Antenna-Independent Deep Learning for Self-Interference Cancellation in In-Band Full-Duplex MIMO Systems

Anh Tuyen Le[1], Xiaojing Huang[1], Le Chung Tran[2], J. Andrew Zhang[1], Peiyuan Qin[1], and Y. Jay Guo[1]

[1]Global Big Data Technologies Centre, University of Technology Sydney, Ultimo, NSW, 2007, Australia
Emails: {anhtuyen.le,xiaojing.huang,andrew.zhang,peiyuan.qin,jay.guo}@uts.edu.au
[2]Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong, NSW, 2522, Australia
Email: lctran@uow.edu.au

***Abstract*—In-band full-duplex (IBFD) operation in multiple-input multiple-output (MIMO) systems faces significant challenges due to self-interference (SI). The cancellation complexity of SI grows quadratically with the number of antennas, posing a critical scalability challenge for 6G and integrated sensing and communications (ISAC) applications. Existing self-interference cancellation (SIC) techniques, including both model-based and recent deep learning approaches, typically operate on a per-antenna basis and thus inherit this prohibitive complexity growth. This paper proposes a novel beam-based deep learning framework that fundamentally overcomes this limitation. By operating in the beam domain, our method models the interference at each receiver as a linear combination of the transmit beam signals, allowing the neural network to learn the end-to-end system effects without explicit channel estimation. The key advantage of this approach is that the network's complexity scales with the number of beams, making it independent of the number of transmit antennas and highly suitable for massive MIMO systems. Simulation results demonstrate that the proposed framework achieves over 33 dB of SIC, with performance remaining consistent even as the number of transmit antennas increases. This beam-domain approach offers a scalable and effective solution for practical IBFD implementation in future wireless networks.**

***Index Terms*—Integrated sensing and communications, in-band full-duplex, MIMO, deep learning, neural networks.**

I. Introduction

In-band full-duplex (IBFD) technology has emerged as a transformative solution for next-generation wireless communications, offering the potential to double spectral efficiency by enabling simultaneous transmission and reception on the same frequency band [1]. This capability is particularly crucial for 6G networks, where the exponential growth in data demands necessitates more efficient spectrum utilization to support emerging applications such as immersive cloud services, holographic communications, and augmented reality [2]. Beyond traditional communications, IBFD technology plays a pivotal role in enabling monostatic integrated sensing and communications (ISAC) systems, where the same infrastructure performs both communication and radar sensing functions [3]. The ITU's vision for 6G, known as IMT-2030, identifies ISAC as a revolutionary feature alongside enhanced mobile broadband (eMBB) and ultra-reliable low-latency communications (URLLC), with monostatic sensing configurations offering superior integration between sensing and communication functionalities [4].

The primary technical obstacle preventing widespread adoption of IBFD technology in MIMO systems is the presence of self-interference (SI), which becomes increasingly complex as the number of antennas grows. In an IBFD $N \times N$ antenna MIMO system, each receive chain experiences SI from N transmit antennas, resulting in a quadratic complexity growth that scales as $\mathcal{O}(N^2)$ [5]. This complexity explosion manifests in multiple dimensions: the naive replication of single-input single-output (SISO) designs requires N^2 cancellation circuits and digital signal processing implementations, consuming prohibitive hardware resources and power [6]. Furthermore, the spatial correlation between closely-spaced MIMO antennas creates strong cross-talk interference, often 75-80 dB stronger than desired received signals, which cannot be adequately addressed by simply replicating SISO cancellation techniques. The estimation error also scales linearly with the number of MIMO chains, as independent estimation algorithms at each receiver chain accumulate errors, leading to performance degradation that worsens with the increasing antenna count [7].

Recent advances in machine learning have led to the development of neural network-based alternatives to traditional polynomial models for self-interference cancellation (SIC), offering superior performance in modelling nonlinear distortions with reduced computational complexity [8]. Among these, feed-forward neural networks (FFNNs) have shown strong potential, achieving cancellation performance comparable to that of polynomial models but with significantly fewer parameters [9]. In particular, the Nonlinear Autoregressive Exogenous (NARX) neural network architecture has demonstrated an 83.3% reduction in computation compared to polynomial methods, while maintaining an equivalent cancellation performance of approximately 44 dB [10].

A persistent challenge in SIC lies in managing time-varying channels and hardware nonlinearities. Kim et al. [11] conducted a comprehensive comparison of model-based and model-free approaches, showing that while deep learning-aided SIC performs well under static conditions, it often struggles to track dynamic channel variations. To address the demands of online training, transfer learning techniques—such as Time-Invariant Distortion (TID) and Time-Varying Distortion (TVD)—have been proposed. These methods effectively reduce training overhead while achieving cancellation performance in the range of 42–43 dB [12].

However, most of the existing deep learning-based SIC methods have been developed for single-antenna systems. As highlighted in [8], extending these approaches to MIMO systems introduces substantial complexity, particularly when designing separate cancellation modules for each transmit antenna. To mitigate this, a recent deep learning architecture proposed in [13] leverages spatial correlation between MIMO channels to reduce model complexity. Nevertheless, the level of cancellation achieved by this method remains modest, reaching only around 18 dB.

This paper proposes a novel beam-based deep learning framework for digital SIC in IBFD MIMO systems that fundamentally addresses the complexity scalability challenge. Unlike existing approaches that process signals from individual transmit chains, our method operates in the beam domain by treating the interference at each receiver as a linear combination of beamformed data streams for different users. The proposed FFNN architecture is designed with complexity scaling as $\mathcal{O}(K)$ where K is the number of beams, making it independent of the number of transmit antennas and thus highly suitable for massive MIMO deployments. By leveraging the spatial correlation inherent in beamspace processing, the network implicitly learns the combined effects of beamforming matrices, transmit chains, and SI channels without requiring explicit channel state information or mathematical modeling of individual impairments. Simulation results demonstrate that the proposed approach achieves more than 32 dB of SIC performance regardless of the number of transmit antennas.

The rest of this paper is organized as follows. Section II presents the system models of an IBFD MIMO system. Section III describes the proposed beam-based deep learning method for SIC. Section IV presents simulation results, and finally, conclusions are drawn in Section V.

II. System Model

We consider an IBFD MIMO system equipped with N_t transmit (Tx) antennas and N_r receive (Rx) antennas as shown in Fig. 1. The system simultaneously transmits K independent data streams, each directed towards a specific angle θ_k for $k = 1, \ldots, K$. These data streams, $s_k(n)$, are multiplied by the corresponding steering vector $\mathbf{a}(\theta_k)$ and combined for transmission. The baseband equivalent vector of the transmitted down-link signals across the N_t antennas, denoted as $\mathbf{x}_{dl}(n)$, is expressed as:

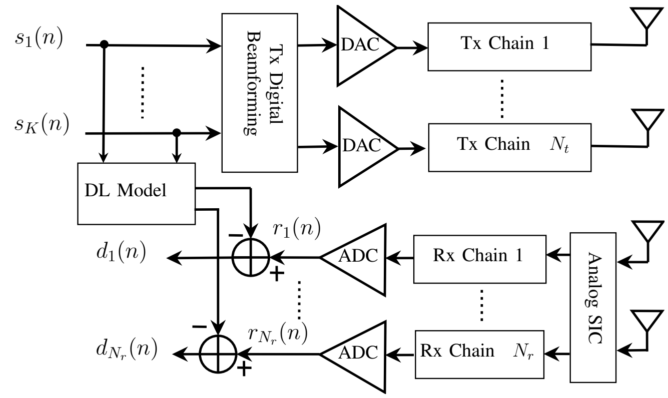

Fig. 1. A FD MIMO system model with deep learning digital SIC.

$$\mathbf{x}_{dl}(n) = \mathbf{A}\mathbf{s}(n) = \begin{bmatrix} a_0(\theta_1) & \cdots & a_0(\theta_K) \\ \vdots & \ddots & \vdots \\ a_{N_t-1}(\theta_1) & \cdots & a_{N_t-1}(\theta_K) \end{bmatrix} \begin{bmatrix} s_1(n) \\ \vdots \\ s_K(n) \end{bmatrix}, \tag{1}$$

where $\mathbf{A} \in \mathbb{C}^{N_t \times K}$ is the transmit steering matrix, whose columns are the array steering vectors; $\mathbf{a}(\theta_k) = [a_0(\theta_k), \ldots, a_{N_t-1}(\theta_k)]^T$ is the steering vector for the k-th beam. Each element is defined as $a_n(\theta_k) = e^{jn\frac{2\pi}{\lambda}d\sin\theta_k}$ for $n = 0, \ldots, N_t - 1$, where d is the antenna element spacing and λ is the wavelength. The Tx data streams are assumed to be mutually independent and have unit power, i.e.,

$$E\{s_k(n)^* s_{k'}(n')\} = \begin{cases} 1, & \text{if } k = k' \text{ and } n = n' \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

Furthermore, we assume all Tx and Rx chains are linear, and any amplifier gains are absorbed into the self-interference channel model.

In IBFD MIMO systems, the SI originating from the N_t Tx antennas poses a saturation risk to the N_r Rx antennas. To prevent this, an analog SIC architecture such as [6] is first employed to suppress the bulk of the SI signal. Following this stage and subsequent processing by the analog-to-digital converter (ADC), matched filter, and down-sampling, a residual SI component remains. Assuming the SI channel comprises L distinct multi-path components, this baseband equivalent residual SI vector, $\mathbf{z}(n) \in \mathbb{C}^{N_r \times 1}$, can be modeled as:

$$\mathbf{z}(n) = \sum_{l=0}^{L-1} \mathbf{H}^H(l)\mathbf{x}_{dl}(n-l), \tag{3}$$

where $\mathbf{x}_{dl}(n-l)$ is the transmitted signal vector delayed by l taps. The matrix $\mathbf{H}(l) \in \mathbb{C}^{N_t \times N_r}$ represents the *effective residual* SI channel for the l-th path. Its entries, $h_{n,m}(l)$, characterize the channel coefficients that persist after the analog cancellation signal has been subtracted, thereby modeling the SI that must be handled in the digital domain. By substituting the expression for the transmitted signal $\mathbf{x}_{dl}(n)$ from (1) into

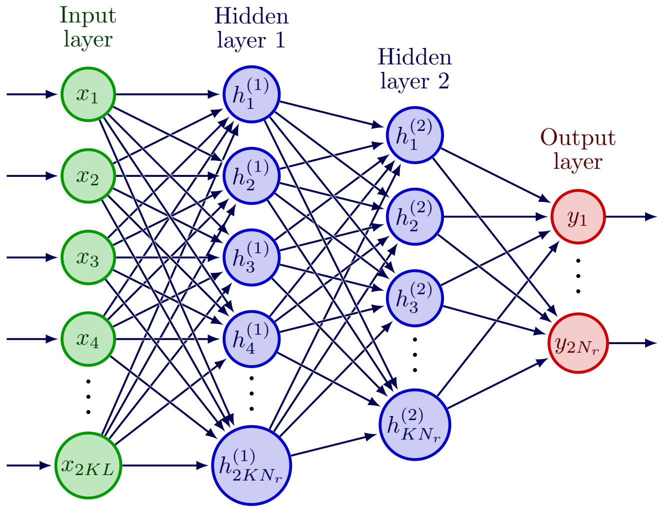

Fig. 2. The proposed neural network.

(3), we get

$$\mathbf{z}(n) = \sum_{l=0}^{L-1} \mathbf{H}^H(l)\mathbf{A}\mathbf{s}(n-l). \tag{4}$$

Denoting $\tilde{\mathbf{H}} = \mathbf{H}^H\mathbf{A} \in \mathbb{C}^{N_r \times K}$, we have $\mathbf{z}(n) = \sum_{l=0}^{L-1} \tilde{\mathbf{H}}(l)\mathbf{s}(n-l)$, representing the residual digital SI signal as a linear combination of the Tx beam data. This observation suggests that the cancellation signal can be generated directly from the Tx beam data, while the combined effects of Tx beamforming, SI channel, residual SI channel after analog SIC, and receiver chain can be estimated by the deep-learning method.

Denoting the received up-link signals from K far-end users in the digital domain as $\mathbf{x}_{ul}(n)$, the received signal before digital SIC is expressed as

$$\mathbf{r}(n) = \mathbf{z}(n) + \mathbf{x}_{ul}(n) + \mathbf{n}(n), \tag{5}$$

where $\mathbf{r}(n) = [r_1(n), \ldots, r_{N_r}(n)]^T$ and $\mathbf{n}(n)$ are the received baseband signals and noise from N_r Rx chains, respectively.

III. Beam-based Deep Learning Network

To learn the complex mapping from the Tx beam signals to SI signal, we employ a deep FFNN whose design and training process are tailored specifically for the beam-based cancellation problem. The FFNN structure is not arbitrary but is methodically designed based on the physical parameters of the MIMO system: the number of beams (K) and the number of Rx antennas (N_r). This beam-centric design ensures that the network complexity scales with the operational parameters of the system, not just the raw number of Tx antennas. The network consists of a sequence of fully connected layers, each followed by a Rectified Linear Unit (ReLU) activation function and a batch normalization layer to stabilize and accelerate training. The architecture shown in Fig. 2 is defined as follows:

- **Input Layer:** An input layer that accepts the normalized feature vector of size $2KL$.
- **First Hidden Layer:** A fully connected layer with $n_1 = 2KN_r$ neurons. This layer is designed to learn the initial, complex combinations of all beam signals as they contribute to the interference at each of the N_r receivers.
- **Second Hidden Layer:** A fully connected layer with $n_2 = KN_r$ neurons, designed to refine the feature representations learned in the first layer.
- **Output Layer:** A fully connected layer with $n_3 = 2N_r$ output neurons, corresponding to the real and imaginary parts of the estimated SI signal at each of the N_r receive antennas. This is followed by a regression layer to produce the final continuous-valued output.

The network is trained using a supervised learning paradigm. Prior to training, both the input feature vectors and the target output vectors are normalized by subtracting the mean and dividing by the standard deviation of the training set. This per-feature normalization is critical for ensuring stable and efficient convergence. The training process aims to minimize a cost function, defined as the Mean Squared Error (MSE) between the network's predicted SI signal, $\hat{\mathbf{z}}(n)$, and the true SI signal, $\mathbf{z}(n)$, observed in the training data. To perform this minimization, we employ the Adam optimizer, an adaptive moment estimation variant of stochastic gradient descent. The gradients of the cost function with respect to each of the network's weights are efficiently computed using the back-propagation algorithm.

Training is conducted over 300 epochs with a mini-batch size of 64, promoting stable gradient updates and improved generalization. The initial learning rate is set to 0.001, with a piecewise decay schedule that reduces the rate by a factor of 0.8 every 50 epochs to ensure convergence. L2 regularization with a factor of 0.0005 is applied to penalize large weights, mitigating overfitting, particularly given the large hidden layers ($2K \cdot Nr$). Validation data is used to monitor performance, with a validation frequency of 50 iterations, and early stopping is employed to halt training if the validation loss plateaus, restoring the best weights to prevent overfitting.

The proposed beam-based deep learning approach offers several distinct advantages over traditional model-based SIC methods, primarily in complexity, scalability, and robustness. Foremost among these is a significant reduction in computational overhead. By learning the end-to-end mapping from the transmit beam signals to the received SI, the method circumvents the need to explicitly estimate and track the large set of $N_t \times N_r \times L$ channel parameters, which is a major bottleneck in conventional techniques. This leads directly to superior scalability; the network's complexity scales favorably with the number of beams (K) rather than exhibiting the quadratic growth with the antenna number (N_t) that plagues model-based methods, making it exceptionally well-suited for massive MIMO systems. Furthermore, the data-driven nature of the neural network provides inherent robustness to real-world imperfections [14]. It implicitly learns to compensate for model mismatches and unknown hardware impairments that are difficult to characterize analytically, which is a common challenge for methods reliant on precise system models. It

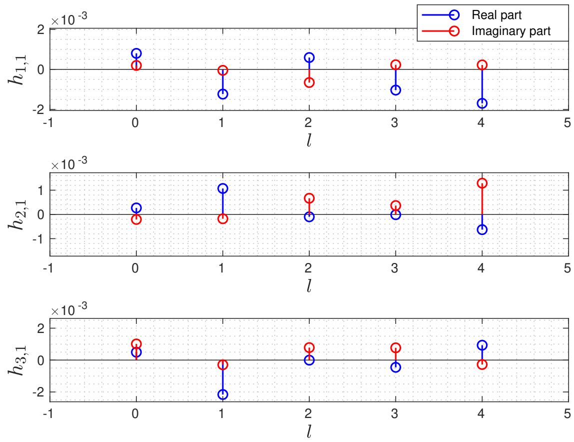

Fig. 3. One realization of SI channels from three Tx antennas to the first receiver.

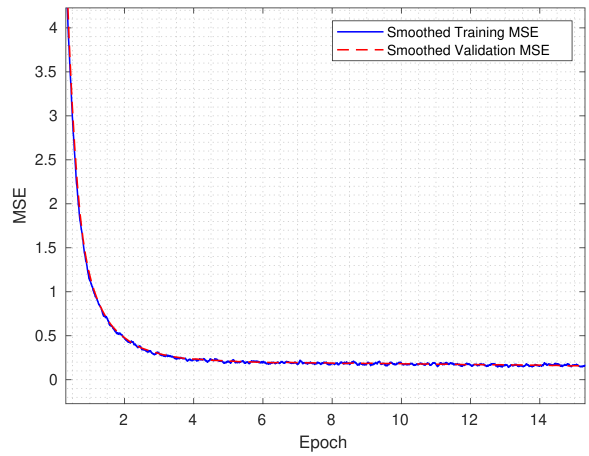

Fig. 4. Smoothed training and validation mean squared error (MSE) versus epoch.

is important to note that the successful training of this data-driven model is contingent on the availability of synchronized Tx-Rx data pairs, which are necessary to establish the ground truth for the supervised learning process.

IV. Simulation Results

A. Simulation Setup

To evaluate the performance and scalability of the proposed beam-based deep learning method, we conduct a series of simulations using MATLAB. We configure an IBFD MIMO system with a fixed number of Rx antennas, $N_r = 4$, and $K = 4$ simultaneous transmit beams. The key parameter under investigation is the number of Tx antennas, which is varied as $N_t \in \{8, 12, 16, \ldots, 32\}$, to demonstrate that the cancellation performance is independent of N_t. The SI channel is modeled with a memory length of $L = 5$ taps to account for multi-path effects. Fig. 3 shows one realization of SI channels from the three Tx antennas to the first receiver. To train the neural network, a dataset is generated under ideal conditions, containing only SI signal and additive noise with power 20 dB lower than that of the SI. The system transmits a total of 50,000 QPSK-modulated symbols for each of the K beams, which are spatially distributed across a 120-degree sector. The dataset is then divided into 80% for training and 20% for validation.

B. Simulation Results

We first evaluate the convergence of the proposed deep learning model during training for a representative scenario with $N_t = 8$ transmit antennas, $N_r = 4$ receive antennas, and $K = 4$ beams. Figure 4 shows the smoothed mean squared error (MSE) curves for the training and validation datasets over the course of model training. Clearly, both MSEs rapidly decline during the first few epochs, indicating effective early learning. By around epoch 5, both training and validation MSEs fall below 0.1, after which they stabilize with minimal fluctuation. The close match between the two curves throughout training suggests strong generalization and no significant overfitting. This behavior demonstrates that the model is well-regularized and effectively fitted to the data.

Once training is complete, the neural network is used to predict the SI signal, allowing for its cancellation. To evaluate the model's performance, a separate test dataset is generated, consisting of 50,000 QPSK-modulated symbols per beam across the same K beams and 120-degree spatial sector. Unlike the training set, the test data includes both desired far-end signals—60 dB weaker than the SI—and receiver noise at a level 20 dB below the far-end signal power. This configuration reflects a more realistic operational scenario, enabling assessment of the model's cancellation performance under practical conditions.

To demonstrate the robustness and scalability of our beam-domain approach, we then evaluate the cancellation performance while varying the number of transmit antennas from $N_t = 8$ to $N_t = 32$. Throughout this experiment, the number of beams ($K = 4$), receive antennas ($N_r = 4$), and the neural network architecture itself are held constant. Denoting the cancellation signals generated by the FFNN as $\mathbf{y}(n) = [y_1(n), \ldots, y_{N_r}(n)]^T$, the SIC performance is evaluated using the Self-Interference Suppression (SIS) metric:

$$\text{SIS (dB)} = 10 \log_{10} \left(\frac{E\{|\mathbf{z}(n)|^2\}}{E\{\left|\mathbf{z}(n) - \mathbf{y}(n)\right|^2\}} \right). \tag{6}$$

Fig. 5 plots the achieved cancellation depth as a function of the number of transmit antennas. The results confirm that our method consistently achieves a cancellation level of over 33 dB across the entire range of N_t values tested. This outcome validates the central premise of our work: the cancellation performance is effectively decoupled from the size of the transmit array, providing a truly scalable solution essential for massive MIMO IBFD systems.

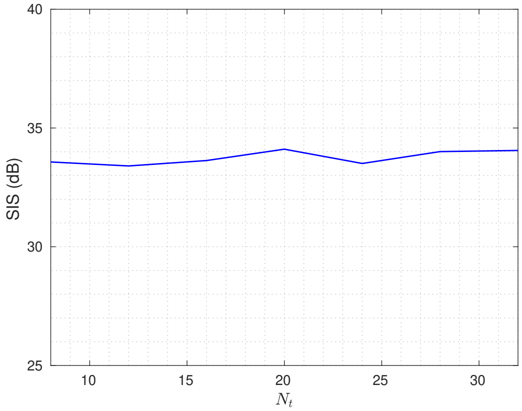

Fig. 5. SIS levels versus different number of transmit antennae.

C. Complexity Analysis

The proposed neural network consists of four fully connected layers: an input layer of size $2KL$, two hidden layers of sizes $2KN_r$ and KN_r, and an output layer of size $2N_r$, where K is the number of transmit beams (or transmit antennas), L is the memory length, and N_r is the number of receive antennas. The trainable parameters in the network consists of weights and biases between each pair of consecutive layers. Hence, the total number of trainable parameters are

$$P_{\text{total}} = 4K^2LN_r + 2KN_r + 2K^2N_r^2 + KN_r + 2KN_r^2 + 2N_r, \tag{7}$$

or in Big-O notation: $P_{\text{total}} = \mathcal{O}(K^2LN_r + K^2N_r^2)$. During inference, each layer requires a matrix multiplication and bias addition. For a single input sample, the number of floating-point operations (FLOPs) is dominated by matrix multiplications between layers. The computation from the input layer to the first hidden layer involves $2 \times (2KL)(2KN_r) = 8K^2LN_r$ FLOPs. From the first to the second hidden layer, the operation requires $2 \times (2KN_r)(KN_r) = 4K^2N_r^2$ FLOPs. Finally, the transformation from the second hidden layer to the output layer contributes an additional $2 \times (KN_r)(2N_r) = 4KN_r^2$ FLOPs. Thus, the total FLOPs per input sample is calculated as

$$\begin{aligned} \text{FLOPs}_{\text{total}} &= 8K^2LN_r + 4K^2N_r^2 + 4KN_r^2 \\ &= \mathcal{O}(K^2LN_r + K^2N_r^2). \end{aligned} \tag{8}$$

This analysis highlights that both the parameter count and inference cost grow quadratically with the number of transmit beams K and receive antennas N_r. However, in practical full-duplex systems, K is often much smaller than the number of antennas. Under the assumption that $K \ll N_r$, the overall complexity reduces to:

$$P_{\text{total}} = \mathcal{O}(LN_r + N_r^2), \quad \text{FLOPs}_{\text{total}} = \mathcal{O}(LN_r + N_r^2).$$

Clearly, both the train variables and FLOPs complexity of the network are functions of the number of beams (K), the number of receive antennas (N_r), and the channel memory length (L), but are entirely independent of the number of Tx antennas (N_t). This property is the fundamental advantage of the proposed beam-based approach, ensuring that the model's complexity does not grow with the size of the Tx array, making it an ideal and scalable solution for massive MIMO systems.

V. Conclusion

This paper presented a deep learning framework for SIC in IBFD MIMO systems that addresses the critical challenge of complexity scaling. By operating in the beam domain, the proposed method's complexity scales with the number of beams rather than the number of Tx antennas. This approach circumvents the quadratic complexity growth found in conventional per-antenna SIC techniques. Simulation results confirm that the framework achieves over 33 dB of SIC and this performance level remains consistent as the Tx array size increases. Future work will focus on extending this framework to dynamic environments. The application of transfer learning or reinforcement learning presents a promising path for adapting the model to time-varying SI channels with minimal retraining overhead.

References

[1] A. Sabharwal, *et al.*, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sept. 2014.

[2] B. Smida, *et al.*, "Full-duplex wireless for 6G: Progress brings new opportunities and challenges," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 9, pp. 2729–2750, Sept. 2023.

[3] A. T. Le, *et al.*, "Joint analog and digital interference cancellation for in-band full-duplex ISAC systems," Jan. 2025. [Online]. Available: http://dx.doi.org/10.36227/techrxiv.173808296.67688198/v1

[4] ITU-R, "Framework and overall objectives of the future development of IMT for 2030 and beyond," International Telecommunication Union, Geneva, Switzerland, Recommendation ITU-R M.2160-0, Nov. 2023.

[5] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proc. 11th USENIX Conf. Networked Systems Design Implementation*, ser. NSDI'14. Berkeley, CA, USA: USENIX Association, 2014, pp. 359–372.

[6] A. T. Le, *et al.*, "Beam-based analog self-interference cancellation in full-duplex MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2460–2471, 2020.

[7] J. Vieira, *et al.*, "Reciprocity calibration for massive MIMO: Proposal, modeling, and validation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3042–3056, 2017.

[8] M. Elsayed, *et al.*, "Machine learning-based self-interference cancellation for full-duplex radio: Approaches, open challenges, and future research directions," *IEEE Open Journal of Vehicular Technology*, vol. 5, pp. 21–47, 2024.

[9] A. Balatsoukas-Stimming, "Non-linear digital self-interference cancellation for in-band full-duplex radios using neural networks," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018, pp. 1–5.

[10] Q. Dong, A. C. M. Austin, and K. W. Sowerby, "Augmentation of self-interference cancellation for full-duplex using NARX neural networks," *IEEE Wireless Communications Letters*, vol. 13, no. 3, pp. 810–813, 2024.

[11] J. Kim, *et al.*, "On the learning of digital self-interference cancellation in full-duplex radios," *IEEE Wireless Communications*, vol. 31, no. 4, pp. 184–191, 2024.

[12] K. Muranov, *et al.*, "On deep learning assisted self-interference estimation in a full-duplex relay link," *IEEE Wireless Communications Letters*, vol. 10, no. 12, pp. 2643–2647, 2021.

[13] Y. Chen, *et al.*, "Mimo full duplex radios with deep learning," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2020, pp. 1–6.

[14] M. H. Meng, *et al.*, "Adversarial robustness of deep neural networks: A survey from a formal verification perspective," *IEEE Transactions on Dependable and Secure Computing*, vol. 99, no. 99, pp. 184–191, 2022.