva (udesi001@fiu.edu)

ABSTRACT This paper addresses the challenge of self-interference in in-band full-duplex radios, which can double the operational bandwidth and wireless channel capacity in 5G New Radio's sub-6 GHz spectrum. To achieve high isolation between simultaneously transmitted and received signals, the study explores envelope neural networks for self-interference cancellation. These networks model non-linear artifacts arising from both the transmit power amplifier and the receiver's low-noise amplifier. Trained model parameters are subsequently applied in real-time via a neural network-based digital signal processor to mitigate self-interference. A real-time prototype operating at 2.4 GHz, featuring direct-RF sampling at 4.096 GS/s and a 20 dBm transmit power through an external PA, was implemented using an AMD-Xilinx ZCU-111 RF-SoC. The system demonstrates digital self-interference cancellation exceeding 30 dB in real-time over a 32 MHz passband bandwidth, utilizing a novel augmented envelope neural network realized as a systolic array architecture.

INDEX TERMS IBFD, STAR, digital cancellation, machine learning, neural networks, RF-SoC.

I. INTRODUCTION

Modern radio standards, such as 5G NR, aim to connect not only people but also vehicles, drones, and household equipment using massive machine-type communications in machine-to-machine, vehicle-to-vehicle and vehicle-to-infrastructure networks [1], [2], [3]. Therefore, the sustainable growth of wireless communications and networking technology would only be possible if new technologies are invented and adapted to improve spectrum efficiency in spectral bands where access to spectrum is scarce. In particular, there is spectral scarcity in sub-6 GHz "legacy" bands, where bandwidth is at a premium. Although the mm-wave spectrum is abundant and the future of wireless relies on moving to mm-wave bands, the legacy frequencies remain extremely important for communications due to favorable physics for highly-scattering urban environments

The associate editor coordinating the review of this manuscript and approving it for publication was Pedro Miguel Cabral.

(e.g., low path-loss) [4]. As a consequence, there is and always will be much demand for the sub-6 GHz spectral band from commercial, public safety, and military systems. In-band full-duplex (IBFD) communication is a promising technology focused on addressing this impending spectrum scarcity challenge [5]. IBFD radios can simultaneously transmit and receive (STAR) different information-carrying signals over the same bandwidth [6], [7], thus doubling the bandwidth of operation and therefore the capacity of the wireless channel [8]. Such a doubling of capacity is a tremendous advantage for wireless communications in legacy spectral bands with massive applications for future 5G/6G systems [9], [10], [11].

The primary challenge faced by IBFD communication is that the transmit-signal (Tx) leaks to the receiver (Rx). Such leakage is known as self-interference (SI), which can shadow the signal of interest (SOI) and may even damage a sensitive receiver. In typical wide-area cellular communications, the mobile receiver sensitivity can be as low as −80 dBm,
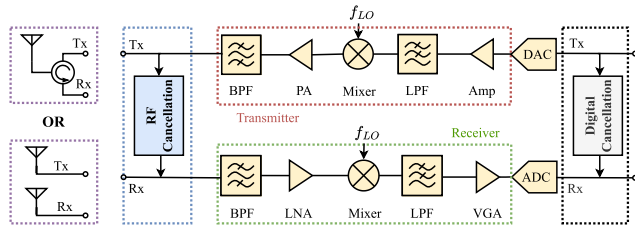
FIGURE 1. The main SI cancellation stages in IBFD radios: antenna cancellation (purple), RF cancellation (blue), and digital cancellation (black).

while the base station transmits power can be as high as 40 dBm [12]. In an example scenario, to achieve a 20 dB signal-to-interference and noise (SINR) ratio, the IBFD radio should achieve 140 dB SI cancellation. This is a very challenging engineering specification due to a host of technical reasons. Despite these challenges, previous work has shown encouraging performance using a combination of electromagnetic, analog, and digital cancellation [8], [13], [14], [15], [16]. In [15], we demonstrated a two-stage electromagnetic-analog SI canceller with up to 85 dB of SI cancellation.

In this work, we present a digital canceller that utilizes envelope neural networks for real-time SI cancellation that achieves up to an additional 30 dB SI cancellation using machine learning (ML). This digital canceller can be combined with an RF/analog canceller, as shown in Fig. 1 to achieve the target SI cancellation performance typically required in practical communication systems. This paper makes two primary contributions to the domain of IBFD systems. First, we introduce and evaluate three envelope neural network architectures designed for SI cancellation using simulation data. Second, we demonstrate a real-time digital hardware implementation using a system-on-chip (SoC) architecture, which facilitates real-time data acquisition, and augmented envelope neural network-based SI cancellation. Our experimental results confirm that the system achieves over 30 dB of digital SI cancellation and supports simultaneous transmission and reception. The rest of the paper is organized as follows. Section II presents a review of current SI cancellation methods. The effect of RF front-end components on SI is modeled in Section III. The proposed envelope models for SI cancellation is introduced in Section IV, which is followed by a performance comparison using simulation data in Section V. Section VI describes the hardware architecture for real-time digital SI cancellation. The experimental measurements are presented in Section VII, which also demonstrates the STAR capability using a remote quadrature phase shift keying (QPSK) transmitter as the SOI. The final section summarizes our findings and present possible future work.

II. REVIEW OF SI CANCELLATION

Research on IBFD radios can be loosely divided into three domains of realization: antenna, RF/analog [17], [18], [19], [20], [21], and digital techniques [22], [23], [24], [25].

Antenna cancellation can be achieved using differential feeding [26], [27] or polarization diversity [28]. In arrays or multiple input multiple output (MIMO) systems, antenna cancellation can be achieved by placing the Tx elements at a specific distance from the Rx elements, such that, the SI experiences destructive interference at the Rx antenna [16], [29]. In this work, we simply used Tx-Rx antenna separation as the front-end isolation technique. The typical approach to achieve RF cancellation is to use a ferrite circulator to isolate the Tx and Rx [30]. Recently, alternatives to ferrite circulators have been proposed that can operate on wider bandwidth, with low insertion loss, and with the ability to integrate into monolithic microwave integrated circuits (ICs) [31], [32], [33], [34]. We demonstrated how a sequentially switched delay line (SSDL) circulator achieves over 25 dB Tx-Rx isolation over a 1GHz bandwidth [35]. Another method to achieve RF cancellation is balanced duplexing [36]. This technique has shown even better isolation at the cost of poor insertion loss [37], [38], [39]. Tapped delay line filters can be used to generate an estimate of the SI signal, which in turn, is subtracted from the received signal to perform RF cancellation [15], [22], [40], [41]. Antenna- and RF/microwave circuit-based SI cancellation play a crucial role in IBFD radios because if the SI signal is much stronger than the received signal of interest (SOI) when it reaches the analog-to-digital converter (ADC), the SOI can be completely shadowed by quantization noise. Digital SI cancellation is useful to capture higher-order non-linearities that are very difficult to model using analog circuitry. Several approaches to achieve SI cancellation in the digital domain are reviewed next.

A simple but effective method to perform digital SI cancellation is to use an adaptive finite-impulse-response (FIR) filter [42], [43] where the coefficients are calculated using the least square error method. In [44], variable fractional delay filters are used to further improve the performance. The performance of FIR interpolation significantly reduces when the SI signal becomes more non-linear at high transmit power levels. As an alternative, non-linear models like memory polynomial [45], [46] and Hammerstein model [47] are used to model non-linear SI signals. In [48], spline-based interpolation is used to reduce the computation complexity of the Hammerstein model-based SI cancellers.

In recent years, there has been a growing interest in leveraging neural networks for digital self-interference (SI) cancellation [23], [49], [50], [51], [52], [53]. One notable approach, presented in [54], utilizes a real-valued time-delay neural network (TDNN) for performing digital SI cancellation. In [55], two modified structures for the TDNN are proposed, namely the ladder-wise grid structure (LWGS) and the moving-window grid structure (MWGS), resulting in significant complexity reduction. Additionally, in [56], the authors explore the utilization of dual-neuron neural networks, which effectively model memory effects, further enhancing the efficiency and performance of the SI cancellation process. Our research draws inspiration from

TDNN-based SI cancellation and proposes an approach that improves performance by incorporating envelope terms for self-interference modeling.

III. SELF-INTERFERENCE MODELING

Digital SI cancellation is done by subtracting a predicted SI signal from the received signal [25], [57]. Therefore, digital cancellation performance relies on the accuracy of the self-interference modeling. This section highlights the key contributors to self-interference and the challenges in accurately modeling these effects. For a more comprehensive study of self-interference modeling, the readers are referred to the dissertation by Korpi [58].

If we consider a perfectly linear RF front-end, the SI signal $y(t)$ can be expressed by Eq. (1), where $x(t)$ is the transmit signal and h_{21} is the unit impulse response of the mutual coupling between the Tx and Rx. In this work, two separate Tx and Rx antennas are used for the RF front end. In such cases, the mutual coupling can be expressed as the sum of the direct propagation path and other significant nearby reflected paths. Then $y(t)$ can be described by Eq. (2) where w_i and t_i is the attenuation and the delay of each propagation path [59].

$$y(t) = h_{21}(t) * x(t) \tag{1}$$

$$y(t) = \sum_{i=1}^{m} w_i \cdot x(t - t_i) \tag{2}$$

In reality, RF front-ends have analog impairments which introduce both linear impairments, as well as non-linear harmonic- and intermodulation- distortion to the transmit signal (baseband) $x(t) \in \mathbb{C}$. The dominant source of non-linear distortion is the power amplifier [60]. The prevalent issue of the mixer is the IQ imbalance which introduces an image component to the signal whose spectrum is inverted [61]. The output of the mixer $x^{IQ}(t) \in \mathbb{C}$ in the presence of IQ imbalance is given by Eq. (3) where $g_1^{mix}(\tau)$ and $g_2^{mix}(\tau)$ are I- and Q-phase impulse responses of the transmitter IQ mixer [47]. However, this effect can be ignored for RF systems with IQ calibration [62], [63] or systems that perform digital down-conversion using a numerically controlled oscillator (NCO); such is the system proposed in this paper.

$$x^{IQ} = g_1^{mix}(\tau) * x(t) + g_2^{mix}(\tau) * x^*(t) \tag{3}$$

In the presence of high-power jamming/blockers, the receiver's LNA or gain stages will cause signal distortion. However, the power amplifier (PA) is the main source of nonlinear distortion. PA non-linearities are introduced due to PA saturation and become more dominant when the transmit power is high since power efficiency reasons call for biasing into a nonlinear regime of operation. PA non-linearity can be modeled using the memory polynomial shown in Eq. (4) [64], [65]. An effective digital SI cancellation should be able to model PA output at both low and high transmit power levels.

$$x^{PA}(n) = \sum_{m=0}^{M} \sum_{k=1}^{K} a_{mk} \cdot x(n - m) \cdot |x(n - m)|^{k-1} \tag{4}$$

In section IV, we propose an envelope-based NN model to efficiently cancel the SI where the non-linear artifacts contaminating the SI is dominated by the transmit PA.

IV. PROPOSED ENVELOPE NN BASED DIGITAL CANCELLATION

In this section, we describe three variations of envelope neural network models for capturing non-linear distortions. The baseband SI canceller operates in discrete time $t = T_{clk}k \ t \in \mathbb{R}, \ k, \ k = 0, 1, \ldots$ and T_{clk} is the sample period. A memory-less PA non-linearity can be represented by AM/AM and AM/PM conversion functions [64]. This relationship can be expressed using Eq. (5), where $A(t) \in \mathbb{R}$ and $\theta(t) \in [-2\pi, 2\pi]$ represents the amplitude and phase of the input signal, respectively; furthermore, $G_A(\cdot)$ and $\phi_G(\cdot)$ are non-linear AM/AM and AM/PM conversion functions, respectively [64].

$$y(t) = G_A(A(t))A(t)e^{j[\theta(t) + \phi_G(A(t))]} \tag{5}$$

Eq. (5) can be rearranged as follows,

$$\begin{aligned}
y(t) &= G_A(A(t))e^{j[\phi_G(A(t))]} \cdot A(t)e^{j(\theta(t))} \\
&= f(A(t)) \cdot A(t)e^{j(\theta(t))} \\
&= f(|x(t)|) \cdot x(t)
\end{aligned}$$

where $x(t) = A(t)e^{j(\theta(t))} \in \mathbb{C}$ is the baseband input signal and $f(\cdot) \in \mathbb{C}$ is a complex-valued function that represents the non-linear distortion; here, $f(\cdot)$ depends on the PA type, and operational condition (i.e., DC bias point, temperature, and age) [66]. The dynamic nature of the wireless environment will introduce time-dependent changes to the SI signal. Therefore, an adaptive approximation function is required to model $f(\cdot)$. Let f_{nn} be an approximator for $f(\cdot)$ then, $\forall x(t) \in T$ Eq. (6) is satisfied where $T \in \mathbb{C}$ is the set of all complex baseband transmit symbols and $\epsilon \in \mathbb{R}$ is a positive tolerance parameter.

$$|f(x(t)) - f_{nn}(x(t))| \le \epsilon \tag{6}$$

A. ENVELOPE NEURAL NETWORK (ENN)

We employ a feed-forward network with a single hidden layer that contains a number of hidden neurons as the approximator. The output of such a network is given by Eq. (7) where w_i, and b_i represent the output weight and the bias of the ith hidden neuron, respectively [67]. Here, a_{ij} are the input weights between the jth input neuron and i^{th} hidden neuron, and φ is a non-linear activation function.

$$f_{nn}(x(t)) = \sum_{i=1}^{n} w_i \varphi(\sum_{j=1}^{m} a_{ij} x_j(t) + b_i) \tag{7}$$

From Universal Approximation Theory [68] it follows that $\forall \epsilon > 0 : \exists n \in \mathbb{N}$ such that Eq. (6) is satisfied. Therefore,

a feed-forward NN with a sufficiently large number of hidden neurons can approximate the non-linearity in a SI signal with an arbitrarily small tolerance. We need to however keep in mind that computations occur in real-time on a digital hardware processor, which inherently uses digital arithmetic having finite precision. For example, the processor may use the twos complement fixed-point format. Therefore, the non-linear effects of the computation arising from digital arithmetic must be negligible compared to the tolerance one is aiming to achieve. To ensure this, we will be using 32-bit precision throughout the implementation, to serve as a baseline for future reduced-precision digital hardware processors.

We so far had assumed a memory-less PA behavior. However, the output of a PA is dependent on the current input and past input waveforms due to memory effects, which is caused by electromagnetic coupling between the PA's output and the input ports. The input to the digital NN is fed by a tapped delay line (i.e., a form of digital FIR filter) to account for the memory effect. The above type of NN is referred to as a time-delay neural network (TDNN) [69]. The modified NN output is described by Eq. (8), where m represents the memory length of the PA, and where $\tau_j = T_{clk}j \in \mathbb{R}$ are time-delays which are integer multiples of the digital sampling/clock period.

$$f_{nn}(|x(t)|) = \sum_{i=1}^{n} w_i \varphi(\sum_{j=0}^{m-1} a_{ij}|x(t - \tau_j)| + b_i) \qquad (8)$$

Since $f_{nn}(\cdot)$ is an approximation for $f(\cdot) = G_A(\cdot)e^{j[\phi_G(\cdot)]}$, the output of Eq. (8) should be complex. Therefore input/output weights and biases should be complex-valued. Instead of using complex weights and biases, the network can be represented using a real-valued network [70]. Since our input to the network is real, we can simply treat the single complex output as two independent real outputs and represents the network using only real weights and biases. We will call this network the envelope model for the rest of the paper. The final output of the predicted SI signal is given by Eq. (9).

$$y_1(t) = f_{nn}(|x(t)|) \cdot x(t) \qquad (9)$$

B. AUGMENTED ENVELOPE NEURAL NETWORK (AENN)

We can absorb the multiplication in Eq.9 to the neural network input to generate a new model. To keep the model as a real-valued neural network, we will provide the real and imaginary parts of the $x(t)$ signal as two independent inputs. We will refer to this model as the augmented envelope model for the remainder of the paper. The model output is described in Eq. (10).

$$y_2(t) = f_{n2}(x(t)) \qquad (10)$$

where, for $\tau_j = jT_{clk}$

$$f_{n2}(x(t)) = \sum_{i=1}^{n} w_i \varphi(\sum_{j=0}^{m-1} a_{ij} \left| x(t - \tau_j) \right| + c_{ij} Re(x(t - \tau_j))$$

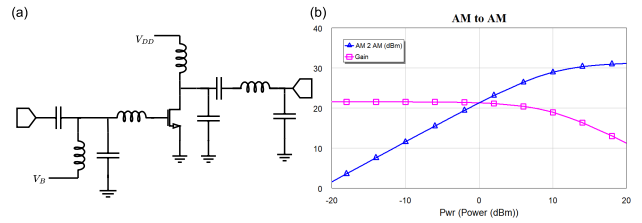$$+ d_{ij}Im(x(t - \tau_j)) + b_i)$$

In section V, we compare performance of both an envelope model and an augmented envelope model NN against the reference NN model proposed in [71].

C. AENN WITH PIECE-WISE LINEAR ACTIVATION (AENN+PWL)

Traditional neural networks use continuous functions such as sigmoid or hyperbolic tangent functions to achieve the non-linear activation operation. The relatively high computational complexity of such functions when realized in parallel digital hardware results in high consumption of both energy and logic resources. We propose using a simple piece-wise linear function inspired by the PA AM/AM curve as the activation function. Although a typical PA has multiple transistors, a single transistor amplifier (Fig. 2(a)) can be used as a simplified model to understand the non-linear behavior. The relationship between output and input power for a common-source amplifier [72] is shown in Fig. 2(b)[blue curve]. The P_{out}/P_{in} relationship is linear at low input power levels and starts saturating after the 1 dB compression point. Using this characteristic as an inspiration, a piece-wise-linear (PWL) function described by Eq. (11) is proposed for the activation function where n is the weighted input to the neuron. The proposed function acts as a linear activation function for small input values and then saturates (i.e., the gradient reduces to $0.125 \ll 1$) after exceeding the threshold value δ, a hyperparameter set to 0.6 in all our simulations and experiments. The gradient 0.125 is chosen because the multiplication can be implemented using shift operators, thus simplifying the digital implementation.

$$\varphi_p(n) = \begin{cases} 0.125 \cdot n - 0.875 \cdot \delta & n \leq -\delta \\ n & -\delta < n \leq \delta \\ 0.125 \cdot n + 0.875 \cdot \delta & n > \delta \end{cases}$$

$$(11)$$

It is shown in [68] that piece-wise linear functions satisfies the universal approximator property. For instance, the piecewise linear function known as rectified linear unit (ReLU) function is said to learn faster and perform more efficiently [73]. Therefore, the proposed function can be considered a valid activation function for the augmented
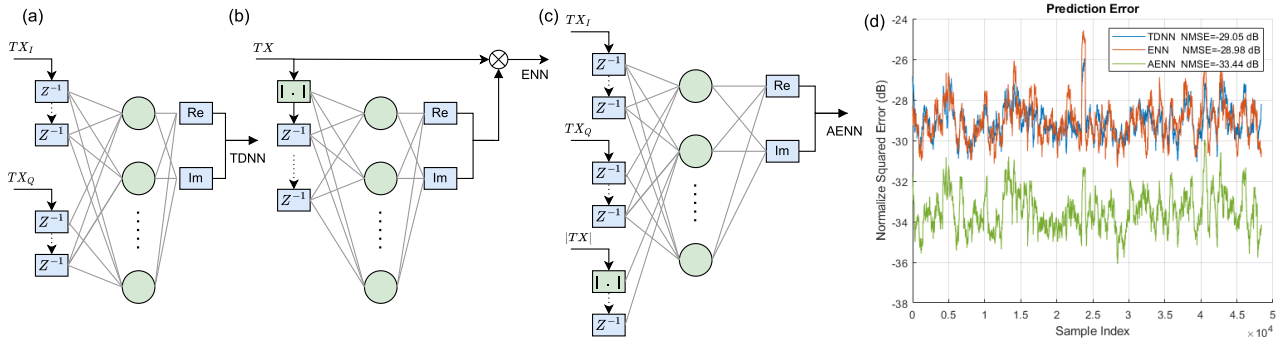
FIGURE 3. (a) TDNN model [71] (b) ENN model (c) AENN model (d) Prediction error comparison of the three models that shows NMSE of -29.0 dB, -29.0 dB, and -33.4 dB for the (a), (b), and (c) models, respectively.

envelope neural networks. The output of this model is given by Eq. (12). Our simulation results show that this activation function achieves better performance when compared with hyperbolic tangent activation function while simultaneously reducing the hardware complexity.

$$y_3(t) = f_{n3}(x(t)) \qquad (12)$$

where, for $\tau_j = jT_{clk}$

$$f_{n3}(x(t)) = \sum_{i=1}^{n} w_i \varphi_p \left(\sum_{j=0}^{m-1} a_{ij} \left| x(t - \tau_j) \right| + c_{ij} Re(x(t - \tau_j)) \right. $$
$$\left. + d_{ij} Im(x(t - \tau_j)) + b_i \right)$$

V. SIMULATION RESULTS

In our simulations, a MATLAB dataset encompassing the measured input and output signals of an NXP Airfast PA was employed [74]. Although there are additional non-idealities that affects the SI signal, such as IQ imbalance and oscillator phase noise, this dataset was chosen for our simulations with the assumption that the PA is the primary source of signal distortion [60]. The transmit waveform is an OFDM signal with 100 MHz bandwidth. MATLAB's statistics and machine learning toolbox is used to train and test the models. The comparison is done for three models: TDNN, ENN, and AENN. All models have a single hidden layer with 20 hidden neurons and a memory length of 10.

Fig. 3(d) shows the normalized mean squared error (NMSE) comparison of the three models. AENN model has the best performance of −33 dB, while the other two models have almost the same performance of -29 dB. An important factor to consider in the comparison is the number of learnable parameters since it determines the model complexity. Since all models have the same number of hidden neurons, the difference in learnable parameters is proportional to the number of inputs to the neural network. The number inputs to the three models are 1, 2, and 3 for the ENN, TDNN, and AENN respectively. Since all three models have 20 hidden neurons and a memory length of 10, the number of learnable parameters in the hidden layer is equal to 200, 400, and 600 for the ENN, TDNN, and AENN

models. The ENN model has only 0.07 dB increase in NMSE despite using 200 parameters less than the TDNN model, which makes it more suitable for deployment in low-power IBFD radios.

A. SI PREDICTION ERROR VS. TRANSMIT POWER

The SI signal can be modeled using a finite impulse response (FIR) digital filter when the transmitter is linear. Linear FIR filters typically require fewer hardware resources than an NN-based nonlinear digital predictor for approximating the SI signal. However, real-world transmitters exhibit significant non-linearity, resulting in a non-linear mapping between the transmit signal and the received SI signal. We first explore the advent of non-linear distortion as a function of system parameters using computer simulation of microwave components. To wit, a circuit envelope simulation [75] was carried out using the vendor-provided simulation model parameters(i.e, gain, 1dB compression point, and IP3) of the Mini-Circuits ZX60-83LN-S+ amplifier, which was chosen for purposes of experimental verification, to compare the performance of FIR interpolation with the proposed augmented envelope neural network for different output power levels. The simulated results of predicted SI using FIR interpolation (blue), augmented envelope neural network with hyperbolic tangent activation function (green), and piecewise-linear activation function (red) are shown in Fig. 4.

The simulation results in Fig. 4 show that the ability of the FIR interpolation to cancel SI degrades when the output power level increases; on the contrary, envelope neural network can maintain a similar performance across a wide range of transmit power levels. This occurs because the PA output becomes more non-linear at higher output power levels, which can be verified by observing the shape of the constellation of the SI signal. Simulations further show that the augmented envelope model with hyperbolic tangent activation functions performs rather poorly compared to linear digital FIR interpolation when the PA operates in its linear region. However, the augmented envelope model with the piecewise linear activation function consistently performs better than the other models in both linear and non-linear regions.
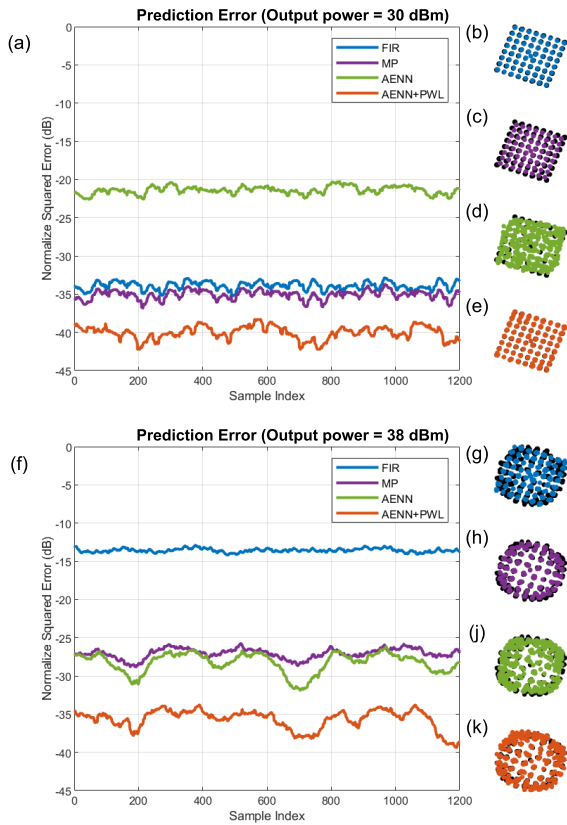
FIGURE 4. Prediction error of the FIR interpolation (blue), memory polynomial (purple), augmented envelope model with hyperbolic tangent activation (green) and piece-wise linear activation (red) when the output power is 30 dBm (a) and 38 dBm (f). Predicted constellation points by the four methods are shown on the right when the output power is 30 dBm (b)(c)(d)(e), and 38 dBm (g)(h)(j)(k). In the constellation plots, the actual received SI is shown in black.

The normalized mean squared prediction error for different output power levels are shown in Fig. 5. The digital FIR linear interpolation works better at low PA power levels but quickly deteriorates when the PA is close to saturation. The first envelope model performs better than FIR interpolation with both activation functions but suffers from the same performance degradation at high power levels. The augmented envelope model shows more resilience for high output power than other models and delivers the best performance in both linear and non-linear regions. Based on these results, we decided to implement a digital design to perform real-time SI cancellation using an augmented envelope model. The real-time implementation of AMD-Xilinx RF-SoC ZCU-111 is described in the next section.

VI. DIGITAL IMPLEMENTATION OF SI CANCELLATION USING AMD-XILINX ZCU-111

We explored a real-time implementation of the proposed augmented envelope neural network-based digital canceller using AMD-Xilinx ZCU-111 RF-SoC, which is a field programmable gate array (FPGA) having integrated ADC
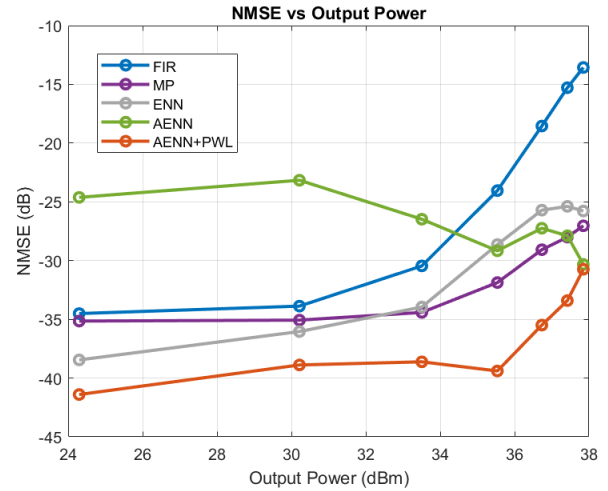

FIGURE 5. Normalized mean squared error (NMSE) plotted against the output power for digital FIR interpolation (blue), memory polynomial (purple), envelope model with hyperbolic tangent activation (gray) and augmented envelope model with hyperbolic tangent activation (green), piecewise linear activation (red).

and digital to analog converters (DACs). Overall system architecture on the RF-SoC is shown in Fig. 6(a).

A. OVERVIEW OF RF-SOC TECHNOLOGY

The Zynq UltraScale+ RF-SoC ZCU-111 board serves as a versatile platform for high-performance RF applications, including wireless communication and radar systems [76], [77], [78], [79]. AMD-Xilinx ZCU-111 RF-SoC features 8 spatial channels for ADCs with 12-bit precision and sample rates up to 4.096 GS/s, alongside another 8 channels for DACs with 14-bit precision at sample rates up to 6.554 GS/s [80]. The system also incorporates 8 soft-decision forward error correction units (SD-FECs) for enhanced reliability. The architecture includes an embedded ARM A53 processor, facilitating system-on-chip designs. Complemented by Arm Cortex-A53 and Cortex-R5 subsystems, the platform integrates UltraScale+ programmable logic, offering unparalleled signal processing bandwidth in a digital signal processing environment.

B. EXPERIMENTAL SETUP

The Tx and Rx blocks are 64 quadrature amplitude modulators (QAM) that encode and decode a text message. A preamble that uses the barker code [81] of length 13 is added to every frame for synchronization purposes. The frame size is set to 56 symbols. The baseband signal is up-converted to a RF carrier centered at 2.4 GHz in the digital domain using the inbuilt numerically controlled oscillator (NCO) that is available inside the RF-SoC. The DAC output is fed into the PA (Mini-Circuits ZHL-2W-63-S+). The amplified Tx signal is then fed to a 2.4 GHz panel antenna consisting of a static array of patch antenna elements. The receiver is fed by an identical patch antenna array. The receiver uses the Mini-Circuits ZRL-2400LN+ LNA
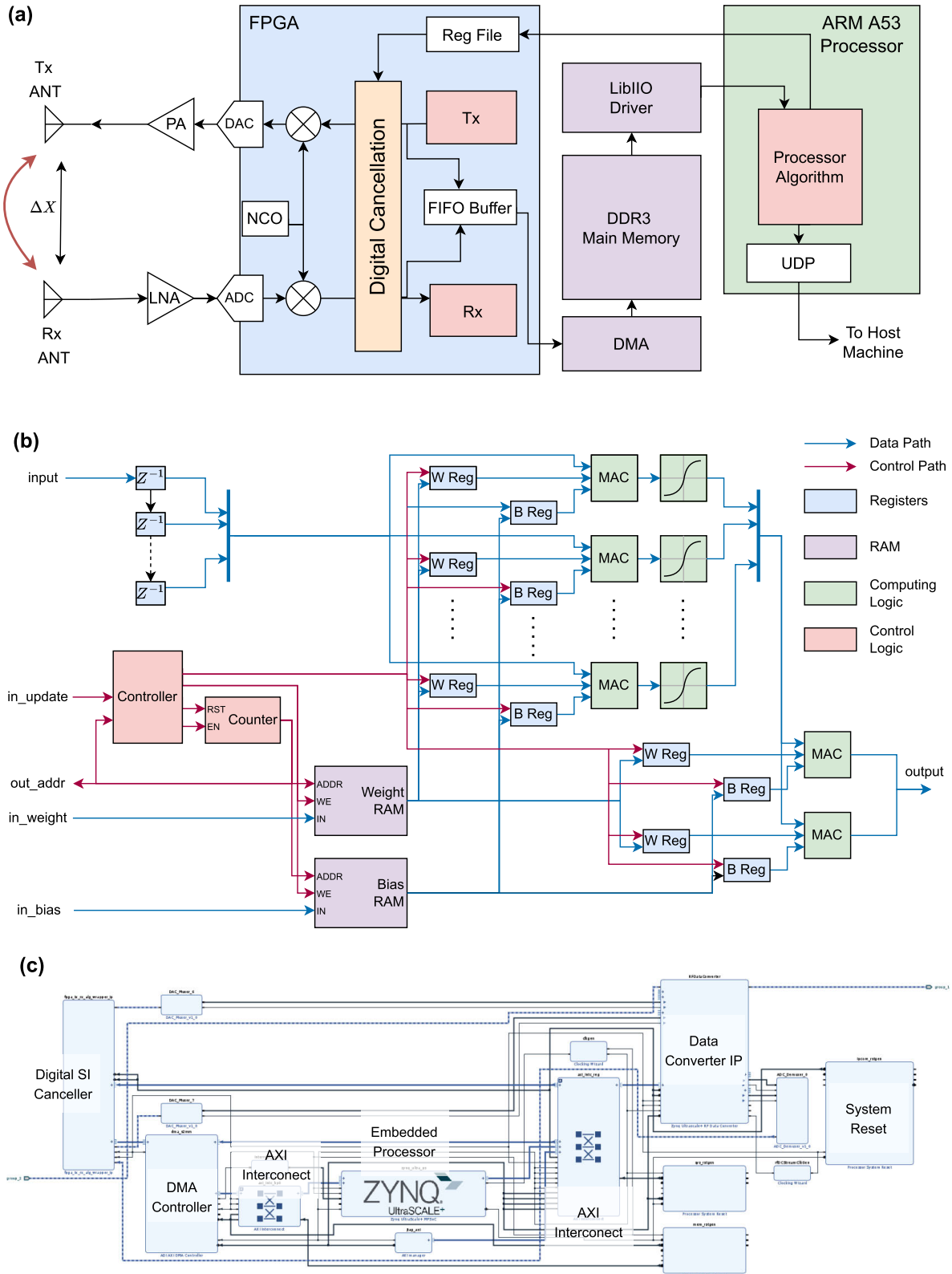
FIGURE 6. (a) System on chip architecture for the IBFD radio with digital cancellation (b) Digital canceller design that performs neural-network-based inference (c) Vivado block diagram of the system.

	TDNN	ENN	AENN
Multipliers	16	16	16
Adders	345	215	475
Registers	529	420	638
Multiplexers	359	289	429
Shift operators	60	40	80
Critical path (ns)	18.65	18.65	18.65

and performs digital down-conversion using the NCO. The experiments do not utilize multi-level analog cancellation because the objective is to evaluate the performance of the SI cancellation at the baseband level using proposed NN-based non-linear canceller operating in real-time. We however stress that eventual communication system examples would require multi-level SI cancellation across electromagnetic, analog and digital domains.

C. DIGITAL ARCHITECTURE

The NN processor architecture is shown in Fig. 6(b). In the digital design, we limit the number of hidden neurons to ten. Since the network has many tunable parameters, a digital controller is designed to serially feed the NN weights and bias values. The controller will start the weight update routine when the processor signals the availability of new weights by setting the in_update signal. The internal counter requests the weight and bias values by their address and stores these in the internal block random access memories (RAMs). Next, the controller will use a bit-serial connection to store these values in weight and bias registers. The main computation blocks of the design are the multiply and accumulate (MAC) unit and activation function. Implementation of the activation function is reduced to a comparator, multiplexer, and bit-shifter because of the proposed low-complexity activation function. These are realized on the RF-SoC's programmable logic fabric as a field programmable custom computing machine. The parallel MAC units form a systolic processor which perform vector operations at clock rate $F_{clk} = 1/T_{clk}$; this part of the digital design consumes most of the computational resources pertaining to the digital design. The complexity of the MAC unit grows in proportion to the input size, which can be noticed from the resource utilization Table 1. These results are based on the high-level synthesis report from MATLAB HDL coder [82]. The generated design is then used in AMD-Xilinx Vivado for FPGA implementation. The resource utilization of the implemented design showed 27k, 15k, and 36k configurable logic block (CLB) utilization for TDNN, ENN, and AENN models respectively.

The NN model needs to be periodically updated because of the dynamic nature of the wireless environment. The embedded processor is preferred for training because it can perform complex floating-point arithmetic effectively. However, this requires low-latency access to IQ data samples from the processor. The FPGA is running at 16 MHz, which

means it will produce IQ samples every 62.5 ns. Although the ARM processor is clocked at 1 GHz, an application running in the processor takes several milliseconds to access the data stored in RAM. Therefore, two FIFO buffers are created to store 2048 samples of Tx and Rx IQ data, which will be transferred to dual data rate-3 (DDR3) RAM on the ZCU-111 RF-SoC prototyping board. Multiple buffers are stored in the RAM and synchronously transferred as a single data frame using a batch transfer to the processor in order to perform the NN-training. Furthermore, the Tx and Rx frames are communicated to a host computer running Linux via 1 Gbps ethernet connection running the user datagram protocol (UDP) protocol. This ethernet based data link is used to verify the SI cancellation performance on the Linux host computer.

VII. REAL-TIME EXPERIMENTAL DIGITAL SI CANCELLATION RESULTS

A. EXPERIMENTAL HARDWARE SETUP

The experimental setup is shown in Fig. 7(a). A 20 dB attenuator is used to limit the radiated power to 20 dBm - selected to emulate the transmit power of a typical WiFi access point. The AMD-Xilinx ZCU-111 RF-SoC device is utilized as the digital signal processing platform to generate the baseband transmit symbols, and receive the signals containing SI such that the proposed algorithm can be applied on a per clock-cycle basis. A Mini-Circuits ZHL-2W-63-S+ is used as the PA for the transmitter and a Mini-Circuits ZRL-2400LN+ is used as the LNA for the receiver. The measured S-parameters of Tx and Rx antennas and their mutual coupling are shown in Fig. 7(b). The S-parameters demonstrate excellent power matching, and mutual coupling between the Tx and Rx arrays at about −30 dB (nominal) which sets the level of SI in the received signal.

B. MODULATION AND WIRELESS CHANNEL

For the purpose of demonstration of augmented envelope NN-based SI cancellation, we chose to modulate the Tx signal with random binary information using 64 quadrature amplitude modulation (64-QAM). The modulated Tx signal is used in full-duplex mode together with another quadrature phase shift key (QPSK) signal is applied to the remote transmitter. The remote transmitter communicates to the full-duplex system to form its Rx SOI; this signal is an over-the-air wireless signal. The remote source Tx uses a horn antenna as shown in Fig. 7(c) having gain 10 dBi, and transmit power of 20 dBm.

C. DIRECT-RF DIGITIZATION

ADC resolution is crucial for effective digital SI cancellation. Generally speaking, we analyze the ADC requirements for advanced modulation used in high-capacity communications. Let the Tx signal contain M-nary QAM loaded orthogonal frequency division multiplexing (OFDM) modulation having a Fast Fourier Transform (FFT) of discrete frequency N-bins.
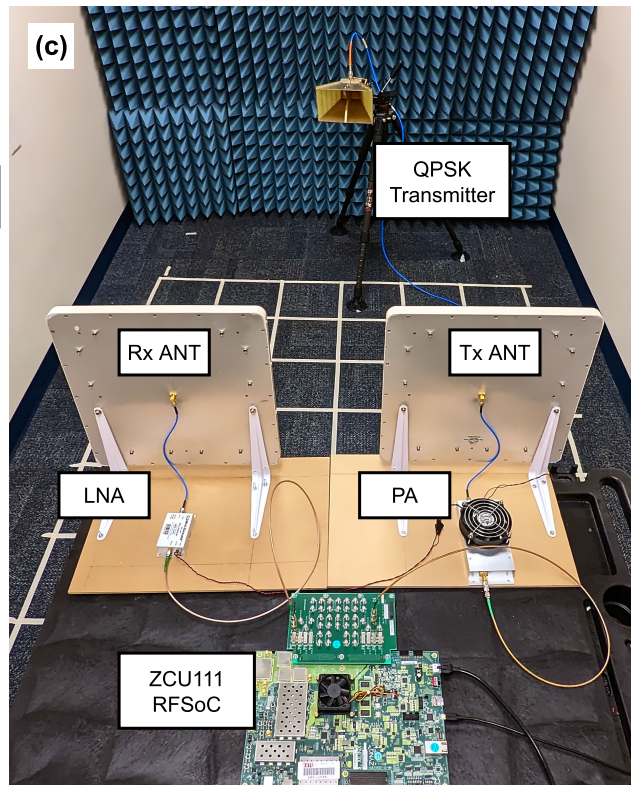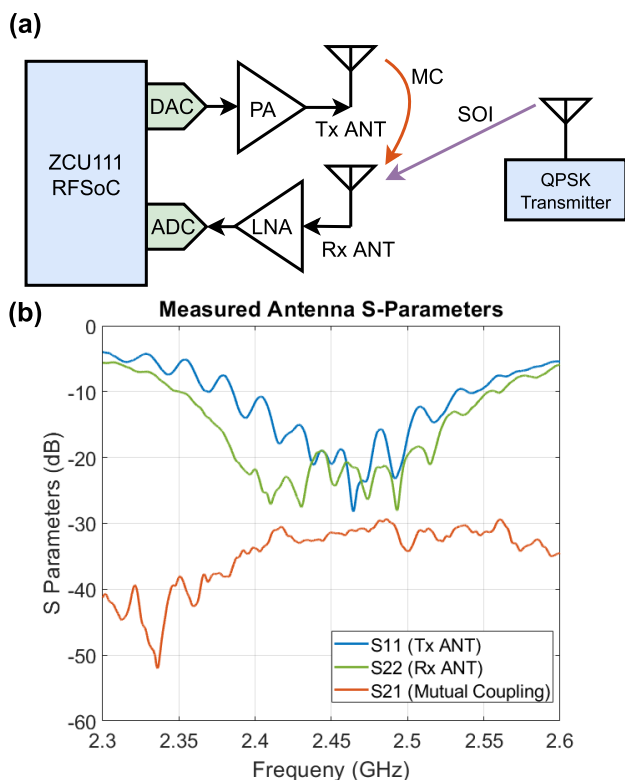
FIGURE 7. (a) Diagram of the experimental setup for verifying the digital canceller performance (b) Measured antenna S-Parameters of the Tx and Rx Antennas (c) Photograph of the experimental setup.

The dynamic range of the Tx signal is equivalent to that of an $\log_2 M + \log_2 N$ bit quantized signal. Recall that the SNR of an n-bit ADC is approximately $1.76 + 6n$ dB. Let the desired SINR of the remote Rx SOI that is required for error-free demodulation by at least γ dB where $L = \lceil ((\gamma - 1.76)/6) \rceil$ bits. Therefore, the ADC must quantize the baseband signals (assume real-valued channels) at $B = \log_2 M + \log_2 N + L$ effective number of bits (ENOB) in order that digital SI cancellation may subtract the Tx signal without corrupting the SOI.

For example, for $N = 2048$ bin FFTs loaded by 64-QAM as commonly found in 4G cellular communications, and assuming there is a remote site transmitting QPSK with the desired SINR at (say) 20 dB at the FD Rx, we need to quantize at a minimum resolution of $B = 6 + 11 + (20 - 1.76)/6 \approx 19$ ENOB at the ADC for SI cancellation to be viable. The ZCU-111 has RF ADCs have 12-bit outputs. However, its ENOB is less than 12 bits as the real-world performance of RF ADCs is poorer than the number of physical bits at the output port due to non-linearity in the internal circuitry: clearly, the ADCs on the RF-SoC do not have sufficient ENOB resolution for effective SI cancellation in the above case. Consider a second example: For 64-QAM at the Tx with no OFDM, and L=12 from the ADC of which 6-bits are required to accommodate the 64QAM constellation. Therefore, there is $L_{SIC} = 6$-additional ENOB worth of dynamic range left for SI cancellation. The best-case SIR achievable via digital SI reduces to about $1.76 + 6L_{SIC} \approx 37.76$ dB: the best-case

performance that can be achieved. Unless ENOB is very large, ADC performance is dominated by quantization noise, and therefore, ADC thermal noise has been neglected.

In the subsequent sections, we show our real-time measurements reach 35 ± 0.7 dB: *only marginally below the expected best-case SINR.* The loss of performance (about 2 dB) is likely due to thermal noise in the ADC, antenna/microwave component mismatches, or experimental/measurement errors.

D. ADC/DAC RATES AND POLYPHASE SAMPLE RATE CONVERSION

The DAC operates at a sample rate of 4.096 GS/s via four parallel time-interleaved streams having sample rates of 1.024 GS/s. Each DAC sample stream operates at a clock rate of 128 MHz using an up-sample factor of ×8 that yields the 1.024 GS/s stream rate. The digital design for processing 64QAM modulation is targeted at 16 MHz clock rate (32 MHz passband bandwidth), yielding a data rate of 96 Mbps (maximum). The processing core updates samples at the rate of 16 MHz, which is applied to the 128 MHz rate DAC sample stream using over-sampling and interpolation using sample copying (i.e., zero-order hold). Up-conversion to 2.4 GHz occurs within the DAC core using a polyphase NCO.

On the ADC side, the RF-ADC samples at 4.096 GS/s and DDC operating via an NCO produce the down-shifted signal in a multi-rate format. The baseband samples at 4.096 GS/s
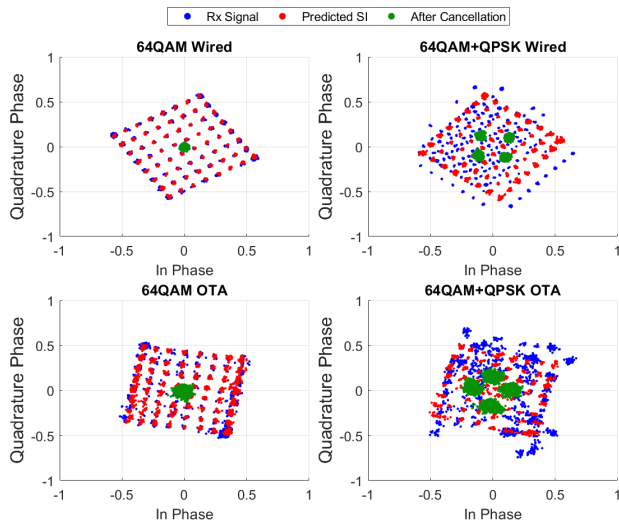
FIGURE 8. Constellation diagrams of (a) direct wired connection from Tx to Rx without using PA and LNA, (b) the signal of interest with QPSK modulation (at remote transmitter) is combined with the Tx signal using a wideband microwave/RF power splitter and RF feed to the Rx using a 50-ohm transmission line based wired connection, (c) Over-the-air (OTA) tests using the setup shown in Fig. 7 without using QPSK transmitter, (d) OTA test with the QPSK transmitter operating at 20 dBm output power.

FIGURE 9. Measured digital SI cancellation for consecutive frames using wired loopback (blue) and OTA testing (red) in the absence of SOI.

are split into 4 parallel streams in an internal decimator core to produce a single stream containing IQ baseband samples having ENOB 12-bits of precision operating at 1.024 GS/s. The stream is further decimated in a polyphase signal processing block available within the RF-ADC unit to yield a sample of 128 MS/s. I.e., the receiver produces two streams (for I and Q) each sampled at 128 MHz and having baseband bandwidth up to a maximum of 64 MHz. However, due to guard bands and filtering requirements, our example is limited to 16 MHz and is produced via a rate-transition block within the RF-SoC. The final data stream for processing consists of two 12-bit streams each updated at a rate of 16 MHz. These streams are applied to the machine learning processor for real-time inference using the augmented envelope NN. The machine learning processor may operate up to a maximum clock rate of about 50 MHz although this example requires only a 16 MHz clock frequency. At the moment, for testing purposes, the measured performance was estimated off-line by streaming IQ data to the host PC using the universal datagram protocol (UDP) packet engine running on the RF-SoC, due to the need for packet synchronization between the predictions of the ML cores and the transmit waveform available within the RF-SoC.

E. MEASURED PERFORMANCE

The real-time measurements of the digital SI cancellation is shown in Fig. 8. To create a reference to compare the effect of the RF front-end on the cancellation, (a)(b) measurements were made with a direct wired connection through a 50-ohm transmission line and a power splitter to combine the intended SOI which uses QPSK modulation. In the absence of SOI
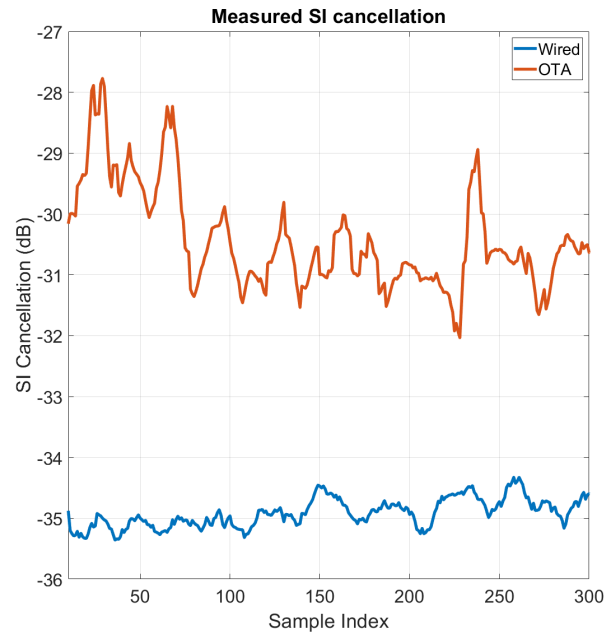
(Fig 8(a)(c)), the received signal only contains the self-interference component; i.e., in this wired configuration, the interference picked up by the Rx antenna, including WiFi signals in the real-world environment, are removed, thereby allowing a baseline best-case performance graph to be established. The measurement results show that the predicted SI signal (green) mostly overlaps with the actual received signal (blue). The difference between the machine learning predicted (from the augmented NN-based canceller) and measured SI from the Rx antenna seems quite negligible. To be specific, the error between the two waveforms (red) is centered at zero with a very small variance for the wired (Fig 8(a)) and OTA (Fig 8(c)) tests. The calculated average error is -35 dB and -31 dB for the wired and OTA tests respectively. The SI cancellation performance variation over the consecutive frames is shown in Fig. 9 which shows a larger variation in OTA testing due to the noise and interference in the real world "in the wild" wireless environment. The wired connection shows a best-case digital SI cancellation of about 35 ± 0.7 dB, and "in the wild" digital SI cancellation for OTA experiments around 30 ± 2 dB, respectively.

A performance comparison with other digital cancellation methods is shown in Table 2. The comparison of the digital cancellation performance independent of the RF front-end is challenging since different front-end cancellation techniques introduce different distortions to the received signal. Therefore, for this comparison, we selected published work that uses one of the two simple measurement setups: directly wired loopback (Wired), Tx-Rx antenna separation (OTA). The work [83] is used to provide a reference for

TABLE 2. Performance comparison of this work with other published work.

	Technique (# of parameters)	Setup	Platform	Tx Power (dBm)	Bandwidth (MHz)	SIC (dB)	M
This work	Augmented Envelope NN (640)	OTA	ZCU111 RF SoC	20	32	31	53.3
		Wired				35	55.5
[84]	FIR (N/A)	Wired	WARP v2	20	10	30	16.5
[85]	Parallel Hammerstein (189)	Wired	PXIe-7972	20	20	35	34.7
[59]	FIR (10)	OTA	KC705	20	10	18.6	14.7
	Hammerstein (15)					18.9	14.8
	Winner (155)					22.4	15.3
[86]	Augmented Spline-Based Hammerstein (524)	OTA	USRP	18	10	40	17.9
[83]	TDNN (550)	Dataset	N/A	10	10	44	17.2

neural-network-based digital SI cancellation which use a dataset as the input waveforms.

The work compared in Table 2 operates in 2.4 GHz band except for [59] which operates at 260 MHz band. An important point to note is that models such as Hammerstein, Winner, and MP require the computation of high-order terms of the input signal, whereas, the proposed method relies only on the input signal and its envelope. The performance of a SI canceller varies based on the operating bandwidth, transmit power and amount of SIC [61]. To compare our measured results with those published by other researchers, we define a unitless metric as $M = B \cdot e^{(P_T+SIC)/100} \geq 0$ where B is the bandwidth in MHz, P_T is the FD transmit power in dBm, and SIC is the best-case cancellation in dB. We defined our own metric as there does not seem to be a standard metric that takes into account bandwidth, transmit power, and SI cancellation. Although this metric enables comparison of different methods, it is not universally applicable as the importance of each factor varies depending on the application. The metric M linearly associates performance to bandwidth while assigning exponentially increasing levels of difficulty (performance) for both the transmit power level and SI cancellation - since these system parameters typically increase in dB increments in system design. The metric M shows increased values for better performance; i.e., bigger means better.

VIII. POWER VS. PERFORMANCE TRADEOFFS
The SI cancellation stages tradeoff power to achieve higher cancellation. In digital cancellation, the canceller increases the digital circuit complexity which leads to an increase in power consumption. Our primary objective in this work is to demonstrate the in-band full-duplex communication capability in base stations, which does not have a stringent power budget limitation compared to mobile devices. Implementing SI cancellation in mobile devices is more challenging due to the stringent power budget limitation.

In order to adopt SI cancellers in mobile devices they should have relatively low complexity. Our ENN-based canceller (Section. IV.A) was shown to achieve higher isolation with a shallow network (i.e., fewer parameters and thus lower complexity) compared to the TDNN-based canceller proposed in [71] (Section V). Also, ENN is capable of capturing both linear and nonlinear components

of the self-interference which eliminates the requirement to implement two different canceller modules as proposed in [56]. However, optimizing the digital canceller for mobile devices requires further investigation because the coupling between the transmitter and receiver is quite different compared to base stations; this will be one focus for future work.

IX. CONCLUSION
In-band full-duplex technology is a promising candidate for improving the spectral efficiency of 5G NR and other future wireless systems. In this paper, it was proposed that an augmented envelope neural network with an activation function inspired by the PA transfer function provides practically relevant levels of SI cancellation in the digital baseband across both linear and non-linear ranges of PA operation. Experimental results from a real-time custom hardware processor realized on the AMD-Xilinx RF-SoC showed good performance within about 2 dB of the theoretical best case of 37.76 dB. Further, the measurements for the over-the-air testing in a real-world spectral environment (laboratory space with multiple 2.4 GHz WiFi routers) showed over 30 dB SI cancellation. A remote QPSK transmitter operating at 20 dBm at a distance of about 2m from the FD system was used to demonstrate the capability of the implemented system to simultaneously receive while transmitting in the same frequency band. The implementation is done utilizing the RF ADCs and ARM A53 embedded processor of the AMD-Xilinx ZCU-111 RFSoC platform. In the future, we will utilize RF and antenna cancellation to build an IBFD radio capable of wide-area IBFD communication that has over 100 dB SI cancellation where the proposed augmented envelope NN-based digital canceller will be implemented at the baseband in the digital software-defined radio backend to achieve up to 30-35 dB of cancellation in addition to multi-level SI cancellation in the electromagnetic and analog domains.

REFERENCES
[1] S. Ali, N. Rajatheva, and M. Latva-aho, "Full duplex device-to-device communication in cellular networks," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2014, pp. 1–5.
[2] M. H. C. Garcia, A. Molina-Galan, M. Boban, J. Gozalvez, B. Coll-Perales, T. Sahin, and A. Kousaridas, "A tutorial on 5G NR V2X communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1972–2026, 3rd Quart., 2021.
[3] P. Pourbaba, K. B. S. Manosha, S. Ali, and N. Rajatheva, "Full-duplex UAV relay positioning for vehicular communications with underlay V2V links," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, Apr. 2019, pp. 1–6.
[4] J. Deng, O. Tirkkonen, R. Freij-Hollanti, T. Chen, and N. Nikaein, "Resource allocation and interference management for opportunistic relaying in integrated mmWave/sub-6 GHz 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 94–101, Jun. 2017.
[5] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.
[6] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3025–3041, Jul. 2019.

[7] M. Kohli, T. Chen, J. Welles, M. B. Dastjerdi, J. Kolodziejski, M. Sherman, I. Seskar, H. Krishnaswamy, and G. Zussman, "Remote experimentation with open-access full-duplex wireless in the COSMOS testbed," in *Proc. 26th Annu. Int. Conf. Mobile Comput. Netw.* New York, NY, USA: Association for Computing Machinery, Sep. 2020, pp. 872–874, doi: 10.1145/3372224.3417324.

[8] S. Bojja Venkatakrishnan, E. A. Alwan, and J. L. Volakis, "Wideband RF self-interference cancellation circuit for phased array simultaneous transmit and receive systems," *IEEE Access*, vol. 6, pp. 3425–3432, 2018.

[9] C. De Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, and M. Liyanage, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 836–886, 2021.

[10] Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, "AI and 6G security: Opportunities and challenges," in *Proc. Joint Eur. Conf. Netw. Commun. 6G Summit*, Jun. 2021, pp. 616–621.

[11] A. J. Alqaraghuli, A. Singh, and J. M. Jornet, "Compact high-gain dual-band antenna for full-duplex terahertz communication in CubeSat mega-constellations," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, Dec. 2021, pp. 1827–1828.

[12] M. Barazzetta, D. Micheli, P. Gianola, F. Moglie, and V. M. Primiani, "4G-LTE base station output power estimation from statistical counters during over-the-air tests in reverberation chamber," in *Proc. Int. Symp. Electromagn. Compat.*, Sep. 2014, pp. 284–289.

[13] S. Khaledian, F. Farzami, B. Smida, and D. Erricolo, "Robust self-interference cancellation for microstrip antennas by means of phase reconfigurable coupler," *IEEE Trans. Antennas Propag.*, vol. 66, no. 10, pp. 5574–5579, Oct. 2018.

[14] M. R. Rahman, S. Bojja Venkatakrishnan, T. Nakatani, and J. L. Volakis, "Wideband self interference cancellation (SIC) RF front end for simultaneous transmit and receive (STAR) radios," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (AP-S/URSI)*, Jul. 2022, pp. 948–949.

[15] H. Zhao, U. De Silva, S. Pulipati, S. B. Venkatakrishnan, S. Bhardwaj, J. L. Volakis, S. Mandal, and A. Madanayake, "A broadband multistage self-interference canceller for full-duplex MIMO radios," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 4, pp. 2253–2266, Apr. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9371299/

[16] M. N. A. Tarek, R. Hokayem, S. R. Govindarajulu, M. H. Novak, and E. A. Alwan, "A two-stage wideband RF cancellation of coupled transmit signal for bi-static simultaneous transmit and receive system," *IEEE J. Microw.*, vol. 2, no. 3, pp. 429–441, Jul. 2022.

[17] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," 2013, *arXiv:1302.2185*.

[18] U. De Silva, A. Madanayake, S. Pulipati, L. Belostotski, T. Kulatunga, H. Zhao, and S. Mandal, "A comparison of AI-enabled digital twins for DSP-based self-interference cancellation in wideband full-duplex communications," in *Proc. IEEE-APS Topical Conf. Antennas Propag. Wireless Commun. (APWC)*, Aug. 2021, p. 136.

[19] H. Luo, M. Holm, and T. Ratnarajah, "On the performance of active analog self-interference cancellation techniques for beyond 5G systems," *China Commun.*, vol. 18, no. 10, pp. 158–168, Oct. 2021.

[20] R. Vann Cleff Raro, G. Gems Mendoza, and M. Francis De Guzman, "Analog domain cancellation in hybrid self-interference cancellers for in-band full-duplex radios," in *Proc. IEEE Asia–Pacific Microw. Conf. (APMC)*, Dec. 2019, pp. 646–648.

[21] S. Khaledian, F. Farzami, B. Smida, and D. Erricolo, "Inherent self-interference cancellation at 900 MHz for in-band full-duplex applications," in *Proc. IEEE 19th Wireless Microw. Technol. Conf. (WAMICON)*, Apr. 2018, pp. 1–4.

[22] D. Korpi, Y.-S. Choi, T. Huusari, L. Anttila, S. Talwar, and M. Valkama, "Adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio: Algorithms and RF measurements," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–7.

[23] A. T. Kristensen, A. Burg, and A. Balatsoukas-Stimming, "Advanced machine learning techniques for self-interference cancellation in full-duplex radios," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, Nov. 2019, pp. 1149–1153.

[24] E. Ahmed and A. M. Eltawil, "All-digital self-interference cancellation technique for full-duplex systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3519–3532, Jul. 2015.

[25] J. Li, H. Zhang, and M. Fan, "Digital self-interference cancellation based on independent component analysis for co-time co-frequency full-duplex communication systems," *IEEE Access*, vol. 5, pp. 10222–10231, 2017.

[26] G. Chaudhary, J. Jeong, and Y. Jeong, "Differential fed antenna with high self-interference cancellation for in-band full-duplex communication system," *IEEE Access*, vol. 7, pp. 45340–45348, 2019.

[27] L. Sun, Y. Li, Z. Zhang, and H. Wang, "Antenna decoupling by common and differential modes cancellation," *IEEE Trans. Antennas Propag.*, vol. 69, no. 2, pp. 672–682, Feb. 2021.

[28] J. Ha, M. A. Elmansouri, P. Valale Prasannakumar, and D. S. Filipovic, "Monostatic co-polarized full-duplex antenna with Left- or right-hand circular polarization," *IEEE Trans. Antennas Propag.*, vol. 65, no. 10, pp. 5103–5111, Oct. 2017.

[29] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, X. Zhang, and S. Barghi, "The case for antenna cancellation for scalable full-duplex wireless communications," in *Proc. 10th ACM Workshop Hot Topics Netw.*, Nov. 2011, pp. 1–6.

[30] A. Geiler and V. Harris, "Atom magnetism: Ferrite circulators—Past, present, and future," *IEEE Microw. Mag.*, vol. 15, no. 6, pp. 66–72, Sep. 2014.

[31] N. Reiskarimian, J. Zhou, and H. Krishnaswamy, "A CMOS passive LPTV nonmagnetic circulator and its application in a full-duplex receiver," *IEEE J. Solid-State Circuits*, vol. 52, no. 5, pp. 1358–1372, May 2017.

[32] M. M. Biedka, R. Zhu, Q. M. Xu, and Y. E. Wang, "Ultra-wide band non-reciprocity through sequentially-switched delay lines," *Sci. Rep.*, vol. 7, no. 1, pp. 1–16, Jan. 2017.

[33] N. A. Estep, D. L. Sounas, and A. Alù, "Magnetless microwave circulators based on spatiotemporally modulated rings of coupled resonators," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 2, pp. 502–518, Feb. 2016.

[34] N. Reiskarimian and H. Krishnaswamy, "Magnetic-free non-reciprocity based on staggered commutation," *Nature Commun.*, vol. 7, no. 1, Apr. 2016.

[35] U. De Silva, H. Malavipathirana, S. Pulipati, S. Bhardwaj, S. Mandal, and A. Madanayake, "Implementation and testing of a switching circulator for twin-pair STAR radio architectures," in *Proc. Moratuwa Eng. Res. Conf. (MERCon)*, Jul. 2020, pp. 331–336.

[36] H. Alves, T. Riihonen, and H. A. Suraweera, *Full-Duplex Communications for Future Wireless Networks*. Cham, Switzerland: Springer, 2020.

[37] B. van Liempd, B. Hershberg, B. Debaillie, P. Wambacq, and J. Craninckx, "An electrical-balance duplexer for in-band full-duplex with <-85 dBm in-band distortion at $+10$ dBm TX-power," in *Proc. 41st Eur. Solid-State Circuits Conf. (ESSCIRC)*, Sep. 2015, pp. 176–179.

[38] B. van Liempd, B. Hershberg, S. Ariumi, K. Raczkowski, K.-F. Bink, U. Karthaus, E. Martens, P. Wambacq, and J. Craninckx, "A $+70$-dBm IIP3 electrical-balance duplexer for highly integrated tunable front-ends," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 12, pp. 4274–4286, Dec. 2016.

[39] U. D. Silva, S. Pulipati, S. B. Venkatakrishnan, S. Bhardwaj, and A. Madanayake, "A passive STAR microwave circuit for 1–3 GHz self-interference cancellation," in *Proc. IEEE 63rd Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2020, pp. 105–108.

[40] M. R. Rahman, S. Bojja-Venkatakrishnan, and J. L. Volakis, "Ultra-wideband RF self interference cancellation filter for STAR radios," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, Dec. 2021, pp. 523–524.

[41] M. N. A. Tarek, M. Novak, and E. A. Alwan, "RF coupling suppression circuit for simultaneous transmit and receive systems," in *Proc. IEEE Int. Symp. Antennas Propag. North Amer. Radio Sci. Meeting*, Jul. 2020, pp. 1833–1834.

[42] S. Li and R. D. Murch, "An investigation into baseband techniques for single-channel full-duplex wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4794–4806, Sep. 2014.

[43] B. King, J. Xia, and S. Boumaiza, "Digitally assisted RF-analog self interference cancellation for wideband full-duplex radios," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 3, pp. 336–340, Mar. 2018.

[44] C. Li, H. Zhao, F. Wu, and Y. Tang, "Digital self-interference cancellation with variable fractional delay FIR filter for full-duplex radios," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 1082–1085, May 2018.

[45] Y. He, H. Zhao, W. Guo, S. Shao, and Y. Tang, "A time-robust digital self-interference cancellation in full-duplex radios: Receiver design and performance analysis," *IEEE Access*, vol. 8, pp. 185021–185031, 2020.

[46] C. Liu, S. Wang, and H. Zhao, "A memory polynomial-based iterative orthogonal least square algorithm for reconstructing self-interferences in full-duplex communications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, 2023, doi: 10.1109/TCSII.2023.3342214.

[47] K. Komatsu, Y. Miyaji, and H. Uehara, "Frequency-domain Hammerstein self-interference canceller for in-band full-duplex OFDM systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.

[48] P. P. Campo, D. Korpi, L. Anttila, and M. Valkama, "Nonlinear digital cancellation in full-duplex devices using spline-based Hammerstein model," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–7.

[49] K. E. Kolodziej, A. U. Cookson, and B. T. Perry, "RF canceller tuning acceleration using neural network machine learning for in-band full-duplex systems," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1158–1170, 2021.

[50] H. Guo, J. Xu, S. Zhu, and S. Wu, "Realtime software defined self-interference cancellation based on machine learning for in-band full duplex wireless communications," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Mar. 2018, pp. 779–783.

[51] U. De Silva, T. Koike-Akino, R. Ma, A. Yamashita, and H. Nakamizo, "A modular 1D-CNN architecture for real-time digital pre-distortion," in *Proc. IEEE Topical Conf. RF/Microwave Power Model. for Radio Wireless Appl. (PAWR)*, Jan. 2022, pp. 79–81.

[52] Q. Zhou, S. Shen, Y.-W. Chen, R. Zhang, J. Finkelstein, and G.-K. Chang, "Simultaneous nonlinear self-interference cancellation and signal of interest recovery using dual input deep neural network in new radio access networks," *J. Lightw. Technol.*, vol. 39, no. 7, pp. 2046–2051, Apr. 15, 2021.

[53] A. Rathnayake, L. Silva, H. Rezaei, and N. Rajatheva, "Augmented-LSTM and 1D-CNN-LSTM models for linearization of wideband power amplifiers," in *Proc. IEEE 34th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2023, pp. 1–6.

[54] A. T. Kristensen, A. Burg, and A. Balatsoukas-Stimming, "Advanced machine learning techniques for self-interference cancellation in full-duplex radios," 2019, *arXiv:1912.06818*.

[55] M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Low complexity neural network structures for self-interference cancellation in full-duplex radio," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 181–185, Jan. 2021.

[56] M. Elsayed, A. A. Aziz El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Full-duplex self-interference cancellation using dual-neurons neural networks," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 557–561, Mar. 2022.

[57] F. J. Soriano-Irigaray, J. S. Fernandez-Prat, F. J. Lopez-Martinez, E. Martos-Naya, O. Cobos-Morales, and J. T. Entrambasaguas, "Adaptive self-interference cancellation for full duplex radio: Analytical model and experimental validation," *IEEE Access*, vol. 6, pp. 65018–65026, 2018.

[58] D. Korpi, "Full-duplex wireless: Self-interference modeling, digital cancellation, and system studies," Ph.D. dissertation, Fac. Comput. Elect. Eng., Tampere Univ. Technol., Tampere, Finland, 2017.

[59] S. Ge, J. Meng, J. Xing, Y. Liu, and C. Gou, "A digital-domain controlled nonlinear RF interference cancellation scheme for co-site wideband radios," *IEEE Trans. Electromagn. Compat.*, vol. 61, no. 5, pp. 1647–1654, Oct. 2019.

[60] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: Pushing the limits," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 80–87, Sep. 2016.

[61] K. Kolodziej, *In-Band Full-Duplex Wireless Systems Handbook*. Norwood, MA, USA: Artech House, 2021. [Online]. Available: https://books.google.com/books?id=oH0qEAAAQBAJ

[62] M. Valkama and M. Renfors, "Digital filter design for I/Q imbalance compensation," in *Proc. 10th Eur. Signal Process. Conf.*, Sep. 2000, pp. 1–4.

[63] A. T. Le, X. Huang, L. C. Tran, and Y. J. Guo, "On the impacts of I/Q imbalance in analog least mean square adaptive filter for self-interference cancellation in full-duplex radios," *IEEE Trans. Veh. Technol.*, vol. 71, no. 10, pp. 10683–10693, Oct. 2022.

[64] F. Ghannouchi, O. Hammi, and M. Helaoui, *Behavioral Modeling and Predistortion of Wideband Wireless Transmitters*. Hoboken, NJ, USA: Wiley, 2015. [Online]. Available: https://books.google.com/books?id=EcnlCAAAQBAJ

[65] A. S. Sappal, "Simplified memory polynomial modelling of power amplifier," in *Proc. Int. Conf. Workshop Comput. Commun. (IEMCON)*, Oct. 2015, pp. 1–7.

[66] M. Tanio, N. Ishii, and N. Kamiya, "A sparse neural network-based power adaptive DPD design and its hardware implementation," *IEEE Access*, vol. 10, pp. 114673–114682, 2022.

[67] M. Tanio, N. Ishii, and N. Kamiya, "Efficient digital predistortion using sparse neural network," *IEEE Access*, vol. 8, pp. 117841–117852, 2020.

[68] B. Csáji, "Approximation with artificial neural networks," M.Sc. thesis, Fac. Math. Comput. Sci., Eötvös Loránd Univ., Budapest, Hungary, 2001.

[69] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang, "Phoneme recognition using time-delay neural networks," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 3, pp. 328–339, Mar. 1989.

[70] A. Yadav, D. Mishra, S. Ray, R. N. Yadav, and P. K. Kalra, "Representation of complex-valued neural networks: A real-valued approach," in *Proc. Int. Conf. Intell. Sens. Inf. Process.*, 2005, pp. 331–335.

[71] A. Balatsoukas-Stimming, "Non-linear digital self-interference cancellation for in-band full-duplex radios using neural networks," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.

[72] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New York, NY, USA: McGraw-Hill, 2016. [Online]. Available: https://books.google.com/books?id=hFzmCwAAQBAJ

[73] S. Sonoda and N. Murata, "Neural network with unbounded activation functions is universal approximator," *Appl. Comput. Harmon. Anal.*, vol. 43, no. 2, pp. 233–268, Sep. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1063520315001748

[74] *Power Amplifier Characterization*. Accessed: Sep. 27, 2022. [Online]. Available: https://www.mathworks.com/help/simrf/ug/power-amplifier-characterization.html

[75] *Circuit Envelope Simulation*. Accessed: Oct. 10, 2022. [Online]. Available: https://www.mathworks.com/help/simrf/circuit-envelope-simulation.html

[76] S. Pulipati, V. Ariyarathna, U. D. Silva, N. Akram, E. Alwan, A. Madanayake, S. Mandal, and T. S. Rappaport, "A direct-conversion digital beamforming array receiver with 800 MHz channel bandwidth at 28 GHz using Xilinx RF SoC," in *Proc. IEEE Int. Conf. Microw., Antennas, Commun. Electron. Syst. (COMCAS)*, Nov. 2019, pp. 1–5.

[77] K. Ullah, S. B. Venkatakrishnan, and J. L. Volakis, "28 GHz millimeter-wave digital beamformer: Design and experimental evaluation," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, Dec. 2021, pp. 1015–1016.

[78] K. Ullah, S. B. Venkatakrishnan, and J. L. Volakis, "RFSoC-based digital beamformer for millimeter-wave MIMO applications," in *Proc. United States Nat. Committee URSI Nat. Radio Sci. Meeting*, Jan. 2022, pp. 147–148.

[79] D. Kim, H. Kang, Y. Son, S. Kim, and S. Nam, "Commercial RFSoC-based wideband MIMO-FMCW radar design with effective pre-distortion," in *Proc. IEEE Radar Conf.*, Mar. 2022, pp. 1–4.

[80] K. Ullah, S. B. Venkatakrishnan, and J. L. Volakis, "Millimeter-wave digital beamforming receiver using RFSoC FPGA for MIMO communications," in *Proc. IEEE 22nd Annu. Wireless Microw. Technol. Conf. (WAMICON)*, Apr. 2022, pp. 1–4.

[81] M. A. Hossain, "Performance analysis of barker code based on their correlation property in multiuser environment," *Int. J. Inf. Sci. Techn.*, vol. 2, no. 1, pp. 27–39, Jan. 2012.

[82] *HDL Coder Documentation*. Accessed: Jan. 30, 2024. [Online]. Available: https://www.mathworks.com/help/hdlcoder/

[83] Y. Kurzo, A. Burg, and A. Balatsoukas-Stimming, "Design and implementation of a neural network aided self-interference cancellation scheme for full-duplex radios," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 589–593.

[84] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2011, pp. 301–312.

[85] D. Korpi, M. AghababaeeTafreshi, M. Piilila, L. Anttila, and M. Valkama, "Advanced architectures for self-interference cancellation in full-duplex radios: Algorithms and measurements," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 1553–1557.

[86] L. Anttila, V. Lampu, S. A. Hassani, P. P. Campo, D. Korpi, M. Turunen, S. Pollin, and M. Valkama, "Full-duplexing with SDR devices: Algorithms, FPGA implementation, and real-time results," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2205–2220, Apr. 2021.

• • •