A Real-Time Wideband Subband LMS Algorithm for Full-Duplex Communications

KEVIN S. LORENZ[ID], **JOEL GOODMAN**[ID]**, (Senior Member, IEEE), MICHELLE MCBETH, MATTHEW B. MCKEON, AND DAVID J. PARRETT**
U.S. Naval Research Laboratory, Washington, DC 20375, USA

Corresponding author: Kevin S. Lorenz (kevin.lorenz@nrl.navy.mil)

ABSTRACT In-band full-duplex (IBFD) communications has the potential to nearly double the spectrum efficiency of existing 5G communications. Digital signal processing to mitigate self-interference is one way to realize this full-duplex capability. However, the aggregated wideband nature of 5G coupled with the potential for large delay spreads and short coherence times makes the realistic implementation of such an IBFD processor challenging. In this paper, we address the architectural considerations and describe the practical implementation of a real-time IBFD digital processor and its performance when operating on wideband data collected in a highly dynamic environment.

INDEX TERMS Adaptive filters, least mean square methods, radio transceivers, radio communication.

I. INTRODUCTION

In-band full-duplex communications can significantly improve the spectral efficiency of 5G cellular communications, with the potential to double both the uplink and downlink throughput [1]. Although there are signal processing approaches to realize digital cancellation for full-duplex communications [2]–[4], most papers in the open literature that describe practical full-duplex implementations have focused on RF and analog cancellation circuits [5]. For digital implementations, 5G presents some unique challenges—most notably, wireless data transfers must occur in as much as 400 MHz of instantaneous bandwidth [6]. With the potential for root mean square (RMS) delay spreads of 100–300 ns [7] and sample rates in excess of 1 Gsps, a finite impulse response (FIR) filter with several hundreds of coefficients may be needed to suppress self-interference. These difficulties, along with a relatively short coherence time [8], make the task of implementing a real-time canceler in digital hardware that is capable of tracking a rapidly changing channel a significant challenge.

In this paper, we consider both the problem and candidate solution to the digital baseband implementation of a full-duplex communications system that is hosted by a Xilinx field-programmable gate array (FPGA). We describe the design of a subband adaptive filter (SAF) least mean

squares (LMS) architecture operating in 1 GHz of instantaneous bandwidth. The SAF LMS architecture described in this paper is preferable to that of a frequency-domain block LMS algorithm [9], as it is better suited to following rapid changes in the channel impulse response and is not susceptible to extreme dynamic range challenges arising from quantization error. To the best of our knowledge, there are no other papers in the open literature that describe the implementation of a real-time LMS algorithm operating in excess of 2 Gsps (1 GHz of instantaneous bandwidth) for not only full-duplex operation [10], [11] but for any application.

This paper is organized as follows: Section II describes the technical approach to achieve such high data rates, Section III describes experimental results demonstrating the approach's efficacy, and Section IV concludes with the novelty of this approach.

II. TECHNICAL APPROACH

Baseband full-duplex operation was realized using a "two-channel" architecture [12] presented in [13] and depicted in Fig. 1. In this architecture, a reference channel (REF) samples the output of the transmitter, y_{Tx}, after its convolution with the REF impulse response, h_{ref}, of the receiver which includes the response of the analog-to-digital converter (ADC). A second channel, the over-the-air (OA) channel, receives the external signal of interest (SOI), y_{soi}, and a backscattered transmitted signal after its convolution with the OA impulse response, h_{oa}. Note that h_{oa} represents both
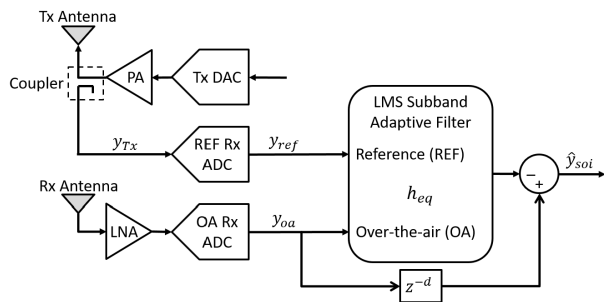
FIGURE 1. Two-channel RF architecture that uses SAF LMS for full-duplex operation.

the channel backscatter and the impulse response of the RF receiver. The SOI experiences a different multipath environment than the transmitted signal and thus is only subject to the impulse response of the RF receiver and not h_{oa}. As such, the signals associated with the REF and OA channels are, respectively,

$$y_{ref} = y_{Tx} * h_{ref}$$
$$y_{oa} = y_{Tx} * h_{oa} + y_{soi} \qquad (1)$$

where the $*$ operator denotes convolution. Within the context of the canonical LMS system model, y_{ref} is the system input $x(n)$ and y_{oa} is the desired signal $d(n)$.

Hardware impairments arising from, for example, the high-power amplifier (HPA) in the transmitter can significantly decrease full-duplex cancellation performance. However, by employing the two-channel architecture depicted in Fig. 1, the deleterious effects of transmitter noise and distortion are mitigated by sampling the transmitter output in the process of forming a linear equalization filter [14]. This enables cancellation performance which would otherwise be unachievable using a single-channel receiver with linear equalization alone [15]. Additionally, sampling the transmitter output with a two-channel architecture enables the system to cancel self-interference irrespective of the modulation used by—or other signal statistics of—either the transmitter or the external SOI. A SAF LMS approach is then used to identify the equalizer, h_{eq}, whose ideal impulse response is $h_{ref}^{-1} * h_{oa}$. Lastly, the SOI is recovered by subtracting the equalized REF channel from the OA channel:

$$\widehat{y_{soi}} = y_{oa} - y_{ref} * h_{eq} \qquad (2)$$

The LMS algorithm is among the most popular adaptive filtering algorithms because of its algorithmic simplicity. However, LMS suffers from various issues, including convergence speed and stability in the presence of in-band interference [16]. With the canonical LMS formulation, the error signal $e(n)$ is expected to converge to zero. However, in the presence of other in-band signals such as with orthogonal frequency-division multiplexing (OFDM) as well as self-interference in the case of full-duplex communications, the error signal should decrease but not converge to zero. This is because, after the adaptive filter has converged and reached

steady-state, the error signal is expected to be identical to the signal desired to be received (i.e., the SOI). More specifically, the estimate of the SOI $\widehat{y_{soi}}$ in Eq. (2) is identical to the error signal $e(n)$.

Here we propose the integration of two separate enhancements to the traditional LMS algorithm: subbanding and parallelization. Additional unique and novel enhancements and adaptations are performed to realize real-time operation at 2 Gsps. Therefore, both digital signal processing methodology and approaches for practical hardware implementation must simultaneously be considered.

A. SUBBAND LMS
Consider the canonical system model [17]

$$d(n) = \mathbf{w}^T \mathbf{x}(n) + \eta(n) \qquad (3)$$

where n denotes the sample index, T denotes the vector transpose operation,

$$\mathbf{w} = \begin{bmatrix} w_0 & w_1 & \cdots & w_{P-1} \end{bmatrix}^T$$

denotes the vector of P unknown system coefficients,

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-P+1) \end{bmatrix}^T$$

denotes the vector of the most recent P system inputs, and $\eta(n)$ denotes any source of noise or interference that is independent of the system input. With the *a priori* error signal defined as

$$e(n) = d(n) - \hat{\mathbf{w}}^T(n-1)\mathbf{x}(n) \qquad (4)$$

where $\hat{\mathbf{w}}(n)$ denotes the vector of M adaptive filter coefficients, the canonical normalized adaptive coefficient update is [17]

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}^T(n-1) + \mu \frac{\mathbf{x}(n)}{\mathbf{x}^T(n)\mathbf{x}(n)} e(n) \qquad (5)$$

where μ denotes the step-size.

The implication of Eq. (5) is that the adaptive filter coefficients $\hat{\mathbf{w}}(n)$ are updated with every new sample of the received signal $x(n)$. However, this canonical approach cannot be realized in real-time at 2 Gsps because current hardware is not able to achieve such high clock rates. Therefore, subband LMS architectures are considered for a practical implementation because they enable the adaptive filter coefficients to be updated at a much slower rate than the data rate while still being able to quickly adapt to a rapidly changing channel.

Subband LMS approaches have been widely explored [18]–[20], most often to reduce computational complexity and improve convergence speed. However, these methods often require specially designed filter banks. Other methods have mitigated the additional delay introduced by subband filtering [21]. However, such an approach uses block Fourier transforms which introduce fixed-point challenges with extraordinarily high dynamic range requirements when implementing on an FPGA. Due to these challenges, an approach from [22] is adopted.
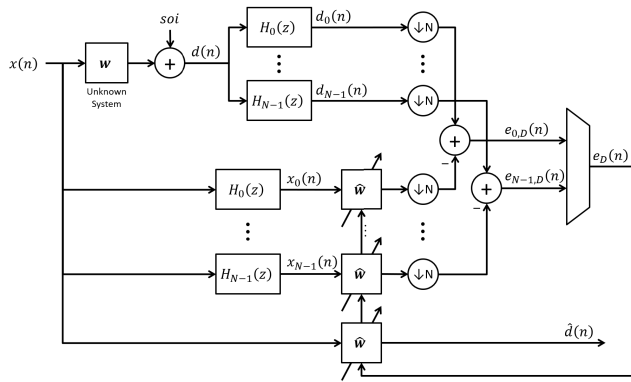
FIGURE 2. Block diagram of subband LMS algorithm using naive architecture implementation. Note that the adaptive filter $\hat{\mathbf{w}}$ is identically instantiated $N + 1$ times, resulting in excessive FPGA area consumption.

FIGURE 3. Block diagram of subband LMS algorithm using proposed implementation optimization. Note that the adaptive filter $\hat{\mathbf{w}}$ is now only instantiated once, significantly reducing FPGA area consumption while achieving equivalent functionality.

Using the adopted subband approach [22], the coefficient update becomes

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}^T(n-1) + \mu \sum_{k=0}^{N-1} \frac{\mathbf{x}_k(n)}{||\mathbf{x}_k(n)||^2} e_{k,D}(n) \quad (6)$$

where N denotes the number of subbands,

$$\mathbf{x}_k(n) = \begin{bmatrix} x_k(nN) & x_k(nN-1) & \cdots & x_k(nN-M+1) \end{bmatrix}^T \quad (7)$$

is the system input vector for the kth subband, $e_{k,D}(n)$ denotes the decimated error signal for the kth subband, M denotes the number of coefficients of the adaptive filter $\hat{\mathbf{w}}(n)$, and μ denotes the step-size. This adopted SAF approach also allows the flexibility to employ a user-selected filter bank architecture. Therefore, a generic cosine modulated filter bank with prototype filter based on [23] is selected. Then, the filter bank is constructed by [24]

$$h_k(n) = 2h_p(n) \cos\left((2k+1)\frac{\pi}{2N}\left(n - \frac{L-1}{2}\right) + (-1)^k \frac{\pi}{4}\right) \quad (8)$$

where $k \in [0, N-1]$ denotes the subband index, $h_p(n)$ is the prototype filter, and L denotes the number of coefficients in $h_p(n)$. Note that $h_k(n)$ only represents the analysis filter bank. The synthesis filter bank is unnecessary for the real-time operation of the SAF LMS architecture and is omitted to simplify and reduce computational complexity.

A naive implementation of the adopted SAF architecture is shown in the block diagram in Fig. 2, and while the methodology is correct, this architecture cannot operate at a sample rate of 2 Gsps due to excessive FPGA resource consumption. This is primarily because the adaptive filter $\hat{\mathbf{w}}(n)$ is replicated across all N subbands, which suggests that many optimizations may be realized. Therefore, options to reduce FPGA area consumption for a practical hardware implementation must be considered that are not generally considered for a non-real-time software implementation. Equivalent functionality may be achieved with instantiating the adaptive filter only once by transposing $\hat{\mathbf{w}}(n)$ with the subband filters $H_k(z)$
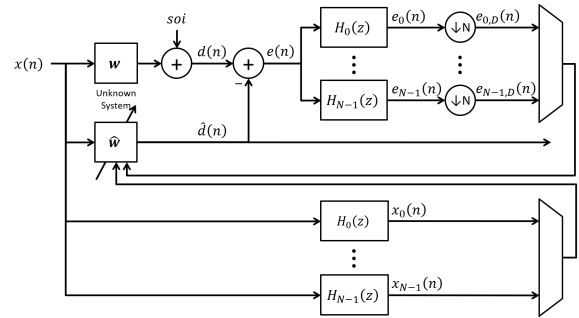
to the beginning of the signal processing chain. Consequently, the loop architecture is rearranged such that the error signal $e(n)$ undergoes subband filtering rather than the desired signal $d(n)$. As a result, this will eliminate the computation of an error signal for each subband and instead, a single error signal will be computed as in the canonical LMS architecture. As a result of these optimizations, a block diagram of this optimized subband architecture is shown in Fig. 3. The development of this architecture targeted for an FPGA implementation operating at 2 Gsps is unique and novel and is solely motivated by the real-time requirements and constraints imposed by such extremely high data rates.

B. PARALLELIZED LMS
Current FPGA clock rates are generally limited to approximately 500 MHz in practical operation despite being advertised to operate at faster clock speeds. To achieve real-time data rates in excess of 500 Msps, multiple data samples per clock cycle must be processed. This requires the use of parallel FIR filters which must be integrated into the architecture to enable processing of more than one sample per clock cycle. With R denoting the number of samples processed simultaneously per clock cycle, each of the subband analysis filters $H_k(z)$ and the adaptive filter $\hat{\mathbf{w}}(n)$ in Fig. 3 are replaced with its equivalent parallel FIR architecture, an example of which is shown in Fig. 4 for $R = 4$. In Fig. 4, each sub-filter $h_{s,r}$ for $r \in [0, R-1]$ includes every Rth coefficient and is defined as

$$h_{s,r}(n) = \begin{bmatrix} h_s(r) & h_s(r+R) & h_s(r+2R) & \cdots \end{bmatrix}^T \quad (9)$$

where $h_s(n)$ represents the serial FIR filter coefficients to be implemented in parallel.

C. HARDWARE IMPLEMENTATION
A real-time hardware implementation poses a number of challenges—most notably, the very high data rate relative to the FPGA clock rate. Specifically, the data rate of 2 Gsps greatly exceeds the theoretical maximum FPGA clock rate—in our case, 775 MHz [25]. Even if the data rate requirement was not as demanding, generally it is very difficult to achieve
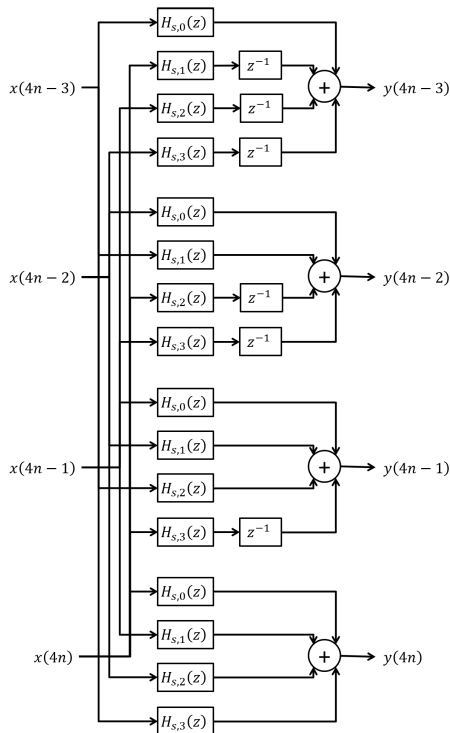
FIGURE 4. Parallel FIR filter architecture with four samples processed simultaneously per clock cycle.

TABLE 1. FPGA Resource Usage of the Xilinx Zync UltraScale+ XCZU28DR.

Resource	Usage	Percentage
LUT	78327	18%
LUTRAM	23033	11%
FF	146675	17%
BRAM	26	2%
DSP	2984	70%

requirements while still ensuring L and M are large enough to meet system performance requirements must all be simultaneously achieved for a successful hardware implementation.

This balance was achieved with a clock rate of 500 MHz and $R = 4$ samples per clock cycle, resulting in a data rate of 2 Gsps. Additionally, if the number of subbands N is chosen such that $N = R$, then the downsampling step as part of the subbanding process simply discards $R - 1$ of the R samples being processed simultaneously. As a result, the entire FPGA design will run at a single clock rate, vastly simplifying the design effort by completely avoiding the implementation of multiple clock domains and clock domain crossing logic.

The aforementioned issues are only some of the factors to be considered where theoretical performance must be sacrificed in order to realize a practical implementation. Another factor that must be considered is the selection of a fixed-point implementation. The dynamic range of every data path within the FPGA must be considered to appropriately choose the number of bits and decimal point placement for each path, which will necessarily introduce quantization error and therefore degrade performance.

To help overcome the numerous aforementioned implementation challenges, the proposed algorithms were implemented using MathWorks HDL Coder to target the Xilinx Zync UltraScale+ XCZU28DR. This commercial off-the-shelf software tool enables rapid prototyping of digital signal processing algorithms in a high-level language and is able to automatically generate the equivalent low-level hardware description language (HDL) code, significantly reducing development time. The resulting HDL code is then used by Xilinx tools to create an FPGA bitstream file. As a result, the entire design may be implemented in real-time hardware without manually writing any HDL code. After successful implementation, Table 1 lists the FPGA resources consumed. Despite the high data rate of 2 Gsps combined with the computational requirements of subbanding, the design consumes less than 70% of the resources available on a relatively small FPGA. To the best of our knowledge, there are no other real-time digital design implementations that realize any variant of an LMS adaptive filter at 2 Gsps to compare this table against.

III. EXPERIMENTAL RESULTS

To test the efficacy of the full-duplex communications architecture developed with consideration of a real-time FPGA implementation in a representative 5G environment, results were gathered in both simulation and a laboratory demonstration. In both cases, data was obtained using an over-the-air
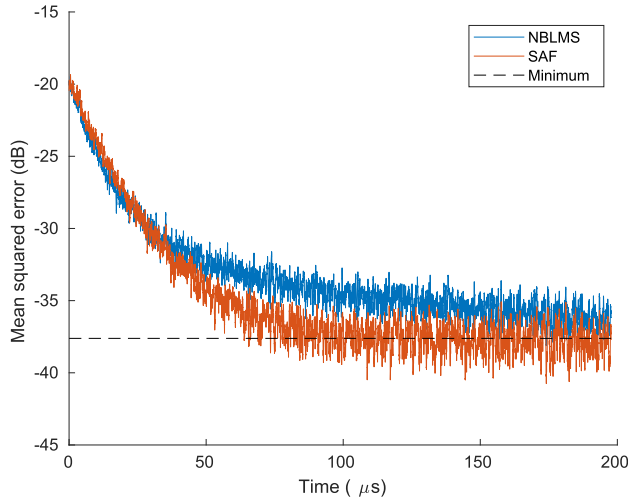
the theoretical maximum FPGA clock rate when the FPGA is highly utilized due to placement constraints, routing congestion, fan-out limitations, as well as other factors. However, the algorithm benefits from parallelization and downsampling via the subband process to meet the high throughput requirements. Specifically, by instantiating the adaptive filter only once as the shown in Fig. 3, this optimized subband architecture will require MR fewer multipliers compared with the naive approach, significantly saving FPGA area.

There are many feasible combinations of R and clock rate to achieve the desired data rate. For example, selecting $R = 5$ with a clock rate of 400 MHz or selecting $R = 8$ with a clock rate of 250 MHz will both achieve a data rate of 2 Gsps. However, specific compromises and tradeoffs only relevant to a hardware implementation must be considered. For example, meeting FPGA timing constraints becomes more challenging as the clock rate increases. Reducing the clock rate to more easily meet timing constraints while still maintaining the same overall data rate may be accomplished by increasing R to process more data samples per clock cycle. However, increasing R will necessarily increase the FPGA area consumed. As FPGA area consumption increases, meeting timing constraints also becomes more challenging. We also recognize that a hardware implementation imposes delay in the signal processing loop, which has a deleterious effect. However, this may mitigated by minimizing L and M as much as practically possible. Therefore, a balance among reducing both clock rate and R to meet timing constraints and also reducing L and M to meet convergence and stability

FIGURE 5. Comparison of MSE when using normalized block LMS (NBLMS) and the proposed subband adaptive filter (SAF). The optimal minimum MSE is shown as a dashed line.

test in a representative 5G environment. The test setup consisted of an S-band radio operating in 1 GHz of instantaneous bandwidth placed in an anechoic chamber along with our real-time full-duplex apparatus. A reflector mounted on a rotating pedestal was placed 3 m away from the radio. The multipath delay spread due to this reflector, as well as the chamber's walls, was approximately 0.12 μs and thus required a cancellation filter with $M = 256$ coefficients at 2 Gsps.

Choosing M such that $M \geq P$ is a necessary—but not sufficient—condition for the adaptive filter to converge correctly. Additionally, poor performance was empirically observed unless the condition $L \geq 8N$ was satisfied. As a result, the following adaptive filter architecture parameters have been selected: $N = 4$, $L = 32$, $M = 256$, $R = 4$, and $\mu = 0.01$.

A. METHODOLOGY COMPARISON

Simulations of results from the proposed SAF LMS architecture were conducted using the measured over-the-air data. The resulting mean squared error (MSE) is shown in Fig. 5. This result is compared against the MSE from using the normalized block LMS (NBLMS) algorithm with the block size equal to N. The NBLMS algorithm is defined as [26]

$$n = kN + i$$
$$e(n) = d(n) - \hat{\mathbf{w}}^T(k-1)\mathbf{x}(n)$$
$$\hat{\mathbf{w}}(k) = \hat{\mathbf{w}}(k-1) + \mu \sum_{i=0}^{N-1} \frac{\mathbf{x}(kN+i)}{\mathbf{x}^T(kN+i)\mathbf{x}(kN+i)} e(kN+i)$$

$$(10)$$

where N denotes the block size, k the block index, n the sample index, and i the sample index within a block. We compared our technique with the NBLMS algorithm because it updates the adaptive filter coefficients $\hat{\mathbf{w}}$ once every N

TABLE 2. Performance Metrics.

Algorithm	Convergence Time (μs)	Steady-State MSE (dB)
Wiener filter	—	-37.6
NBLMS	33.2	-37.0
SAF	28.5	-37.5

samples, equivalent to the proposed SAF technique. In contrast, canonical LMS in Eq. 5 updates $\hat{\mathbf{w}}$ on every sample, which cannot be realized in real-time for 2 Gsps and is therefore an inappropriate comparison. By only comparing the proposed approach with another LMS approach that also could feasibly be implemented in real-time at 2 Gsps, we ensured a fair comparison by simultaneously considering the performance of the methodology and practical implementation limitations.

Also shown in Fig. 5 is the minimum achievable MSE. The minimum MSE, J_{min}, is achieved when using the optimal Wiener filter, $\mathbf{w}_{opt}$, and this are computed as [26]

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{P}$$
$$J_{min} = \mathbb{E}\left[d^2(n)\right] + \mathbf{w}_{opt}^T\mathbf{R}\mathbf{w}_{opt} - 2\mathbf{w}_{opt}^T\mathbf{P} \qquad (11)$$

where $\mathbf{R}$ denotes the autocorrelation matrix of $\mathbf{x}$, $\mathbf{P}$ denotes the cross-correlation vector between $\mathbf{x}$ and $\mathbf{d}$ of length M, and $\mathbb{E}[\cdot]$ denotes the expectation operator.

Two metrics that are often used to evaluate LMS algorithm performance are steady-state MSE and the time to converge to steady-state. While Fig. 5 shows that both of the two approaches converge in steady-state nearly to J_{min}, the time required to do so is drastically different. Quantitative results including time to converge within 95% of J_{min} and the final steady-state MSE are shown in Table 2. As shown, the proposed SAF approach isolated an external signal of interest within 0.1 dB of the optimal steady-state MSE and significantly improved convergence time compared to the NBLMS approach.

B. LABORATORY DEMONSTRATION

In addition to the simulation results, a real-time hardware demonstration was conducted. The system consisted of a Xilinx Zynq UltraScale + RFSoC ZCU111 Evaluation Kit[1] which is pictured in Fig. 6 and includes the Zynq UltraScale + XCZU28DR FPGA discussed in Section II-C. The Zynq FPGA hosted the SAF self-interference cancellation algorithms that were optimized for a real-time hardware implementation and described in Section II. The RFSoC includes two 12-bit analog-to-digital converters (ADCs) and were both clocked to achieve an effective 2 Gsps sampling rate. The samples from each of the two ADCs were demultiplexed into four parallel data streams—as described in Section II-C where $R = 4$—and each data stream was simultaneously presented to the FPGA at 500 MHz. With a FPGA clock
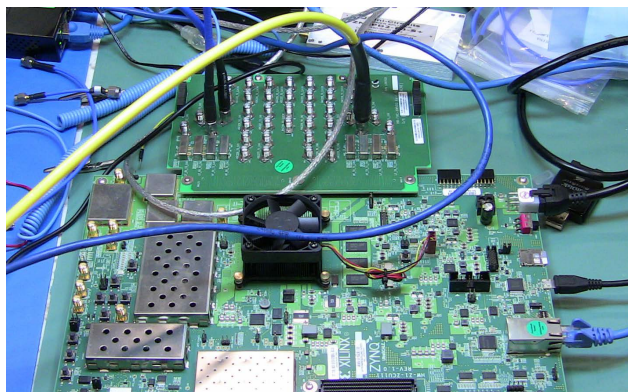
[1]http://www.xilinx.com/zcu111

FIGURE 6. Xilinx Zynq UltraScale+ RFSoC ZCU111 Evaluation Kit which includes the Zynq UltraScale+ FPGA used to implement the proposed SAF LMS algorithm for real-time digital cancellation.
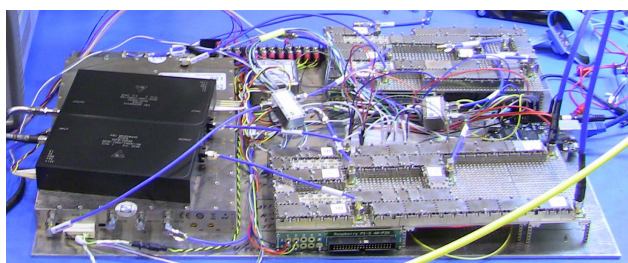

FIGURE 7. Custom two-channel RF front-end transceiver.

TABLE 3. RF Receiver Specifications.

Receiver Parameter	Range and Value
Tunable frequency	1–10 GHz
Cascaded IP3	32 dBm
Noise figure	10 dB
SFDR	56 dBFS
Gain	25 dB

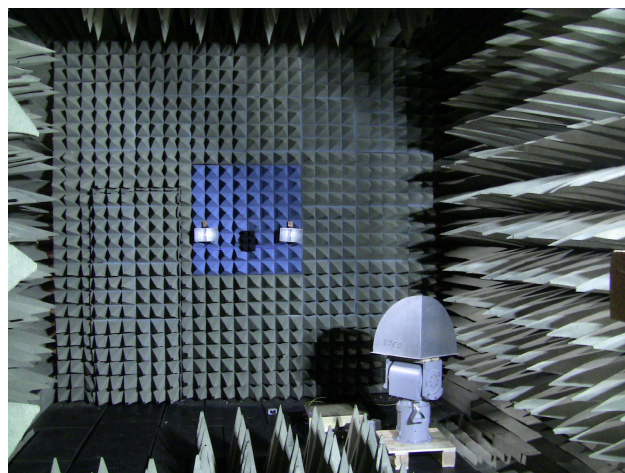**FIGURE 8.** Anechoic chamber with a corner reflector mounted on a rotating pedestal emulating a dynamic multipath environment.

rate of 500 MHz and four samples processed per clock cycle, a data rate of 2 Gsps is realized.

A custom two-channel receiver was developed as an RF front-end and placed before the ADCs of the RFSoC, and is pictured in Fig. 7. The receiver's specifications are given in Table 3. As described in [14], one of the two channels of the receiver was coupled to the output port of the transmitter and formed the reference (REF) channel as illustrated in Fig. 1, while the other channel of the receiver was directly connected to the receive antenna and formed the over-the-air (OA) channel. The RF receiver used two stages of downconversion on each channel to generate signals in the second Nyquist zone in the intermediate frequency (IF) bandwidth from 1 GHz to 2 GHz that fed both of the ADCs on the RFSoC. Approximately −20 dBm of transmit power was coupled into the OA channel of the receiver from both direct-path leakage and backscatter.

With a 10 dB noise figure, the receiver's integrated noise floor across 1 GHz of instantaneous bandwidth was approximately −74 dBm. Given that the receiver was noise-limited rather than distortion-limited (cf. Table 3), this resulted in the receiver having a dynamic range of approximately −20 dBm + 74 dBm, or 54 dB. Because the desired signal $d(n)$ was derived from the reference (REF) channel and the input signal $x(n)$ was derived from the over-the-air (OA) channel, each of which included receiver noise of approximately equal variance, the upper bound on cancellation performance was 3 dB lower than the dynamic range of the

receiver, or 51 dB, which is directly related to the Cramér-Rao lower bound (CRLB) on adaptive filter coefficient ($\hat{w}$) estimation accuracy [27].

To emulate a 5G environment, we modeled a full-duplex base station with an over-the-air downlink-uplink connection and a mobile transceiver in the 5G n78 band (3.3 GHz to 3.8 GHz S-band) [6]. We placed our base station antenna into the 10 m × 5 m anechoic chamber pictured in Fig. 8. At the far end of the chamber, we placed the mobile communication transceiver that was connected to one of the horn antennas (as pictured), as well as a corner reflector mounted on a rotating pedestal. The pedestal spun at a rate of approximately 30 °/s and emulated the dynamic backscatter present in a multipath environment. At the near end of the chamber, we placed the base station antenna which was connected to the high-power amplifier (HPA) pictured in Fig. 9. This signal was coupled into the RF transceiver which in turn was connected to the ADCs of the RFSoC and digital baseband processor.

We emulated full-duplex base station transmission across all channels of the 5G n78 band with a single active mobile transmitter present. The output of the FPGA digital canceler was connected to one of the digital-to-analog (DAC) channels on the RFSoC and in turn was connected to one port of a spectrum analyzer after being converted to IF. Another DAC channel on the RFSoC served as a pass-through port prior to FPGA digital cancellation and was connected to a second port of the spectrum analyzer after IF conversion. The measured signal at the output of the RF receiver with and without digital cancellation as captured by the spectrum analyzer is

FIGURE 9. Waveform generator and transmitter connected to a high-power amplifier with cables leading to the antenna located in the adjacent anechoic chamber.
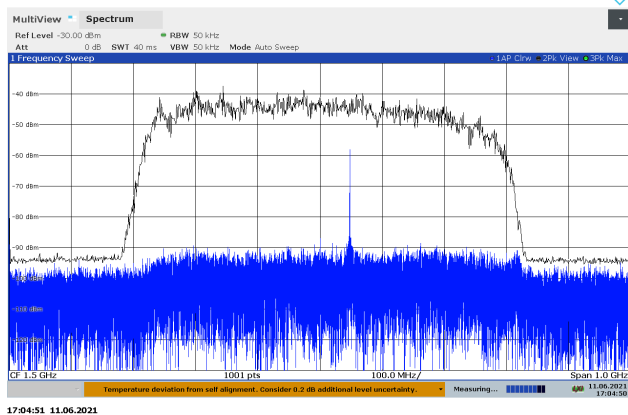

FIGURE 10. Frequency spectrum when self-interference cancellation is enabled. The wideband transmitted signal is indicated in black, and the received signal is indicated in blue. The self-interference is digitally suppressed by 50.5 dB across the entire 5G n78 band. The narrowband emitter (signal shown in blue) is received simultaneously during transmission of the signal shown in black.

pictured in Fig. 10. Without self-interference cancellation, all of the channels in the 5G n78 band are completely filled with self-interference due to the base station transmission coupling back into the RF front-end receiver with both direct path leakage and near-in backscatter. However, after digital cancellation, the mobile radio's transmission is recovered in the absence of error, with digital cancellation suppressing self-interference by five orders of magnitude—in this case, 50.5 dB—to come within 0.5 dB of the CRLB.

IV. CONCLUSION

Wideband operation, a rapidly changing RF channel, and a relatively long delay spread all represent significant challenges for the real-time implementation of a baseband full-duplex communications system. In this paper, we presented a novel joint signal processing and FPGA architecture that is able to achieve near-optimal cancellation performance when operating in 1 GHz of instantaneous bandwidth in a representative 5G environment. Most other LMS approaches in the open literature typically consider either theoretical performance or real-time hardware implementation constraints separately and individually. Instead, by jointly considering both in an interdisciplinary fashion, we developed and realized an LMS adaptive filter with an optimized architecture integrated with subbanding and parallelization that processes data at 2 Gsps in real-time.

REFERENCES

[1] X. Xia, K. Xu, Y. Wang, and Y. Xu, "A 5G-enabling technology: Benefits, feasibility, and limitations of in-band full-duplex mMIMO," *IEEE Veh. Technol. Mag.*, vol. 13, no. 3, pp. 81–90, Sep. 2018.

[2] D. Korpi, L. Anttila, V. Syrjala, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.

[3] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[4] 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. (MobiCom)*, 2011, pp. 301–312.

[5] 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.

[6] ETSI 3rd Generation Partnership Project (3GPP), *Base Station (BS) Radio Transmission and Reception*, European Telecommunications Standards Institute, document 38.104, version 15.9.0 Release 15, Apr. 2020.

[7] S. Henry, A. Alsohaily, and E. S. Sousa, "5G is real: Evaluating the compliance of the 3GPP 5G new radio system with the ITU IMT-2020 requirements," *IEEE Access*, vol. 8, pp. 42828–42840, 2020.

[8] K. Guan, B. Ai, B. Peng, D. He, G. Li, J. Yang, Z. Zhong, and T. Kürner, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part II: Case study for paradigm implementation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9129–9144, Oct. 2018.

[9] K. Sann Chan and B. Farhang-Boroujeny, "Analysis of the frequency-domain block LMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, May 2000, pp. 380–383.

[10] 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.

[11] H. Li, J. Van Kerrebrouck, O. Caytan, H. Rogier, J. Bauwelinck, P. Demeester, and G. Torfs, "Self-interference cancellation enabling high-throughput short-reach wireless full-duplex communication," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6475–6486, Oct. 2018.

[12] J. P. Doane, K. E. Kolodziej, and B. T. Perry, "Simultaneous transmit and receive with digital phased arrays," in *Proc. IEEE Int. Symp. Phased Array Syst. Technol. (PAST)*, Oct. 2016, pp. 1–6.

[13] L. Boglione, C. Davis, J. Goodman, M. McKeon, D. Parrett, S. Shin, and N. Walker, "STAR performance for SPEAR (signal processing electronic attack RFIC)," in *Proc. Government Microcircuit Appl. Crit. Technol. Conf. (GOMACTech)*, Reno, NV, USA, Mar. 2017. [Online]. Available: https://apps.dtic.mil/sti/pdfs/AD1042246.pdf

[14] J. Goodman, L. Boglione, and C. G. Davis, "Reference-based nonlinear digital cancellation," in *In-Band Full-Duplex Wireless Systems Handbook*, K. E. Kolodziej, Ed. Norwood, MA, USA: Artech House, 2021, ch. 12, pp. 199–220.

[15] 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.

[16] Z. Li, Y. Xia, W. Pei, K. Wang, and D. P. Mandic, "An augmented nonlinear LMS for digital self-interference cancellation in full-duplex direct-conversion transceivers," *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4065–4078, Jun. 2018.

[17] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, 2nd ed. Hoboken, NJ, USA: Wiley, Jun. 2013.

[18] M. de Courville and P. Duhamel, "Adaptive filtering in subbands using a weighted criterion," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2359–2371, Sep. 1998.

[19] S. Sandeep Pradham and V. U. Reddy, "A new approach to subband adaptive filtering," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 655–664, Mar. 1999.

[20] S. Gyun Kim, C. D. Yoo, and T. Q. Nguyen, "Alias-free subband adaptive filtering with critical sampling," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1894–1904, May 2008.

[21] R. Merched, P. S. R. Diniz, and M. R. Petraglia, "A new delayless subband adaptive filter structure," *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1580–1591, Jun. 1999.

[22] K. A. Lee and W. S. Gan, "Improving convergence of the NLMS algorithm using constrained subband updates," *IEEE Signal Process. Lett.*, vol. 11, no. 9, pp. 736–739, Sep. 2004.

[23] C. D. Creusere and S. K. Mitra, "A simple method for designing high-quality prototype filters for M-band pseudo QMF banks," *IEEE Trans. Signal Process.*, vol. 43, no. 4, pp. 1005–1007, Apr. 1995.

[24] T. Q. Nguyen, "A class of generalized cosine-modulated filter bank," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 2, May 1992, pp. 943–946.

[25] *Zynq UltraScale+ RFSoC Data Sheet: DC AC Switching Characteristics*, Xilinx, San Jose, CA, USA, Jan. 2022.

[26] S. S. Haykin, *Adapt. Filter Theory*, 5th ed. London, U.K.: Pearson, 2014.

[27] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, Mar. 1993.

JOEL GOODMAN (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Boston University. He is currently pursuing the Ph.D. degree in systems engineering with The George Washington University, Washington, DC, USA.

He worked as a Lecturer with the Virginia Tech Symposium and Summer School on Wireless Communications, in 2014. He was previously employed by Eastman Kodak, worked for venture capital-funded Hammerhead Networks (later acquired by Cisco Systems) and a Technical Staff Member with the MIT Lincoln Laboratory (MIT LL), most recently working with the Chief Technology Officer's Technical Advisory Group. He is currently a Senior Research Engineer with the U.S. Naval Research Laboratory, Tactical Electronic Warfare Division.

Mr. Goodman was a recipient of the Eastman Technical Achievement Award for his work on magnetic imaging, an Invited Lecturer for an IEEE Advanced Signal Processing Symposium on the topic of Nonlinear Signal Processing, in 2008, also a recipient of the 2008 MIT LL Team Excellence Award for his work on Nonlinear Equalization, and also a recipient of the 2012 NRL Alan Berman Research Publication Award and NRL Review Award. He was a member of the Organizing Committee of IEEE GlobalSIP 2016 and an Invited Panel Member of the Full-Duplex Communications Session at the IEEE IMS Conference, in 2019. He is currently serving as the Chair of the IEEE Computational Intelligence Society Chapter in Washington.

MICHELLE MCBETH received the M.S. degree in electrical and computer engineering from Carnegie Mellon University, in 2004, with a focus on mixed-signal integrated circuit design.

Since then, she has worked as an Embedded Systems Designer with tasks, including embedded real-time software development, sensor integration, and FPGA design. She joined the U.S. Naval Research Laboratory, in 2007, as the Lead Software Engineer for several electronic countermeasure testing systems. She has been integral in successful field testing of numerous RF payloads. Her current research interest includes high-speed DSP system design for FPGAs.

MATTHEW B. MCKEON received the B.S. degree in electrical engineering from Clarkson University, in 2007, and the M.S. degree in electrical engineering from Syracuse University, in 2012.

From 2007 to 2011, he was an Electronics Engineer with Sensis/Saab Corporation of Syracuse, NY, USA. Since 2012, he has been an Electronics Engineer with the U.S. Naval Research Laboratory, Washington, DC. He has over 15 years of experience designing digital signal processors using both ASIC and FPGA technology. His current research interests include wideband signal processing, radar, and wireless communications.

Mr. McKeon was a recipient of the 2016 NRL Review Award and has one patent.

DAVID J. PARRETT received the B.S. degree in electrical engineering from the Rochester Institute of Technology, in 2005.

From 2005 to 2014, he was an Embedded Software Engineer with Sensis/Saab Corporation of Syracuse, NY, USA. Since 2014, he has been an Electronics Engineer with the U.S. Naval Research Laboratory, Washington.

KEVIN S. LORENZ received the dual B.S. degrees in electrical engineering and computer engineering from Michigan State University and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University.

Since 2012, he has been a Research Engineer with the U.S. Naval Research Laboratory, Tactical Electronic Warfare Division. He worked as a Lecturer with the Virginia Tech Symposium and Summer School on Wireless Communications, in 2014. He has over 15 years of signal processing, algorithm and software development, and modeling and simulation experience. His current research interests include linear and nonlinear signal processing for various applications, including imaging, radar, and wired and wireless communications. He has authored more than 20 articles.

Dr. Lorenz was a recipient of the 2012 NRL Jerome and Isabella Karle Distinguished Scholar Fellowship, 2016 NRL Review Award, and 2017 and 2020 NRL Alan Berman research publication awards.

● ● ●