Nonlinear Self-Interference Cancellation via U-Net Variational Autoencoder with Multi-Feedback for In-Band Full-Duplex Communications

*HongYu ZHU, WeiWei WU *, ZiHao ZHENG, JingJian HUANG *, QingZhan SHI, XiaoFa ZHANG, NaiChang YUAN*

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

wuweiwei@nudt.edu.cn, ZhuHongYu6868@163.com, zhengzihao19@nudt.edu.cn, hjjfh2013@163.com, qingzhanshi@foxmail.com, zhangxiaofa@163.com, yncwww@qq.com

Submitted July 16, 2026 / Accepted September 30, 2026 / Online first November 20, 2026

Abstract. *In-band full-duplex (IBFD) technology promises to double spectral efficiency, yet its deployment is fundamentally constrained by self-interference (SI) arising from coupled RF impairments. Existing SI cancellation methods predominantly rely on linear models that treat RF impairments in isolation, leaving the nonlinear coupling among in-phase/quadrature (I/Q) imbalance, phase noise, power amplifier (PA) nonlinearity, and multipath propagation unaddressed—a critical barrier to practical IBFD implementation. Here we depart from single-impairment models by constructing a comprehensive SI channel representation that captures cross-domain coupling effects, and propose a multi-feedback SI cancellation framework, termed U-Net Variational Autoencoder with Multi-Feedback (U-VAE-MF). U-VAE-MF integrates U-Net hierarchical feature extraction, VAE latent-space regularization, and a multi-feedback mechanism for iterative residual correction. The multi-feedback mechanism pre-compensates dominant transmitter nonlinearities, simplifying the residual interference pattern and reducing the feature space dimensionality for digital-domain estimation. Under composite non-ideal conditions, the proposed multi-feedback stage suppresses the residual self-interference to –56 dBm at 0 dBm input power, a 38 dB improvement over conventional least-squares (LS) cancellation. The complete U-VAE-MF canceller operates within 2 dB of the theoretical cancellation limit, and it reduces the SI channel estimation latency from 270 μs to 100 μs (63% reduction).*

Keywords

In-band full-duplex technology; RF impairments; U-net variational autoencoder with multi-feedback

1. Introduction

In-band full-duplex (IBFD) technology promises to double spectral efficiency by enabling simultaneous transmission and reception in the same frequency band, thereby substantially increasing user capacity and improving communication quality under spectrum-constrained conditions. IBFD has attracted considerable research interest for its potential applications in satellite communications, relay systems, and 5G/6G networks [1–3]. However, the challenge of suppressing severe self-interference (SI) remains a major technical bottleneck. SI arises from the combined effects of coupled non-ideal RF characteristics and insufficient antenna isolation [4], [5]. High-power transmit signals leak to the receiver with minimal attenuation, raising the system noise floor and severely degrading the demodulation and recovery performance of the received signal [6–8].

Extensive research has examined the various sources of self-interference [9–11]. Non-ideal RF front-ends inherently exhibit nonlinear and time-varying characteristics; meanwhile, device miniaturization has reduced the physical spacing between antennas. Together, these factors cause severe propagation-domain coupling effects [12], [13]. While most studies have focused on improving self-interference isolation, few have addressed the dual challenge of maintaining high real-time performance and sufficient isolation in complex time-varying channels. This paper addresses this gap by focusing on fast nonlinear correction, self-interference channel estimation, and reconstruction-based cancellation for linear frequency-modulated (LFM) signals, which serve here as a representative waveform and are complemented by the orthogonal frequency-division multiplexing (OFDM) waveform in the comparative analysis.

Existing mitigation schemes implement interference cancellation across three domains: the propagation domain, the analog domain, and the digital domain [14–17]. Propagation-domain techniques suppress transmitter signal leakage [18]; analog-domain methods attenuate dominant SI components using analog circuits [19–21]; and digital-domain approaches perform fine-grained residual suppression to recover the signal of interest (SOI) near the noise floor [22–25].

DOI: 10.13164/re.2026.0559

Prior work has made notable progress in each domain separately. J.-H. Lee achieved significant self-interference reduction through analog cancellation based on rotated baseband signals and maximum-likelihood estimation [7]. Fukui et al. pushed digitally-assisted RF-domain cancellation toward theoretical limits by suppressing nonlinear interference in simulation [19]. Kong et al. proposed a neural-network-based digital self-interference cancellation (SIC) method that adapts to time-varying channels and reduces interference toward the noise floor [12]. More recently, a rich body of deep-learning-based digital SIC schemes has emerged, including the deep-learning SIC (DSIC) architecture built on a deep neural network (DNN) [15, 26–29], pioneering neural-network nonlinear cancellers [30], low-complexity and hybrid-layer neural-network (NN) structures for SI modeling [31], [32], and dual-neuron NN [1]. Other representative schemes include an attention-based combination of a temporal convolutional network and a gated recurrent unit (TCN-GRU) [17], convolutional long short-term memory (ConvLSTM) networks [33], a sliding-window GRU with temporal pattern attention (SW-GRU-TPA) [34], and fast Kolmogorov–Arnold networks (KANs) [35]. A comprehensive survey of machine-learning-based SIC approaches is given in [36]. Taken together, these studies differ mainly in the network used to represent the nonlinearity rather than in the set of impairments they model: only a few account for time variation explicitly [12], and none combines transmitter-side pre-compensation with a latent-regularized receiver-side canceller. They either address individual impairments in isolation, rely on a single neural architecture without explicit latent-space regularization, or separate transmitter-side linearization from receiver-side digital cancellation; none has demonstrated joint suppression of the coupled nonlinearities—in-phase/quadrature (I/Q) imbalance, power amplifier (PA) distortion, phase noise, and frequency-selective fading—under real-time constraints [33, 37–39]. Compared with these prior deep-learning methods, our work differs in three respects. First, we model the coupled interaction of all dominant RF impairments in a unified signal model, rather than a single impairment. Second, we jointly design a transmitter-side multi-feedback predistortion stage and a receiver-side U-VAE cancellation stage, so that the network only has to handle the reduced residual. Third, we regularize the U-Net latent space with a variational autoencoder (VAE) module, which prior SIC networks lack. In addition, we derive the closed-form theoretical cancellation limit against which the learned canceller is verified.

To address this gap, we propose a two-stage approach: (1) a transmitter-side multi-feedback linearization technique that pre-compensates I/Q imbalance and PA nonlinearity, emitting a near-ideal linear signal and thereby reducing the algorithmic burden at the receiver; and (2) a digital-domain cancellation method based on a U-Net Variational Autoencoder (U-VAE) architecture that leverages hierarchical feature extraction and latent-space regularization to handle residual nonlinear interference in dynamic scenarios.

Our key contributions are as follows:

1. **Comprehensive RF impairment model.** We develop an integrated signal model that captures I/Q imbalance, phase noise, PA nonlinearity with memory effects, multipath propagation, quantization noise, and frequency response mismatches, including cross-coupling effects neglected by conventional single-impairment models [40–42].

2. **Multi-feedback transmitter linearization.** A closed-loop mechanism dynamically estimates and compensates for I/Q imbalance and PA nonlinearity via digital predistortion. By pre-suppressing dominant distortions at the transmitter, this scheme reduces the feature space dimensionality required for digital-domain channel estimation, cutting the estimation time by 63% relative to conventional least-squares estimation.

3. **U-VAE-based digital SIC.** A U-Net encoder-decoder augmented with VAE regularization directly extracts the SOI from the mixed received signal. The VAE's latent-space regularization ensures robust reconstruction under distribution shifts caused by time-varying nonlinear distortion, and its effect is directly observable in Fig. 8, where the same U-Net without the VAE module deviates by about 15 dB from the ground truth at the main-signal frequency on unseen samples. The complete scheme achieves a cancellation ratio within 2 dB of the theoretical limit across a 100–500 MHz bandwidth.

4. **Positioning with respect to prior deep-learning SIC.** Table 1 compares U-VAE-MF with recent deep-learning SIC methods [1, 12, 15, 17, 31–35] and discusses the roles of the multi-feedback module, U-Net skips, and VAE regularization. Unlike prior schemes with partial or no RF-impairment modeling, U-VAE-MF covers I/Q imbalance, phase noise, PA nonlinearity, multipath fading, and quantization error; combines TX-side feedback compensation with latent-space regularization; and derives a closed-form limit. It is the only compared scheme verified against this physical bound, not just an isolated metric.

The remainder of this paper is organized as follows. Section 2 establishes the comprehensive RF non-ideality model and system architecture. Section 3 details the U-VAE-MF framework, including transmitter linearization and U-VAE-based digital cancellation. Section 4 derives the error analysis and theoretical performance limits. Section 5 evaluates the methodology through simulations under full-impairment channel scenarios, and reports the comparative performance analysis together with the component contribution analysis. Section 6 concludes with a discussion of limitations and future directions.

Method (year, architecture)	Core Features
DSIC [15] (2019, DNN)	No RF impairment modeling
Low-complexity NN [31] (2021, DNN)	Nonlinear RF impairment modeling
Hybrid-layer NN [32] (2022, DNN)	Nonlinear RF impairment modeling
Dual-neuron NN [1] (2022, DNN)	Nonlinear RF impairment modeling
NN-aided SIC [12] (2022, DNN)	Time-varying modeling
ConvLSTM [33] (2022, ConvLSTM)	Nonlinear RF impairment modeling
TCN-GRU+Att. [17] (2024, TCN-GRU+Att.)	Nonlinear RF impairment modeling
SW-GRU-TPA [34] (2024, GRU+Att.)	Nonlinear RF impairment modeling
Fast KAN [35] (2026, KAN)	Nonlinear RF impairment modeling (OFDM self-interference)
U-VAE-MF (proposed, U-Net+VAE)	**Full RF impairment modeling, Tx compensation, closed-form limit**

Tab. 1. Comparison of the proposed U-VAE-MF with representative deep-learning-based SIC methods.

2. System Signal Model

We derive a unified digital-domain signal model for the key non-ideal effects in IBFD systems, covering I/Q imbalance, phase noise, PA nonlinearity, memory effects, additive white Gaussian noise (AWGN), and quantization noise [40–42]. The SI signal is decomposed into a linear combination of a reference signal and its distortion-induced derivatives, with equivalent channel parameters estimated via least-squares fitting.

2.1 Local Full-Duplex Transmitter

We model the transmitter signal chain as follows. The LFM waveform serves as the input excitation; the analysis extends to OFDM and DFT-spread OFDM waveforms without loss of generality.

The baseband LFM signal, centered at DC with RF carrier frequency f_0 and chirp rate $k = B/T_\mathrm{p}$, is given by:

$$x[n] = \exp(\mathrm{j}\pi \frac{B}{T_\mathrm{p}} n^2) \tag{1}$$

where B is the signal bandwidth and T_p is the pulse duration. After baseband signal generation, the waveform passes through a low-pass filter for bandwidth limiting and out-of-band spur suppression prior to digital-to-analog conversion (DAC) and up-conversion. I/Q imbalance in the up-conversion chain introduces undesired conjugate (mirror-frequency) components into the up-converted RF signal. The output of the up-converter is therefore:

$$x_{\mathrm{IQ,TX}}[n] = (x[n] + \eta_{\mathrm{TX}} x^*[n]) \exp(\mathrm{j}2\pi f_0 n + \mathrm{j}\Phi(n)) \tag{2}$$

where $(\cdot)^*$ denotes the complex conjugate. For the zero-IF IBFD transmitter architecture considered here, the local oscillator (LO) operates at the RF carrier frequency f_0. The amplitude of the undesired conjugate (mirror-frequency) component introduced by transmitter I/Q imbalance is proportional to the imbalance coefficient η_{TX}. The transmitter image rejection ratio (IRR) quantifies the suppression of this conjugate image component, defined as:

$$\mathrm{IRR} = 10 \log_{10} \left(\frac{P_{\mathrm{desired}}}{P_{\mathrm{image}}} \right) \tag{3}$$

where P_{desired} denotes the power of the desired fundamental signal, and P_{image} denotes the power of the undesired conjugate image component. We next model the LO phase noise, a critical impairment that limits self-interference cancellation performance. While the phase noise process $\Phi(t)$ is non-stationary, its incremental process $\Delta\Phi(\tau) = \Phi(t+\tau) - \Phi(t)$ follows a zero-mean Gaussian distribution. For a Lorentzian phase noise spectrum, the variance of the incremental phase noise is $\mathrm{var}[\Delta\Phi(\tau)] = 2\pi\beta|\tau|$, where β denotes the single-sided 3-dB bandwidth of the LO phase noise spectrum. Incorporating the nonlinear characteristics of the power amplifier, the output signal is:

$$x_{\mathrm{PA,TX}}[n] = f_{\mathrm{Rapp}}^{\mathrm{PA}}(\widetilde{x}_{\mathrm{IQ,TX}}[n]) \exp(\mathrm{j}\Phi(n)) + Z_{\mathrm{TX}}[n] \tag{4}$$

where

$$\widetilde{x}_{\mathrm{IQ,TX}}[n] = (x[n] + \eta_{\mathrm{TX}} x^*[n]) \exp(\mathrm{j}2\pi f_0 n) \tag{5}$$

where $Z_{\mathrm{TX}}[n]$ denotes the transmitter AWGN. We adopt the widely accepted Rapp model to characterize the memoryless nonlinear behavior of solid-state power amplifiers (SSPAs) commonly used in IBFD systems. The Rapp model was introduced by C. Rapp in 1991 to describe the amplitude-to-amplitude (AM-AM) conversion of SSPAs in digital broadcasting systems [43]. It is an amplitude-only model, so it neglects amplitude-to-phase (AM-PM) conversion; this makes it adequate for typical SSPAs whose phase distortion is small, and it parameterizes the compression curve with only three physically meaningful quantities: the small-signal voltage gain g, the output saturation voltage V_{sat}, and the smoothness factor s. The normalized form of the model reads:

$$f_{\mathrm{Rapp}}^{\mathrm{PA}}(x) = \frac{gx}{\left(1 + \left(\frac{g|x|}{V_{\mathrm{sat}}}\right)^{2s}\right)^{1/(2s)}}. \tag{6}$$

Here, when the input magnitude satisfies $g|x| \ll V_{\mathrm{sat}}$, the model reduces to the linear law $f_{\mathrm{Rapp}}^{\mathrm{PA}}(x) \approx gx$, whereas for $g|x| \gg V_{\mathrm{sat}}$ the output saturates at V_{sat}; the exponent s (smoothness factor) controls the sharpness of the transition from the linear to the saturation region. Because it captures the dominant AM-AM compression of SSPAs with a minimal parameter set, the Rapp model is a de facto standard benchmark in self-interference-cancellation and full-duplex transceiver studies, and it is therefore adopted here.

Figure 1 illustrates the nonlinear characteristics of the adopted Rapp model and the resulting intermodulation distortion under dual-tone excitation. Figure 1(a) plots the input-output power transfer curve of the PA under parameters $V_{\text{sat}} = 20\,\text{dBm}$, $s = 3$, and $g = 43\,\text{dB}$. The curve exhibits a smooth transition from the linear amplification region to the saturation region, with two representative operating points marked at $-28\,\text{dBm}$ and $-17\,\text{dBm}$ input power. Figure 1(b) compares the spectra of the dual-tone input signal and the PA output signal, showing that PA nonlinearity generates numerous spurious intermodulation components around the fundamental frequencies. Figures 1(c) and 1(d) present the time-domain waveforms of the input and output signals, respectively, where amplitude compression and envelope distortion caused by PA saturation are clearly observable.

Figure 2 illustrates two key impairments that limit self-interference cancellation performance; both are analyzed in detail in the following sections. Figure 2(a) presents the frequency-domain response of a 3-transmit-3-receive linear phased array antenna, where the non-flat amplitude response demonstrates the frequency-selective characteristic of the antenna channel—the primary source of the frequency mismatch effect discussed below. Figure 2(b) plots the power levels of quantization noise and Gaussian noise versus receiver input power. Both noise components increase monotonically with input power, and the Gaussian noise power remains consistently above the half-duplex noise floor across the entire dynamic range.

Following power amplification, the transmit signal propagates through the frequency-selective multipath channel of the antenna, which introduces memory effects. The frequency-domain antenna response term $Y_{\text{sig}}[n]$ is defined as:

$$Y_{\text{sig}}[n] = \frac{1}{N}\sum_{k=0}^{N-1} S[k]X_{\text{PA,TX}}[k]\exp\left(\text{j}\frac{2\pi}{N}kn\right). \quad (7)$$

Then the received signal at the antenna output is expressed in the discrete-time domain as:

$$y_{\text{RX}}[n] = Y_{\text{sig}}[n] + x_{\text{SOI}}[n] + Z_{\text{RX}}[n] \quad (8)$$

where $X_{\text{PA,TX}}[k]$ denotes the N-point discrete Fourier transform (DFT) of the PA output signal, defined as:

$$X_{\text{PA,TX}}[k] = \sum_{n=0}^{N-1} x_{\text{PA,TX}}[n]\exp(-\text{j}\frac{2\pi}{N}kn). \quad (9)$$

$S[k]$ is the equivalent frequency-domain transfer function of the antenna channel. In (8), $x_{\text{SOI}}[n]$ represents the signal of interest, and $Z_{\text{RX}}[n]$ denotes the receiver AWGN. The equivalent channel transfer function $S[k]$ is further decomposed to account for the amplitude-phase control vectors used in self-interference cancellation:

$$S[k] = \varpi_{\text{TX}}[k]\cdot H[k]\cdot \varpi_{\text{RX}}[k] \quad (10)$$

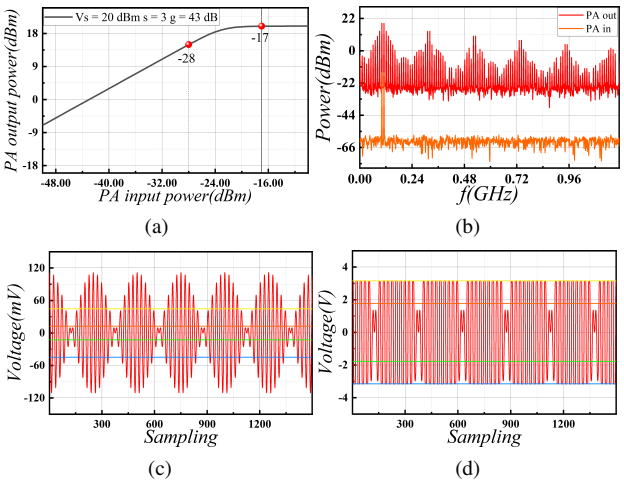


Fig. 1. Intermodulation distortion of dual-tone signals and nonlinear effects of power amplifiers based on the Rapp model with $V_{\text{sat}} = 20\,\text{dBm}$, $s = 3$, and $g = 43\,\text{dB}$. (a) Input-output power transfer characteristics of the power amplifier; the two markers indicate operating points at $-28\,\text{dBm}$ and $-17\,\text{dBm}$ input power. (b) Spectra of dual-tone input and output signals; intermodulation products appear around the fundamentals. (c) Time-domain waveform of the input signal. (d) Time-domain waveform of the output signal, showing amplitude compression and envelope distortion caused by saturation.

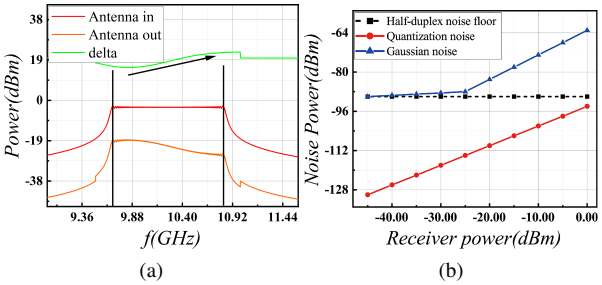


Fig. 2. Input-output performance of transmit-receive antenna systems. (a) Spectrum of linear frequency-modulated signals before and after passing through antennas in frequency-nonuniform channels. (b) Relationship between Gaussian white noise, quantization noise power, and received signal power in receivers.

where $H[k]$ is the intrinsic frequency response of the antenna multipath channel, $\varpi_{\text{TX}}[k]$ and $\varpi_{\text{RX}}[k]$ are the transmit and receive amplitude-phase control vectors, respectively, and $\cdot$ denotes the element-wise (Hadamard) product. The self-interference signal component entering the receiver can therefore be adaptively adjusted by modifying $\varpi_{\text{TX}}[k]$ and $\varpi_{\text{RX}}[k]$.

By the convolution theorem (frequency-domain multiplication corresponds to time-domain convolution), the received signal at the receiver front-end is equivalently expressed in the time domain as:

$$y_{\text{RX}}[n] = \sum_{l=0}^{L-1} s[l]x_{\text{PA,TX}}[n-l] + x_{\text{SOI}}[n] + Z_{\text{RX}}[n] \quad (11)$$

where L denotes the maximum memory length of the equivalent multipath channel, and $s[l]$ is the discrete-time impulse response of the equivalent channel, obtained as the N-point inverse discrete Fourier transform (IDFT) of the frequency-domain transfer function $S[k]$ defined in (10).

The Rician factor K of the equivalent channel, quantifying the ratio of direct-path to multipath power, is defined as:

$$K = \frac{\mathbb{E}\left[\left|s[0]x_{\mathrm{PA,TX}}[n]\right|^2\right]}{\mathbb{E}\left[\left|\sum_{l=1}^{L-1} s[l]x_{\mathrm{PA,TX}}[n-l]\right|^2\right]}. \tag{12}$$

2.2 Local Receiver

Following the antenna output, the signal passes through the receiver low-noise amplifier (LNA). Unlike the PA, which operates in the saturation region, the LNA operates in the small-signal linear region to minimize noise figure, and its nonlinearity is negligible for IBFD self-interference cancellation applications. The output of the LNA is:

$$y_{\mathrm{LNA,RX}}[n] = G_{\mathrm{LNA}} \cdot y_{\mathrm{RX}}[n] \tag{13}$$

where G_{LNA} denotes the linear voltage gain of the LNA. The amplified RF signal is then down-converted to baseband by the shared LO, which introduces phase noise into the received signal. The order of impairments in the receiver down-conversion chain is: LO phase noise first, followed by I/Q imbalance during quadrature demodulation. The baseband signal after down-conversion and I/Q demodulation is:

$$y_{\mathrm{BB,RX}}[n] = y_{\mathrm{main}}[n] + y_{\mathrm{image}}[n] + Z_{\mathrm{RX}}[n] + Q[n] \tag{14}$$

where the main signal component is:

$$y_{\mathrm{main}}[n] = y_{\mathrm{LNA,RX}}[n] \cdot \exp(-\mathrm{j}(2\pi f_0 n + \Phi(n - \tau/T_{\mathrm{s}}))). \tag{15}$$

The image component introduced by receiver I/Q imbalance is:

$$y_{\mathrm{image}}[n] = \eta_{\mathrm{RX}} \cdot y^*_{\mathrm{LNA,RX}}[n] \cdot \exp(\mathrm{j}(2\pi f_0 n + \Phi(n - \tau/T_{\mathrm{s}}))). \tag{16}$$

In (15)–(16), f_0 is the LO carrier frequency (consistent with the transmitter), $\Phi(n)$ denotes the discrete-time phase noise process of the shared LO, τ is the time delay between the transmitter and receiver LO paths, and T_{s} is the system sampling period. η_{RX} is the receiver I/Q imbalance coefficient, which characterizes the amplitude and phase mismatch between the in-phase and quadrature branches. $Q[n]$ is the quantization noise introduced by the analog-to-digital converter (ADC). For sufficiently high ADC resolution, the quantization noise follows an approximately uniform distribution over $[-\Delta/2, \Delta/2]$, where Δ is the quantization step size.

3. Principle of the U-Net Variational Autoencoder with Multi-Feedback

Building on the full-link impairment model for IBFD systems presented in the preceding sections, we introduce the proposed U-Net Variational Autoencoder with Multi-Feedback (U-VAE-MF) framework. The proposed framework integrates two mutually complementary functional modules: transmitter-side nonlinear distortion mitigation and digital-domain SOI extraction under strong self-interference. Unlike conventional methods that address these two issues separately, the U-VAE-MF framework achieves joint optimization of nonlinear suppression and desired signal recovery in a unified end-to-end architecture.

3.1 Multi-Feedback Cancellation

Building on the analysis above, we present a digital-domain predistortion (DPD) scheme that suppresses the conjugate image components induced by transmitter I/Q imbalance. For accurate parameter estimation, the nonlinear output of the I/Q mixer is captured via a closed-loop feedback path and converted to the digital baseband. The observed feedback signal is modeled as:

$$y_{\mathrm{IQ,FB}}[n] = x[n] + \eta_{\mathrm{TX}} x^*[n] \tag{17}$$

where the phase noise effect is incorporated into the equivalent channel for brevity, consistent with the system model established earlier. We construct a measurement matrix from M observed samples to estimate the I/Q imbalance coefficient η_{TX}, given by:

$$\mathbf{X} = \begin{bmatrix} x[1] & x[2] & \dots & x[M] \\ x^*[1] & x^*[2] & \dots & x^*[M] \end{bmatrix} \tag{18}$$

The imbalance parameter vector $\mathbf{h}$ is estimated via the standard least-squares (LS) criterion:

$$\hat{\mathbf{h}} = \left(\mathbf{X}\mathbf{X}^{\mathrm{H}}\right)^{-1}\mathbf{X}\mathbf{y}_{\mathrm{IQ,FB}} \tag{19}$$

where $\mathbf{y}_{\mathrm{IQ,FB}}$ denotes the vectorized feedback signal, and $(\cdot)^{\mathrm{H}}$ denotes the Hermitian transpose. With the estimated coefficient $\hat{\eta}_{\mathrm{TX}} = \hat{h}_1$, the digitally pre-distorted baseband signal for the I/Q modulator is generated as:

$$x_{\mathrm{IQ,pre}}[n] = x[n] - \hat{\eta}_{\mathrm{TX}} x^*[n]. \tag{20}$$

This pre-compensation structure effectively cancels the conjugate image component before up-conversion.

We adopt the standard dual-tone test method, widely used to characterize intermodulation distortion in nonlinear systems, to validate the proposed PA nonlinearity calibration scheme. Figure 3(a) shows the spectra of the dual-tone input signal (red) and the uncalibrated PA output signal (orange). The input signal has a fundamental power of approximately -40 dBm, and the PA operates with the parameters $V_{\mathrm{sat}} = 20$ dBm, $s = 3$, and $g = 43$ dB as defined earlier.

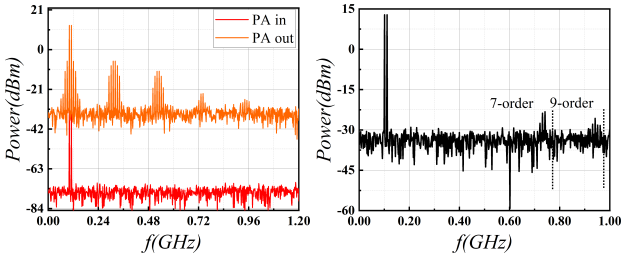


Fig. 3. Performance of nonlinear calibration for power amplifiers. (a) Input and output signal spectra of the power amplifier. (b) Output signal spectrum of the power amplifier after calibration.

The uncalibrated PA output contains significant high-order intermodulation components (3rd, 5th, 7th, and 9th order) induced by nonlinear distortion. Figure 3(b) presents the PA output spectrum after calibration, where the high-order intermodulation components are effectively suppressed to near the noise floor, demonstrating the strong nonlinear correction capability of the proposed method.

To further analyze the nonlinear distortion characteristics of the PA and construct the feature space for self-interference cancellation, we adopt a normalized complex baseband dual-tone excitation signal:

$$x_{\text{dual-tone}}[n] = \exp(\mathrm{j}2\pi f_1 n T_s) + \exp(\mathrm{j}2\pi f_2 n T_s) \tag{21}$$

where f_1 and f_2 are the frequencies of the two tones, and $\Delta f = f_2 - f_1$ is the frequency spacing between them. When passing through a nonlinear PA, the dual-tone signal generates a series of intermodulation distortion components at frequencies of the form $mf_1 + nf_2$, where the integer pair (m, n) specifies the intermodulation order. Each intermodulation component is then written in discrete time, with $[\cdot]$ denoting the discrete-time index, as:

$$x_{\text{fea},mn}[n] = x_1^m[n] \cdot x_2^n[n] \tag{22}$$

where $x_1[n] = \exp(\mathrm{j}2\pi f_1 n T_s)$ and $x_2[n] = \exp(\mathrm{j}2\pi f_2 n T_s)$. By collecting all significant intermodulation components as row vectors, we construct the feature matrix $\mathbf{X} \in \mathbb{C}^{K\times M}$, where K is the number of selected feature signals and M is the number of sampling points. The complete form of the feature matrix is (23)[at the top of the next page].

Figure 4 confirms the presence of these intermodulation components. Figure 4(a) shows the spectrum of the first-order fundamental signals and their adjacent intermodulation products, Figure 4(b) presents the third-order intermodulation components, and Figure 4(c) displays the fifth-order intermodulation components. All components are clearly distinguishable above the noise floor, confirming the validity of the feature signal model.

Figure 5 illustrates the architecture of the proposed mixed-domain nonlinear correction scheme for power amplifiers.

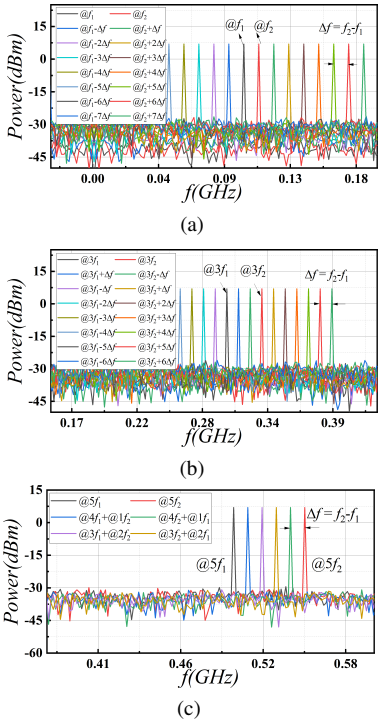


Fig. 4. Spectral representation of individual component signals in the feature matrix of (23) and performance after power amplifier calibration. (a) Spectrum of the first-order fundamental signal and its adjacent intermodulation products. (b) Spectrum of the third-order intermodulation components. (c) Spectrum of the fifth-order intermodulation components. All components lie clearly above the noise floor, confirming that the feature matrix captures the dominant nonlinear basis signals.

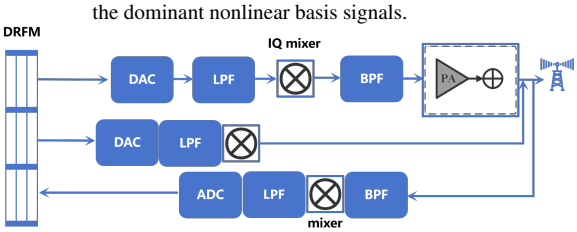


Fig. 5. Illustration of the working principle for nonlinear correction in power amplifiers. The system comprises three parallel paths implemented on a DRFM platform: the main transmit path (DAC → low-pass filter → up-conversion → PA), the analog cancellation path (digital nonlinear correction signal → DAC → injection at the antenna coupling point), and the feedback path (antenna-port sampling → down-conversion → ADC → parameter estimation).

The system consists of three parallel signal paths implemented on a digital radio frequency memory (DRFM) platform:

1. Main transmit path: The baseband signal is converted to analog via a DAC, low-pass filtered, up-converted to RF, and amplified by the PA before transmission.

2. Analog cancellation path: A nonlinear correction signal is generated digitally, converted to analog, and injected at the antenna coupling point to cancel the nonlinear distortion components of the PA output.

$$\mathbf{X} = \begin{bmatrix} x_{\text{fea},10}[1] & x_{\text{fea},10}[2] & \cdots & x_{\text{fea},10}[M-1] & x_{\text{fea},10}[M] \\ x_{\text{fea},01}[1] & x_{\text{fea},01}[2] & \cdots & x_{\text{fea},01}[M-1] & x_{\text{fea},01}[M] \\ x_{\text{fea},2(-1)}[1] & x_{\text{fea},2(-1)}[2] & \cdots & x_{\text{fea},2(-1)}[M-1] & x_{\text{fea},2(-1)}[M] \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{\text{fea},3(-2)}[1] & x_{\text{fea},3(-2)}[2] & \cdots & x_{\text{fea},3(-2)}[M-1] & x_{\text{fea},3(-2)}[M] \end{bmatrix} \tag{23}$$

3. Feedback path: The combined signal at the antenna port is sampled, down-converted to baseband, and digitized by an ADC for parameter estimation.

For algorithmic clarity, the effect of phase noise is isolated and evaluated separately in the performance analysis section. As analyzed above, the PA output contains a series of intermodulation distortion components, represented by the feature matrix $\mathbf{X}$ constructed in (23). The nonlinear correction coefficients are estimated via the standard LS criterion:

$$\hat{\mathbf{h}}_{\text{PA}} = \left(\mathbf{X}\mathbf{X}^{\text{H}}\right)^{-1}\mathbf{X}\mathbf{y}_{\text{FB}} \tag{24}$$

where $\mathbf{y}_{\text{FB}} \in \mathbb{C}^{M\times 1}$ is the vectorized feedback signal from the ADC, and $(\cdot)^{\text{H}}$ denotes the Hermitian transpose. Using the estimated coefficients, the digital nonlinear correction signal is generated as:

$$y_{\text{cor}}[n] = \sum_{m,n} \hat{h}_{\text{PA},mn} x_{\text{fea},mn}[n]. \tag{25}$$

This correction signal is converted to analog and injected into the analog cancellation path, effectively suppressing the nonlinear distortion components of the PA output. The proposed feature-based estimation method remains applicable to other modulated waveforms such as OFDM, as well as to LFM signals. The algorithm proceeds in two sequential stages:

1. Linear component estimation: The linear part of the PA response is estimated using a low-power training signal, where nonlinear distortion is negligible.

2. Nonlinear component extraction: The nonlinear distortion components are obtained by subtracting the estimated linear response from the feedback signal, and the correction coefficients are updated accordingly.

3.2 IBFD U-Net Variational Autoencoder Strategy

While conventional DRFM-based nonlinear correction effectively suppresses low-order intermodulation distortion, it struggles with high-order nonlinearities and time-varying channels in complex IBFD scenarios. Deep learning methods overcome these limitations by learning the nonlinear input–output mapping directly from data.

The encoder-decoder architecture decouples mixed SI and SOI signals through hierarchical feature learning. The encoder takes as its input a signal sequence of length N (the training sequence length) and compresses it into high-dimensional latent features, suppressing noise while retaining distinctive SI and SOI signatures. The decoder reconstructs a cleaner SOI estimate by upsampling latent features and fusing multi-scale information. Among encoder-decoder variants, the U-Net excels in this task due to skip connections that bridge low-level details with high-level semantics, as shown in Fig. 6.

Based on the standard U-Net framework, we design an improved three-stage encoder-decoder architecture for SIC, as illustrated in Fig. 7.

The U-Net consists of three components. The **down-sampling (encoder) path** applies two 3×3 convolutions with ReLU followed by max-pooling in each down block. The **up-sampling (decoder) path** uses transposed convolution for upsampling followed by two 3×3 convolutions with ReLU in each up block. The **skip connections** concatenate each down-block output with the corresponding up-block input, so that the fine-grained details lost during down-sampling are preserved, and a 1×1 convolution finalizes the reconstructed SOI.

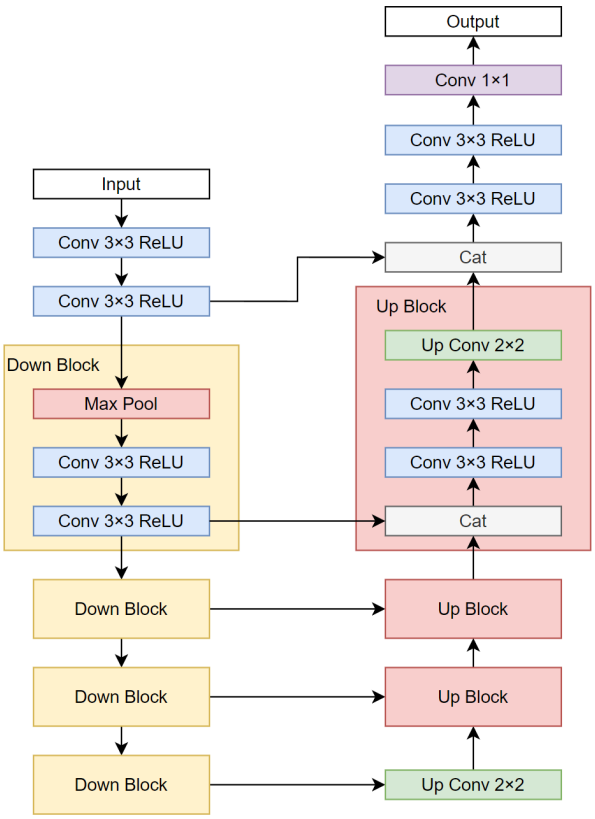


Fig. 6. The U-Net architecture used as the backbone of the proposed canceller: a symmetric encoder-decoder with 3×3 convolutions, ReLU activations, max-pooling down-sampling, transposed-convolution up-sampling, and skip connections that concatenate each encoder stage with the corresponding decoder stage.

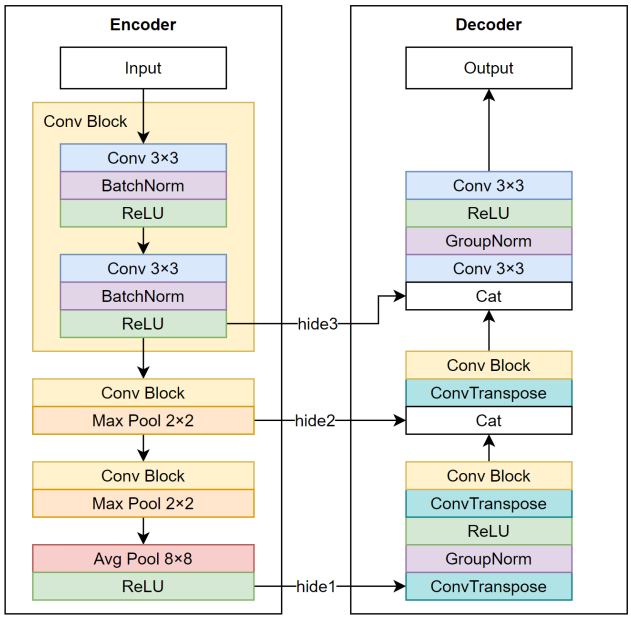


Fig. 7. The improved three-stage U-Net encoder-decoder architecture designed for self-interference cancellation, with channel dimensions $32 \rightarrow 64 \rightarrow 128$ in the encoder and $128 \rightarrow 64 \rightarrow 32$ in the decoder; intermediate feature maps $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ are retained as skip connections.

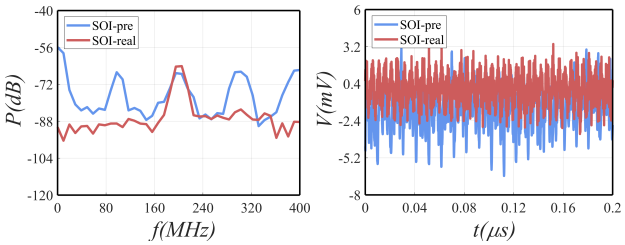


Fig. 8. Input-output performance without VAE structure. (a) Frequency-domain comparison of the predicted and actual signal of interest (SOI). (b) Time-domain comparison of the predicted and actual SOI.

The encoder comprises three cascaded convolutional blocks, each consisting of two 3×3 convolutional layers with batch normalization (BN) and ReLU activation. Max-pooling (2×2) is applied after the first two blocks for down-sampling; the output of the third block passes through global average pooling to form a compact bottleneck representation. At each stage, intermediate feature maps $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ (from shallow to deep) are retained as skip connections. Shallow features preserve fine-grained signal detail, while deep features encode high-level representations of nonlinear distortion, together enabling effective SIC while protecting the desired signal structure.

The decoder mirrors the encoder with a three-stage up-sampling structure. Each stage applies a 2×2 transposed convolution to restore the feature map resolution, followed by a convolutional block with group normalization (GN). Unlike BN, GN is independent of batch size and thus better suited for small-dataset training. The up-sampled maps are then concatenated with the corresponding encoder skip connections to integrate multi-scale features for signal reconstruction.

While the improved U-Net architecture can effectively extract multi-scale features, it suffers from a critical limitation inherent to standard encoder-decoder frameworks: the latent space lacks explicit regularization, resulting in a discontinuous and unstructured distribution. Consequently, when presented with unseen signal samples that deviate from the training data distribution, the decoder cannot generate accurate reconstructions. This phenomenon is clearly demonstrated in Fig. 8, which shows the reconstruction performance of the U-Net without the VAE module. In the frequency domain (Fig. 8(a)), the predicted SOI exhibits significant power deviation from the ground truth across all frequency bands, with a maximum error of approximately 15 dB at the main-signal frequency. In the time domain (Fig. 8(b)), the predicted signal shows severe amplitude distortion and waveform mismatch compared to the actual SOI. To address this generalization limitation, we integrate a VAE into the U-Net framework to regularize the latent space and improve the robustness of signal reconstruction.

Specifically, the encoder learns to map the input feature to two parameters of the Gaussian distribution: the mean μ and the log-variance λ. To enable end-to-end training via backpropagation, we employ the reparameterization trick to sample the latent vector $\mathbf{z}$ as:

$$\mathbf{z} = \mu + \epsilon \cdot \mathrm{e}^{\lambda/2}, \quad \epsilon \sim \mathcal{N}(0, 1) \tag{26}$$

where λ is the element-wise natural logarithm of the posterior variance, so that $\exp(\lambda)$ is the variance itself, and the latent vector $\mathbf{z}$ has dimension D_{latent}. The encoder and the decoder are optimized with Adam [44] on random mini-batches of 8 records; the corresponding settings are summarized in Tab. 3 and are discussed together with the database in Sec. 5.

The regularization effect of the VAE comes from the Kullback-Leibler (KL) divergence term in the loss function, which forces the learned latent distribution to approximate the standard Gaussian distribution. This results in a continuous and smooth latent space, significantly improving the model's robustness to unseen signal samples and time-varying nonlinear distortions.

The architecture of the proposed U-VAE network is illustrated in Fig. 9.

It retains the symmetric encoder-decoder structure and skip connections of the standard U-Net, with the VAE module embedded in the middle layer. Specifically:

1. The shallowest and deepest skip connections are directly transmitted from the encoder to the decoder to preserve fine-grained signal details and high-level semantic features.

2. The middle-layer feature map is fed into the VAE module, which splits it into the mean μ and log-variance λ via a chunk operation. The latent vector $\mathbf{z}$ is then sampled using the reparameterization trick and transmitted to the corresponding decoder layer.

This hybrid architecture combines the multi-scale feature extraction capability of U-Net with the latent-space regularization capability of the VAE, achieving both accurate signal reconstruction and strong generalization performance.

The core objective of the VAE is to maximize the evidence lower bound (ELBO) of the log-likelihood of the training data, expressed as:

$$\mathcal{L}_{\text{ELBO}} = \mathbb{E}_{q(\mathbf{z}|\mathbf{x})}[\log p(\mathbf{x}|\mathbf{z})] - D_{\text{KL}}(q(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z})) \quad (27)$$

where $q(\mathbf{z}|\mathbf{x})$ is the approximate posterior distribution learned by the encoder, $p(\mathbf{x}|\mathbf{z})$ is the likelihood of the decoder generating $\mathbf{x}$ given $\mathbf{z}$, and $D_{\text{KL}}(\cdot \parallel \cdot)$ denotes the Kullback-Leibler divergence. The first term corresponds to the reconstruction loss, and the second term is the regularization penalty that forces the approximate posterior to match the standard Gaussian prior $p(\mathbf{z}) \sim \mathcal{N}(0, 1)$. For the proposed U-VAE network, we design a two-stage training strategy as illustrated in Fig. 10.

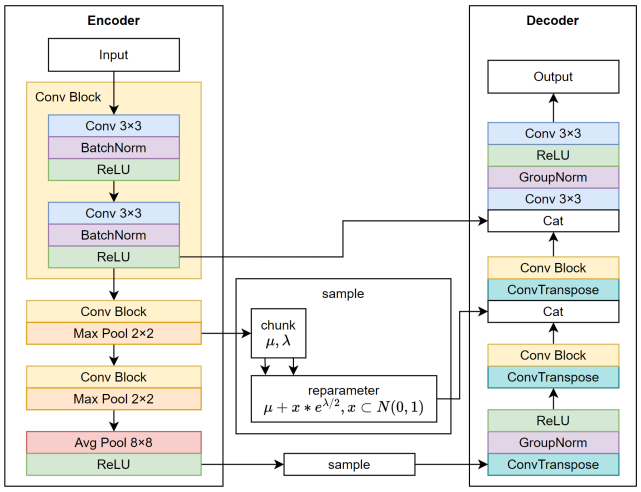


Fig. 9. The U-VAE encoder-decoder architecture: the VAE module is embedded at the bottleneck, splitting the middle-layer feature map into the mean μ and log-variance λ (chunk operation) and sampling the latent vector $\mathbf{z}$ with the reparameterization trick; the shallowest and deepest skip connections bypass the VAE to preserve fine-grained and high-level features.

The encoder loss is the KL divergence between the approximate posterior and the standard Gaussian prior:

$$\mathcal{L}_{\text{encoder}} = D_{\text{KL}}(q(\mathbf{z}|\mathbf{x}) \parallel \mathcal{N}(0, 1)). \quad (28)$$

For the diagonal Gaussian posterior and prior used here, this divergence admits the closed form

$$\mathcal{L}_{\text{encoder}} = \frac{1}{2} \sum_{i=1}^{n} \left(\mu_i^2 + \mathrm{e}^{\lambda_i} - \lambda_i - 1 \right) \quad (29)$$

where μ_i and λ_i are the mean and log-variance of the approximate posterior distribution, respectively. This regularization term prevents the latent space variance from collapsing to zero, a common issue in standard autoencoders. The decoder is conditioned on the difference between the latent code of the received signal and the latent code of the self-interference signal, and its goal is to reconstruct the clean SOI. The decoder loss is defined as the mean square error (MSE) between the reconstructed SOI and the ground-truth SOI, the latter being the interference-free SOI used to synthesize the received signal:

$$\mathcal{L}_{\text{decoder}} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \quad (30)$$

where $\hat{y}_i$ is the decoder's output and y_i is the ground-truth SOI. The total loss function for end-to-end training is the weighted sum of the encoder loss and the decoder loss:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{decoder}} + \gamma \cdot \mathcal{L}_{\text{encoder}} \quad (31)$$

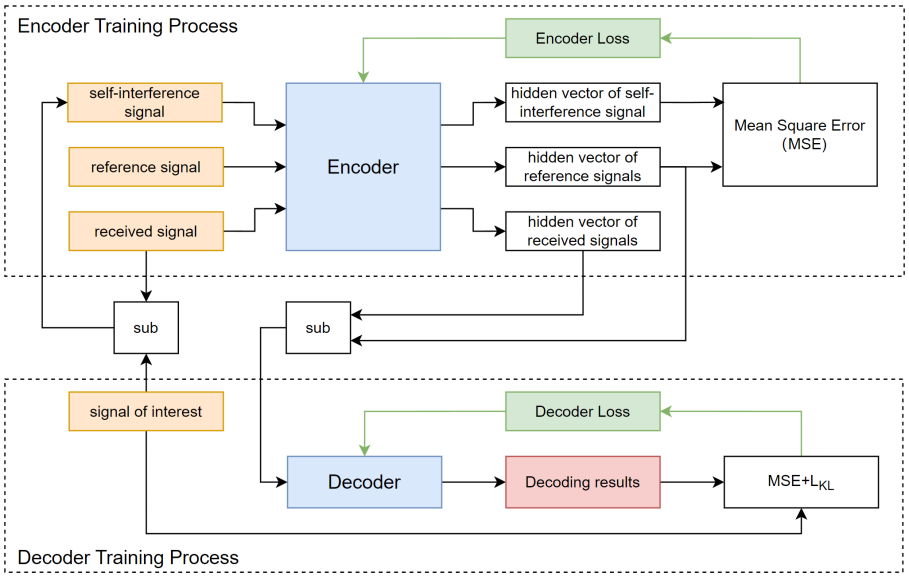


Fig. 10. Two-stage training workflow of the proposed U-VAE network. Stage 1 trains the encoder with the KL-divergence loss to regularize the latent space toward $\mathcal{N}(0, 1)$; Stage 2 trains the decoder with the mean square error between the reconstructed and ground-truth SOI; the two stages are then fine-tuned end-to-end with the combined loss of (31).

where γ is a weighting coefficient that balances reconstruction accuracy and latent space regularization.

Figure 11 demonstrates the SOI extraction performance of the proposed method. The 120 MHz signal bandwidth in this figure differs from the nominal values in Tab. 2, because the dataset was deliberately expanded with additional bandwidth configurations in order to generate more training data and to increase the diversity of the training set. Figure 11 therefore corresponds to one such expanded-bandwidth case rather than to the nominal setting of Tab. 2. In the frequency domain (Fig. 11(a)), the self-interference across the 120 MHz signal bandwidth is effectively suppressed by more than 40 dB, and the predicted SOI spectrum matches the ground truth closely. In the time domain (Fig. 11(b)), the reconstructed waveform agrees closely with the actual SOI, with negligible amplitude and phase distortion.

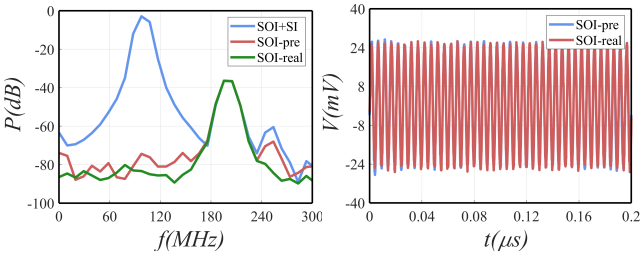
3.3 Generalization Performance

To further quantify generalization, we evaluate the MSE between the predicted SOI and the ground-truth SOI on the 960-record test set, which is fully disjoint from the training and validation sets. As shown in Fig. 12(a), the MSE remains consistently low across all samples, with a maximum value of 1.8×10^{-5}. The histogram in Fig. 12(b) shows that the MSE values follow a right-skewed distribution, with over 80% of the samples achieving an MSE below 4×10^{-6}. Because the test set is independent of training, these statistics also quantify the spread of the reconstruction accuracy rather than a single operating point.

To report the dispersion of the cancellation performance explicitly, we compute the mean cancellation ratio over the same 960 test records together with the 95% confidence interval of the mean: the mean cancellation ratio is 52 dB and the confidence interval is ±3 dB. The residual gap to the theoretical limit of (42) is therefore 2 dB on average, which confirms that the proposed canceller operates close to the bound derived in Sec. 4 while retaining a narrow spread across the independent test records.

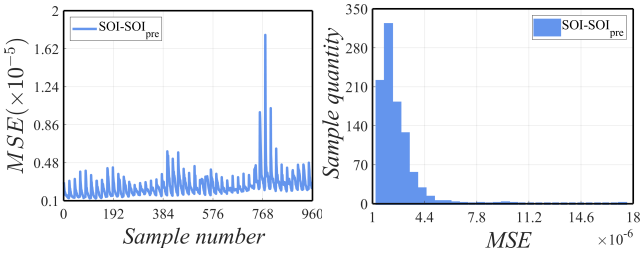
Figure 13(a) presents the self-interference cancellation performance across the 960-record test set. Note that the power levels in Fig. 13(b) differ from those in Tab. 2. This is intentional: the dataset was expanded to include different SI and SOI power ranges, so as to evaluate the robustness of the proposed method under varying power conditions. The 960-record test set therefore covers multiple SI/SOI power combinations, and the reported average cancellation ratio of approximately 52 dB is obtained by averaging over these conditions. The values shown in Fig. 13(b) correspond to one representative sample (SI ∼10 dBm, residual ∼−30 dBm, SOI ∼−32 dBm) rather than to the fixed nominal setting in Tab. 2. They indicate that, for this sample, the residual is dominated by the SOI and that the SI has been suppressed close to the noise floor, thereby confirming the adaptability of the method to different SI and SOI power levels.

Figure 14 further demonstrates the SOI extraction performance using a representative test sample. In the frequency domain (Fig. 14(a)), the self-interference across the 120 MHz signal bandwidth is strongly suppressed, and the predicted SOI spectrum matches the ground truth closely. In the time domain (Fig. 14(b)), the reconstructed waveform agrees closely with the actual SOI, and the residual amplitude and phase error is negligible. A minor power difference of less than 1 dB is observed between the predicted and ground-truth SOI, which is within the acceptable range for most communication applications.



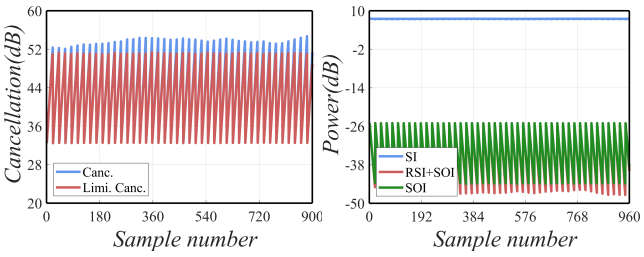
(a) (b)

Fig. 11. SOI extraction performance of the proposed U-VAE method. (a) Frequency-domain comparison of the mixed signal (SOI+SI), predicted SOI, and ground-truth SOI. (b) Time-domain comparison of the predicted SOI and ground-truth SOI.



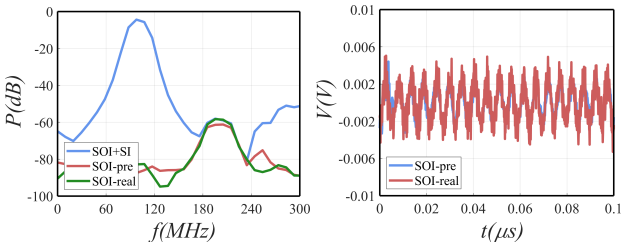
(a) (b)

Fig. 12. Statistical analysis of the cancellation performance over the independent test set. (a) MSE values for individual test samples. (b) Histogram of the MSE distribution across the 960 test records.



(a) (b)

Fig. 13. Statistical cancellation performance of the proposed U-VAE-MF over the 960 test records. (a) Cancellation ratio versus sample index, where "Canc." denotes the actual cancellation performance of the proposed method and "Limi. Canc." denotes the theoretical cancellation limit. (b) Power levels of the original self-interference (SI), of the residual self-interference after cancellation, and of the ground-truth signal of interest (SOI).



(a) (b)

Fig. 14. Single-sample SOI extraction performance. (a) Frequency-domain comparison of the mixed signal (SOI+SI), predicted SOI, and ground-truth SOI. (b) Time-domain comparison of the predicted SOI and ground-truth SOI.

4. Error Analysis and Theoretical Performance Limit

This section presents a comprehensive error analysis of the proposed nonlinear correction and self-interference cancellation scheme, and derives the theoretical performance limit imposed by system impairments.

4.1 I/Q Imbalance Correction Error

The output signal after digital predistortion and upconversion is:

$$y_{\mathrm{IQ,out}}[n] = \left(x_{\mathrm{IQ,pre}}[n] + \eta_{\mathrm{TX}} x^{*}_{\mathrm{IQ,pre}}[n]\right) \exp(\mathrm{j}\Phi(n)) \quad (32)$$

where $x_{\mathrm{IQ,pre}}[n]$ is the pre-distorted baseband signal, η_{TX} is the transmitter I/Q imbalance coefficient, and $\Phi(n)$ is the LO phase noise process. Substituting the pre-distortion signal $x_{\mathrm{IQ,pre}}[n] = x[n] - \hat{\eta}_{\mathrm{TX}} x^{*}[n]$, where $\hat{\eta}_{\mathrm{TX}}$ is the estimated imbalance coefficient, into (32) yields:

$$y_{\mathrm{IQ,out}}[n] = \left((1 - \eta_{\mathrm{TX}}\hat{\eta}_{\mathrm{TX}})\, x[n] + \Delta\eta_{\mathrm{TX}} x^{*}[n]\right) \exp(\mathrm{j}\Phi(n)) \quad (33)$$

where $\Delta\eta_{\mathrm{TX}} = \eta_{\mathrm{TX}} - \hat{\eta}_{\mathrm{TX}}$ is the estimation error of the I/Q imbalance coefficient. The second term in (33) represents the residual conjugate image component caused by estimation error. The average power of the residual image component is:

$$\begin{aligned} P_{\mathrm{image}} &= \mathbb{E}\left[\left|\Delta\eta_{\mathrm{TX}} x^{*}[n] \exp\left(\mathrm{j}\Phi(n)\right)\right|^2\right] \\ &= \left|\Delta\eta_{\mathrm{TX}}\right|^2 \mathbb{E}\left[\left|x[n]\right|^2\right] \end{aligned} \quad (34)$$

where we use the fact that $\mathbb{E}\left[\left|\exp\left(\mathrm{j}\Phi(n)\right)\right|^2\right] = 1$ for any phase noise process.

4.2 PA Nonlinear Correction Error

Similarly, the residual nonlinear distortion after PA correction is:

$$y_{\mathrm{PA,err}}[n] = \sum_{m,n} \Delta h_{mn} x_{\mathrm{fea},mn}[n] + y_{\mathrm{PA,other}}[n] \quad (35)$$

where $\Delta h_{mn} = h_{mn} - \hat{h}_{mn}$ is the estimation error of the mn-th nonlinear coefficient, $x_{\mathrm{fea},mn}[n]$ is the corresponding feature signal, and $y_{\mathrm{PA,other}}[n]$ represents the residual distortion from higher-order intermodulation components not included in the feature matrix. The average power of the residual nonlinear distortion is:

$$P_{\mathrm{PA,err}} = \sum_{m,n} \mathbb{E}\left[\left|\Delta h_{mn}\right|^2\right] \mathbb{E}\left[\left|x_{\mathrm{fea},mn}[n]\right|^2\right] + P_{\mathrm{other}} \quad (36)$$

where $P_{\mathrm{other}} = \mathbb{E}\left[\left|y_{\mathrm{PA,other}}[n]\right|^2\right]$ is the power of the unmodeled higher-order distortion. Equations (34) and (36) show that the total correction error is determined by three key factors:

1. The estimation accuracy of the I/Q imbalance and nonlinear coefficients.

2. The completeness of the feature matrix, which determines the magnitude of the unmodeled higher-order distortion.

3. The residual phase-noise floor, which is independent of the estimation accuracy and cannot be removed by any coefficient adaptation (Equation (42)).

4.3 Impact of Estimation Errors, Model Mismatch, and Channel Dynamics

The analysis above assumes that the impairment coefficients are estimated perfectly and that the channel is static during the estimation interval. We now relax these assumptions and discuss their influence on the achievable cancellation.

Channel estimation errors. Let $\hat{h}[l] = h[l] + e[l]$ denote the estimated impulse response, where $e[l]$ is a zero-mean estimation error with variance σ_{e}^2 (for the LS estimator, σ_{e}^2 scales inversely with the observation length M). Reconstructing the SI replica with $\hat{h}[l]$ leaves a residual power

$$P_{\mathrm{RSI,est}} = \sum_{l=0}^{L} \sigma_{\mathrm{e}}^2\, \mathbb{E}\left[\left|x_{\mathrm{PA}}[n-l]\right|^2\right] \quad (37)$$

which adds directly to the phase-noise floor in (42). In the proposed scheme this term is reduced in two ways: the multi-feedback stage confines the reconstruction to a small residual subspace (fewer coefficients to estimate), and the U-VAE learns to suppress the remaining structured error from data.

Model mismatch. If the true PA characteristic deviates from the Rapp model used for calibration (e.g., additional AM-PM conversion or stronger memory), the unmodeled part appears as extra energy in P_{other} and degrades the cancellation performance. The U-VAE mitigates this mismatch because it does not assume a parametric PA model: it learns the residual mapping directly from the mixture signal. This robustness to PA-model mismatch follows directly from the non-parametric, data-driven nature of the learned canceller rather than from the specific calibration model.

Channel dynamics. Under a time-varying channel with maximum Doppler shift f_D, the channel decorrelates over a coherence time $T_c \approx 1/(2f_D)$. The multi-feedback coefficients and the U-VAE weights are only valid for a fraction of T_c; re-estimation or online fine-tuning is therefore required at a rate proportional to f_D. The low estimation latency reported in Sec. 5 (100 µs) indicates that the coefficients can be refreshed at a rate well above typical fading dynamics in pedestrian and low-mobility scenarios, while high-Doppler scenarios remain an open problem. As long as the coefficient refresh period T_{up} is kept small relative to the coherence time T_c, the additional residual due to channel aging remains below the phase-noise floor of (42), which therefore remains the dominant limitation.

4.4 Theoretical Cancellation Limit

The theoretical performance limit of self-interference cancellation is determined by the minimum achievable residual self-interference power when the estimation errors are zero and the channel is static over the estimation interval. Both assumptions are relaxed and quantified in the preceding subsection, so the limit derived here corresponds to the best case. The estimation errors, model mismatch, and channel dynamics analyzed there cause the performance to degrade from this bound. The received signal at the receiver is:

$$y[n] = y_{SOI}[n] + y_{RSI}[n] + Z_{total}[n] \tag{38}$$

where $y_{SOI}[n]$ is the signal of interest, $y_{RSI}[n]$ is the residual self-interference, and $Z_{total}[n]$ is the total additive noise including receiver thermal noise and quantization noise. When the estimation errors are zero, the residual self-interference is caused solely by LO phase noise:

$$y_{RSI,PN}[n] = \sum_{l=0}^{L} h[l]P(l)x_{PA}[n-l] \tag{39}$$

where the residual phase factor $P(l)$ is the difference between the phase-noise rotations of the transmit-side and receive-side LO paths:

$$P(l) = \exp\big(j\Phi((n+l_1)T_s)\big) - \exp\big(j\Phi((n+l_2)T_s - \tau)\big) \tag{40}$$

with $l = l_1 - l_2$. The underlying phase noise difference is the zero-mean Gaussian random variable

$$\Delta\Phi(l) = \Phi((n+l_1)T_s) - \Phi((n+l_2)T_s - \tau) \tag{41}$$

whose variance is $\sigma^2 = 2\pi\beta|lT_s + \tau|$ (see Appendix A). In (39)–(41), $h[l]$ is the channel impulse response, T_s is the sampling period, τ is the time delay between the transmitter and receiver LO paths, and β is the single-sided 3-dB bandwidth of the LO phase noise spectrum defined in Sec. 2. The average power of the phase-noise-induced residual self-interference is:

$$P_{RSI,PN} = \sum_{l=0}^{L} |h[l]|^2 \, \mathbb{E}_2(l) \, \mathbb{E}\left[|x_{PA}[n-l]|^2\right] \tag{42}$$

where $\mathbb{E}_2(l) = \mathbb{E}\left[|P(l)|^2\right] = 2 - 2\exp(-\pi\beta|lT_s + \tau|)$ is the average power of the residual phase factor, derived in Appendix A. Equation (42) gives the fundamental theoretical limit of self-interference cancellation, which cannot be exceeded by any linear or nonlinear cancellation algorithm.

5. Analysis of Simulation Results

We validate the proposed U-VAE-based self-interference cancellation scheme on a full-link IBFD simulation platform that incorporates all system impairments analyzed in Sec. 2, including PA nonlinearity, I/Q imbalance, LO phase noise, and receiver thermal noise. The LFM waveform serves as the transmitted signal owing to its widespread use in radar systems; the OFDM waveform is additionally considered in the comparative performance analysis presented later in this section. Table 2 summarizes the key simulation parameters, which are consistent with the system model established earlier. Unless otherwise stated, all simulation results presented in this paper are obtained under these parameters.

The database used in this paper is generated entirely by the simulation platform described above, so that the provenance of every record is determined by the models of Sec. 2 rather than by an external measurement. Each record corresponds to one realization of the complete transmit–receive chain. An LFM waveform is generated, distorted by the transmitter I/Q imbalance and by the Rapp power-amplifier nonlinearity (Sec. 2.1), propagated through the frequency-selective self-interference channel, and then processed by the receiver chain (Sec. 2.2). The receiver chain adds the shared-LO phase noise, the receiver I/Q imbalance, the additive noise, and the ADC quantization noise. The impairment parameters are drawn within the nominal operating ranges of Tab. 2, so that each record represents a different but physically admissible combination of RF non-idealities, and the canceller must therefore generalize over the resulting spread of the impairment space. Every record contains three synchronized waveforms: the received mixture, the self-interference reference signal, and the ground-truth signal of interest, the latter obtained by removing the self-interference component from the mixture.

Parameters	Value
Transmitter signal-to-noise ratio (SNR)	60 dB
Rician factor (K)	30 dB
PA gain	43 dB
PA saturation power	20 dBm
Rapp model smoothness factor (s)	3
LO phase noise 3-dB bandwidth	50 Hz
Baseband self-interference (SI) power	0 dBm
Signal of interest (SOI) power	−60 dBm
LO carrier frequency	2.4 GHz
SI baseband carrier frequency	100 MHz
SI bandwidth	300 MHz
SOI baseband carrier frequency	120 MHz
SOI bandwidth	330 MHz

Tab. 2. Key simulation parameters.

Setting	Value
Optimizer	Adam
Learning rate	10^{-3} (constant)
Adam β_1, β_2	0.9, 0.999
Adam ϵ	10^{-7}
Batch size	8
Max Epochs	1000 (early stopping)
Framework and hardware	PyTorch + RTX 4090

Tab. 3. Training configuration of the U-VAE network.

The database is partitioned into a training portion and an independent test portion, comprising approximately 70% and 30% of the records, respectively. The test portion contains the 960 waveform records used for all statistical results reported in this paper, and a subset of the training portion is reserved for validation and used only for model selection. No test record takes part in training or in any parameter or hyper-parameter decision, so that the test set is fully disjoint from the data seen by the network. Because the generator reproduces the coupled impairment mechanisms of Sec. 2 rather than an idealized channel, it also provides a controlled setting in which model mismatch can be examined later in this section.

The U-VAE network is trained on the training portion of this database with the configuration summarized in Tab. 3. The weights are optimized with Adam, whose first and second moment estimates are updated with the default exponential decay rates of 0.9 and 0.999 and a numerical stability constant of 10^{-7}. A constant learning rate of 10^{-3} is used on random mini-batches of 8 records drawn without replacement. One epoch corresponds to one complete pass over the training portion, and training is run for a maximum of 1000 epochs with early stopping when the validation loss stops improving. All experiments reported in this paper share this single training configuration, so that no reported result depends on a per-figure hyper-parameter choice.

The metrics used in this section are defined as follows. The cancellation ratio of a test record is the ratio, expressed in decibels, between the self-interference power at the receiver input and the residual self-interference power after cancellation, the latter being the power of the received mixture once the estimated self-interference component has been removed. The isolation curves of Fig. 20 follow the same definition as a function of the SI baseband power. The mean square error reported in Fig. 12 is computed between the predicted and the ground-truth signal of interest of each test record. The computational latency is the average channel-estimation and cancellation time per record, measured on a common hardware platform and averaged over 1 000 independent runs. The bit error rate is obtained for the communication link. The reconstructed signal of interest is OFDM-demodulated with 64 subcarriers and a 16-sample cyclic prefix (80 samples per OFDM symbol) using QPSK modulation. For each record, 10 000 OFDM symbols are evaluated (128 information bits per symbol), and the decided bits are compared with the transmitted bits. Isolation and cancellation ratios quantify the residual self-interference power for both the LFM (sensing) and the OFDM (communication) excitation, whereas the bit error rate is meaningful only for the communication link and is therefore reported for the OFDM waveform alone. Unless a figure states otherwise, all curves are obtained with the parameter set of Tab. 2, with only the input power and the signal bandwidth varied as indicated in each figure. The same parameter set is used for every comparison, so that the differences between methods are attributable to the cancellation algorithm rather than to a change of operating conditions.

5.1 Nonlinear Self-Interference Components

To verify the effectiveness of I/Q imbalance correction and identify the dominant nonlinear distortion sources, we analyze the power distribution of each SI component as a function of the input power, shown in Fig. 15. The "other" category encompasses unmodeled higher-order intermodulation components (≥7th order), channel memory effects, and receiver thermal noise.

Figure 15(a) shows the component power distribution without I/Q imbalance correction. The conjugate image of the fundamental SI (Image SI) emerges as the dominant distortion component when input power exceeds −56 dBm, surpassing the SOI power. As input power increases, the third-order intermodulation component (HD3) and its conjugate image (Image HD3) grow rapidly. Above −36 dBm, Image HD3 overtakes all other nonlinear components and becomes the dominant nonlinear distortion source. For input powers below −20 dBm, the power difference between the fundamental SI and the "other" components remains constant at 60 dB, indicating that thermal noise dominates the "other" category in this regime. Beyond −20 dBm, higher-order nonlinear components become significant, causing the power of the "other" components to rise with input power.

Figure 15(b) presents the component power distribution after I/Q imbalance correction. The conjugate image components (Image SI, Image HD3, Image HD5) are suppressed by more than 30 dB across the entire power range. Consequently, HD3 becomes the dominant nonlinear distortion source, and the "other" components no longer exceed the SOI power until the input power surpasses −45 dBm, which relaxes the input-power constraint of the system by 11 dB with respect to the uncorrected case.

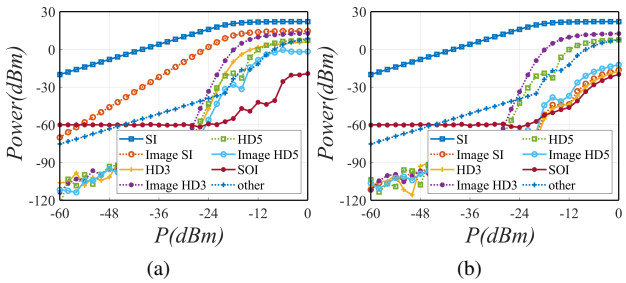


Fig. 15. Power distribution of the self-interference components versus input power, for the parameter set of Tab. 2. (a) Without I/Q imbalance correction. (b) With I/Q imbalance correction.

Notably, when the input power exceeds −20 dBm, several high-order intermodulation components overlap with the SOI in the frequency domain, introducing additional interference into the extracted SOI and causing a slight increase in the measured SOI power.

5.2 Performance Analysis

We evaluate the robustness of the proposed method by testing its cancellation performance over a range of input power levels (−60 to 0 dBm) and signal bandwidths (132 to 297 MHz), as shown in Fig. 16.

Figure 16(a) presents the cancellation ratio under various conditions. The proposed method achieves a maximum cancellation ratio of 85 dB at low input power (−60 dBm) and narrow bandwidth (132 MHz). As the input power increases, the cancellation ratio gradually declines. This degradation is mainly caused by the increasingly strong PA nonlinearity at high input power: the adopted model deviates more significantly from the actual PA behavior, leading to incomplete cancellation of the residual nonlinear components and thus a reduction in the effective isolation. At 0 dBm input power, the PA operates in deep saturation (saturation power = 20 dBm), and the cancellation ratio drops to approximately 60 dB.

Figure 16(b) shows the corresponding residual self-interference (RSI) power. At 0 dBm input power and 297 MHz bandwidth, the RSI power is approximately −40 dBm, which is consistent with the 60 dB cancellation ratio reported in Fig. 16(a), because the transmitted power is 20 dBm at this operating point. This result corresponds to an ablation experiment without the multi-feedback linearization module. In this case, when the PA enters the high-saturation region, the isolation degrades severely, leading to the observed increase in RSI power. In contrast, with the complete proposed method, the RSI power remains below the receiver noise floor even under the most demanding conditions (high power and wide bandwidth), ensuring reliable detection of the signal of interest.

These results confirm that the proposed method maintains stable cancellation performance across the tested range of operating conditions, which indicates its suitability for practical IBFD systems with varying signal bandwidths and power levels.

Figure 17 compares the cancellation performance of the proposed U-VAE-MF with the theoretical limit derived in Sec. 4. This limit, determined by LO phase noise and receiver thermal noise, represents the physical bound that no cancellation algorithm can exceed.

As shown in Fig. 17(a), when the input power is fixed at 0 dBm (PA saturation point), the cancellation ratio remains stable across the entire bandwidth range from 100 to 500 MHz. The proposed method achieves an average cancellation ratio of 81 dB, only 2 dB below the theoretical limit, demonstrating its ability to suppress frequency-dependent nonlinear distortion over a wide bandwidth.

Figure 17(b) depicts the cancellation performance under varying input power with the signal bandwidth fixed at 200 MHz. The cancellation ratio increases linearly with input power in the low-power region and saturates above −20 dBm. At this saturation point, the residual self-interference power is dominated by system noise rather than nonlinear distortion. The proposed method maintains a cancellation ratio of 81 dB for all input powers above −20 dBm.

These results indicate that the proposed U-VAE-MF approaches the physical limit of self-interference cancellation under the reported operating conditions.

Robustness beyond 0 dBm: Real IBFD operation requires stable self-interference cancellation when the transmit power exceeds 0 dBm and the power amplifier is driven further into the nonlinear region. The reference operating point is 0 dBm, at which the PA already operates in the nonlinear region, so the reported cancellation performance already reflects the proposed method's ability to fit a strongly nonlinear system. Because 0 dBm is the upper end of the simulated power sweep, behavior beyond it is inferred from the error model of Sec. 4 rather than read from the reported curves. That model predicts no abrupt failure under deeper

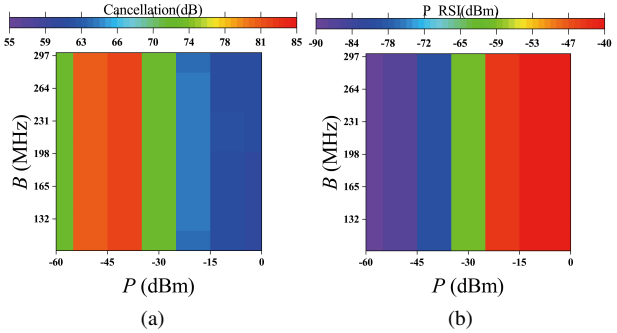


Fig. 16. Self-interference cancellation performance across different bandwidths and input power levels. Input power ranges from −60 to 0 dBm and the signal bandwidth from 132 to 297 MHz. (a) Cancellation ratio versus input power and signal bandwidth. (b) Residual self-interference (RSI) power versus input power and signal bandwidth.

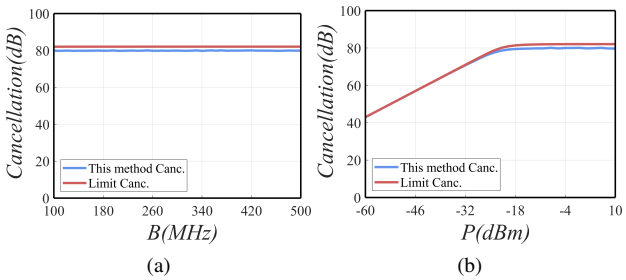


Fig. 17. Cancellation performance of the proposed U-VAE-MF versus the theoretical limit. (a) Cancellation ratio versus signal bandwidth over 100–500 MHz (input power fixed at 0 dBm). (b) Cancellation ratio versus input power (signal bandwidth fixed at 200 MHz).

compression, since the limiting mechanisms degrade smoothly rather than abruptly. As the input power increases further, two structural effects dominate: (i) 7th- and higher-order intermodulation grows faster than the calibrated components and increasingly contributes to P_{other} in (36); and (ii) LFM peak-to-average excursions reach the PA's deep-saturation knee, where the Rapp model becomes more sensitive to parameter mismatch. Both effects are structural rather than algorithmic: they grow monotonically with input power and can be counteracted by enlarging the feature matrix with higher-order terms or by backing off the PA. The 0 dBm reference corresponds to the maximum practical transmit power of the considered short-range IBFD front-end, and the gradual degradation predicted above it indicates that the framework does not rely on a narrow linear PA regime.

5.3 Comparative Performance Analysis

5.3.1 Comparison with Traditional Methods

We compare the proposed multi-feedback nonlinear correction scheme with two widely used linear cancellation methods: the LS estimator and the least mean squares (LMS) estimator. These two estimators are chosen as baselines because they are the standard digital-domain self-interference cancellers in IBFD transceivers and because the proposed multi-feedback structure is designed to replace exactly this class of estimators; comparing against them therefore isolates the effect of the nonlinear correction and of the transmitter-side pre-compensation. All methods are evaluated under the same system parameters defined in Tab. 2. Unless stated otherwise, the results in this subsection isolate the multi-feedback stage, namely the nonlinear-augmented LS (LS-NL) and LMS (LMS-NL) estimators, without the U-VAE canceller; the performance of the complete U-VAE-MF canceller is reported in Figs. 17 and 13.

Figure 18(a) shows the residual self-interference power versus input power. For LS-based methods, the traditional LS method suffers from severe performance degradation as input power increases: when the input power exceeds −24 dBm, its residual power rises sharply from −62 to −18 dBm at 0 dBm input, owing to its inability to suppress nonlinear distortion. In contrast, the proposed LS-based method maintains strong cancellation performance across the entire power range, keeping the residual power as low as −56 dBm even at 0 dBm input, which is a 38 dB improvement over the traditional LS method; the minor degradation at high power stems primarily from unmodeled 7th-order and higher nonlinearities.

For LMS-based methods, the traditional LMS method outperforms the traditional LS method at high power, stabilizing the residual power at approximately −30 dBm when the input power is −12 dBm. The proposed LMS-based method further improves performance, achieving a stable residual power of −33 dBm across all input power levels. Figure 18(b) presents the cancellation performance under varying Rician factor K, with the input power fixed at 0 dBm. As K increases, the strength of the direct-path ($l = 0$) self-interference signal increases, and the proposed method consistently outperforms the traditional methods across all K values. When K exceeds 24 dB, the proposed LS-based method reaches its optimal performance with a residual power of −56 dBm, whereas the proposed LMS-based method achieves a residual power of −33 dBm, which is 10 dB better than the traditional LMS method (−22 dBm, at 0 dBm power) under the same conditions.

These results confirm that the proposed multi-feedback scheme effectively suppresses both linear and nonlinear self-interference components, substantially outperforming traditional linear cancellation methods across a range of operating conditions. We evaluate the real-time processing cost of the proposed multi-feedback scheme against that of the traditional methods using two metrics: the channel estimation time and the self-interference cancellation time. All tests are conducted on the same hardware platform equipped with a 2.4 GHz CPU and 8 GB RAM, and the reported times are averaged over 1,000 independent runs.

Figure 19(a) presents the complexity comparison for LS-based methods. The traditional LS method requires an average of 270 µs for channel estimation and 80 µs for signal cancellation. By contrast, the proposed LS-based method reduces the channel estimation time to 100 µs (a 63% reduction) and the cancellation time to 50 µs (a 37.5% reduction). This substantial speedup arises from the multi-feedback structure, which avoids the repeated matrix inversions required by traditional LS methods.

Figure 19(b) shows the complexity comparison for LMS-based methods. Owing to the iterative nature of the LMS algorithm, the total execution time depends on the number of iterations required for convergence. The traditional LMS method requires an average of 80 µs for channel estimation and 0.01 µs for cancellation, whereas the proposed LMS-based method reduces the channel estimation time to 25 µs (a 68.75% reduction) while maintaining a similar cancellation time of 0.01 µs. The real-time advantage of the proposed method is particularly pronounced for iterative algorithms, as it significantly reduces the number of iterations required for convergence.

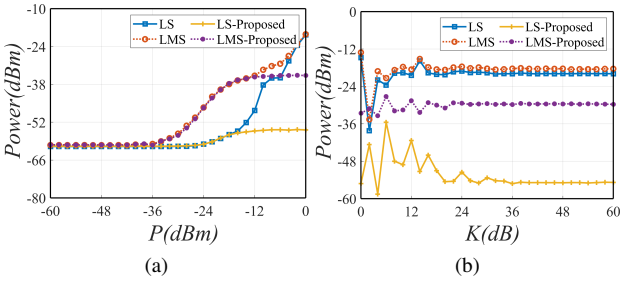
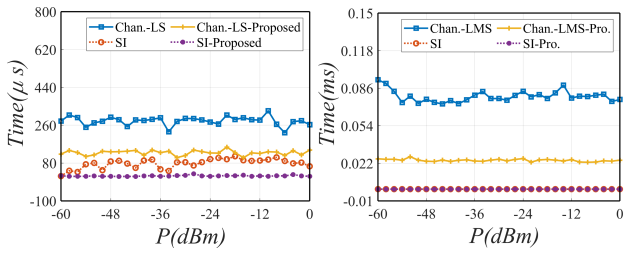
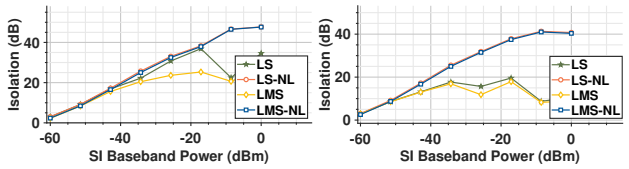


Fig. 18. Residual self-interference power after cancellation: comparison between traditional methods and the proposed multi-feedback nonlinear correction scheme. (a) Performance under varying input power. (b) Performance under varying Rician factor K, with the input power fixed at 0 dBm.



(a) (b)

Fig. 19. Computational complexity comparison: channel reconstruction time and cancellation time. (a) LS-based methods. (b) LMS-based methods.



(a) (b)

Fig. 20. Comparison of isolation performance. (a) Isolation versus SI baseband power for LFM input. (b) Isolation versus SI baseband power for OFDM input.

These results demonstrate that the proposed multi-feedback stage achieves strong cancellation within the latency budget of practical IBFD systems (100 µs for channel estimation and 50 µs for cancellation, as reported in Fig. 19). Although direct comparison with existing neural-network-based SIC methods [12] is difficult because the system assumptions differ (e.g., single-impairment models without coupled nonlinearities), a systematic qualitative comparison with recent learning-based SIC methods is provided in the next subsection. In addition, the LS-NL variants of the proposed multi-feedback scheme remain within 2 dB of the theoretical limit derived in Sec. 4 under the conditions of Fig. 17.

Figure 20 compares the isolation achieved for OFDM and LFM waveforms as a function of the SI baseband power. At low SI power, linear and nonlinear cancellers achieve comparable isolation. As the SI power increases, the conventional LS and LMS methods exhibit a non-monotonic rise-then-fall trend caused by RF nonlinear distortion, whereas the proposed nonlinear-augmented LS (LS-NL) and LMS (LMS-NL) methods improve monotonically and exceed a peak isolation of 40 dB, which makes them better suited to high-power full-duplex communication and radar.

Figure 21 presents the bit error rate (BER) performance of OFDM systems across varying SI baseband power levels. At low SI power, all evaluated algorithms maintain a low BER. As SI power increases, the BER of conventional linear LS and LMS algorithms degrades sharply at a relatively low power threshold. By contrast, the proposed LS-NL and LMS-NL nonlinear approaches preserve low-BER operation over a substantially wider high-power range, effectively enhancing communication reliability under strong self-interference.

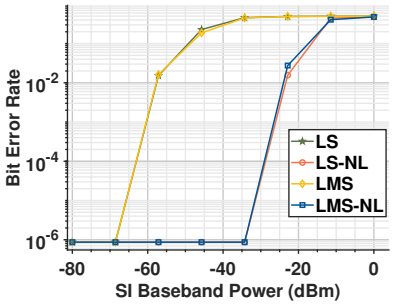


Fig. 21. BER of OFDM signals as a function of SI baseband power.

Three quantities that are commonly quoted for full-duplex links are not presented here as separate curves, and the reason is stated explicitly. The error vector magnitude (EVM) is the same receiver-side distortion that the BER curve of Fig. 21 resolves, expressed per symbol rather than per bit; for the QPSK-OFDM link considered here the two quantities are monotonically related, so the BER is reported in its place. Throughput and spectral efficiency are bounded by the same measurements: because the residual self-interference is suppressed below the receiver noise floor (Fig. 16(b)), the full-duplex link is limited by thermal noise rather than by the canceller, so those two quantities follow from the reported BER curve instead of providing independent evidence about the algorithm. The FLOP count, finally, is fixed by the network architecture and by the length N of the input sequence (Sec. 3.2) and is therefore platform-independent, whereas the per-frame inference time is a property of the deployed hardware and is measured as part of the wireless experiment (Sec. 6).

5.3.2 Comparison with Recent Methods

Table 1 positions the proposed U-VAE-MF against representative recent deep-learning-based SIC methods [1, 12, 15, 17, 31–35]. A quantitative head-to-head comparison cannot be made on an equal footing, because each method is evaluated under different waveforms, impairment sets, and hardware assumptions; we therefore compare the methodological dimensions that determine practical applicability. Three differentiators stand out. First, prior neural SIC methods operate almost exclusively in the digital domain and model a single or a subset of impairments, whereas U-VAE-MF is built on a coupled I/Q-imbalance/phase-noise/PA-nonlinearity/multipath model and jointly designs the transmitter compensation and the receiver canceller. Second, none of the prior architectures regularizes its latent representation; the VAE module of U-VAE-MF is the only one that enforces a continuous latent space, which is precisely what prevents the 15 dB reconstruction deviation observed for the plain U-Net in Fig. 8. Third, U-VAE-MF is the only method that verifies its learned canceller against a closed-form theoretical cancellation limit, providing a physically grounded performance bound rather than an isolated figure of merit.

5.4 Component Contribution Analysis

The contribution of each component of U-VAE-MF follows from its role in the signal-processing chain. The multi-feedback transmitter stage is designed to provide the dominant contribution, because without transmitter pre-compensation the network would have to fit the full nonlinear self-interference directly from the saturated PA output, where the deep-saturation intermodulation components dominate the reconstruction error. The VAE regularization provides a further, clearly visible contribution: its absence is exactly the configuration analyzed in Fig. 8, where the plain U-Net deviates by roughly 15 dB from the ground truth at the main-signal frequency on unseen samples. The U-Net skip connections are designed to preserve the fine-grained signal detail across scales. The three components are therefore complementary: the multi-feedback stage reduces the dynamic range of the interference that the network must represent, the U-Net skip connections preserve signal detail, and the VAE regularization confers robustness to distribution shifts; together, they allow the canceller to approach the theoretical limit.

5.5 Generalization and Model-Mismatch Robustness

Because the U-VAE learns the residual mapping directly from data rather than relying on a parametric PA or phase-noise model, its performance is expected to degrade gracefully when the test conditions deviate from the training assumptions. Three representative deviations illustrate this point. First, if the Rapp PA were replaced by a Saleh-type model with additional AM-PM conversion, the unmodeled distortion would be absorbed by the learned residual mapping instead of violating a fixed calibration model. Second, a wider LO phase-noise bandwidth than the nominal value only increases the variance σ^2 in (41), whose effect is already accounted for by the closed-form limit. Third, the mechanism is not tied to the excitation statistics of the waveform. Both waveforms used in this paper (LFM for the sensing link and QPSK-OFDM for the communication link) are already evaluated in the isolation and BER comparisons of Figs. 20 and 21. A change of excitation changes these statistics but not the mechanism itself, since the network operates on the received mixture rather than on a waveform-specific model. The extension to higher-order modulation formats such as 64-QAM has not been measured here, and it is identified as future work in Sec. 6. Together, these considerations suggest that the proposed method does not rely on a specific PA model, phase-noise profile, or waveform, and that its performance may generalize to conditions not represented in the training set.

6. Conclusions

We have presented U-VAE-MF, a multi-feedback SI cancellation scheme integrating a U-Net variational autoencoder to jointly address PA nonlinearity, I/Q imbalance, and time-varying channels in IBFD systems.

We first derived a closed-form SI cancellation limit, identifying LO phase noise and receiver thermal noise as the irreducible performance floor, and verified it through simulation across input power and bandwidth conditions. The multi-feedback architecture corrects I/Q imbalance and PA nonlinearity via closed-loop digital predistortion, eliminating the repeated matrix inversions of conventional LS estimation. The U-VAE with latent-space regularization further suppresses high-order nonlinear residuals and stabilizes SOI reconstruction against time-varying distribution shifts.

Across all tested operating conditions, U-VAE-MF achieves 60–85 dB SI suppression, outperforming conventional LS by up to 38 dB at 0 dBm input power, with estimation latency reduced by 63% (270 µs to 100 µs) and overall performance within 2 dB of the theoretical limit. The component-wise discussion in Sec. 5 attributes this gain to the complementary action of the multi-feedback stage, the U-Net skip connections, and the VAE regularization, while the generalization discussion in the same section indicates that the performance is not tied to a specific PA model, phase-noise profile, or waveform.

Four limitations remain. First, all results are simulation-based; software-defined radio (SDR) hardware validation and experimental verification of fixed-point loss and phase-noise coherence are left for future work. Second, quasi-static channels are assumed; under high Doppler, feedback latency may approach the coherence time and degrade cancellation. Third, the training pipeline assumes a fixed impairment distribution; RF front-end drift outside the training range would require recalibration or online fine-tuning. Fourth, the results are limited to the impairment ranges in Tab. 2 and the LFM waveform family, and are not claimed to generalize beyond them.

Future work will validate the U-VAE-MF chain on an SDR platform using an impairment set matching the simulations, reproduce the cancellation-versus-power and bandwidth-robustness experiments, extend the BER evaluation to higher-order constellations such as 64-QAM, and measure the per-frame inference time of the deployed network on the same platform. The main hardware risks are already captured by our error model. The transferability of the present results will be assessed on independent datasets containing different kinds of samples, in a systematic component-wise ablation with re-trained variants, and in a head-to-head comparison with recent learning-based cancellers under matched waveform, impairment and hardware settings. In parallel, we plan to pursue three directions of our own interest. The first is online and continual learning, so that the canceller can follow slow drift of the RF front end without periodic recalibration. The second is quantization-aware training and pruning for field-programmable gate array (FPGA) deployment at reduced precision. The third is an extension of the coupled-impairment model and of the closed-form limit

to multi-antenna and integrated sensing and communication (ISAC) transceivers, including massive MIMO, mmWave and 6G operation. The complete model, database generation, parameters, and training configuration are given in Sec. 2–5 and Appendix A, enabling direct re-implementation. The dataset, scripts, and trained weights are available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No. 62301570.

References

[1] ELSAYED, M., EL-BANNA, A. A. A., DOBRE, O. A., et al. Full-duplex self interference cancellation using dual-neurons neural networks. *IEEE Communications Letters*, 2022, vol. 26, no. 3, p. 557–561. DOI: 10.1109/LCOMM.2021.3136030

[2] GIORDANI, M., POLESE, M., MEZZAVILLA, M., et al. Toward 6G networks: Use cases and technologies. *IEEE Communications Magazine*, 2020, vol. 58, no. 3, p. 55–61. DOI: 10.1109/MCOM.001.1900411

[3] HEINO, M., KORPI, D., HUUSARI, T., et al. Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays. *IEEE Communications Magazine*, 2015, vol. 53, no. 5, p. 91–101. DOI: 10.1109/MCOM.2015.7105647

[4] ZHANG, Z., CHAI, X., LONG, K., et al. Full duplex techniques for 5G networks: Self interference cancellation, protocol design, and relay selection. *IEEE Communications Magazine*, 2015, vol. 53, no. 5, p. 128–137. DOI: 10.1109/MCOM.2015.7105651

[5] CHEN, W. 5G-advanced toward 6G: Past, present, and future. *IEEE Journal on Selected Areas in Communications*, 2023, vol. 41, no. 6, p. 1592–1619. DOI: 10.1109/JSAC.2023.3274037

[6] ZHU, H., HU, W., WANG, C., et al. Neural network-based self interference cancellation method for in-band full-duplex system. *IET Conference Proceedings*, 2024, no. 47, p. 1528–1535. DOI: 10.1049/icp.2024.1312

[7] LEE, J.-H. Self interference cancellation using phase rotation in full-duplex wireless. *IEEE Transactions on Vehicular Technology*, 2013, vol. 62, no. 9, p. 4421–4429. DOI: 10.1109/TVT.2013.2264064

[8] KORPI, D., CHOI, Y.-S., HUUSARI, T., et al. Adaptive nonlinear digital self interference cancellation for mobile inband full-duplex radio: Algorithms and RF measurements. In *IEEE Global Communications Conference (GLOBECOM)*. San Diego (USA), 2015, p. 1–7. DOI: 10.1109/GLOCOM.2015.7417188

[9] SMIDA, B., SABHARWAL, A., FODOR, G., et al. Full-duplex wireless for 6G: Progress brings new opportunities and challenges. *IEEE Journal on Selected Areas in Communications*, 2023, vol. 41, no. 9, p. 2729–2750. DOI: 10.1109/JSAC.2023.3287612

[10] KOLODZIEJ, K. E., COOKSON, A. U., PERRY, B. T. RF canceller tuning acceleration using neural network machine learning for in-band full-duplex systems. *IEEE Open Journal of the Communications Society*, 2021, vol. 2, p. 1158–1170. DOI: 10.1109/OJCOMS.2021.3080618

[11] HONG, Z. H., ZHANG, L., WU, Y., et al. Iterative successive nonlinear self interference cancellation for in-band full-duplex communications. *IEEE Transactions on Broadcasting*, 2024, vol. 70, no. 1, p. 2–13. DOI: 10.1109/TBC.2023.3291136

[12] KONG, D. H., KIL, Y.-S., KIM, S.-H. Neural network aided digital self interference cancellation for full-duplex communication over time-varying channels. *IEEE Transactions on Vehicular Technology*, 2022, vol. 71, no. 6, p. 6201–6213. DOI: 10.1109/TVT.2022.3158402

[13] MASMOUDI, A., LE-NGOC, T. Channel estimation and self interference cancelation in full-duplex communication systems. *IEEE Transactions on Vehicular Technology*, 2016, vol. 66, no. 1, p. 321–334. DOI: 10.1109/TVT.2016.2540538

[14] MAO, Q., HU, F., HAO, Q. Deep learning for intelligent wireless networks: A comprehensive survey. *IEEE Communications Surveys and Tutorials*, 2018, vol. 20, no. 4, p. 2595–2621. DOI: 10.1109/COMST.2018.2846401

[15] GUO, H., WU, S., WANG, H., et al. DSIC: Deep learning based self interference cancellation for in-band full duplex wireless. In *IEEE Global Communications Conference (GLOBECOM)*. Waikoloa (USA), 2019, p. 1–6. DOI: 10.1109/GLOBECOM38437.2019.9013521

[16] HU, C., CHEN, Y., WANG, Y., et al. Digital self interference cancellation for full-duplex systems based on deep learning. *AEU - International Journal of Electronics and Communications*, 2023, vol. 168, p. 1–10. DOI: 10.1016/j.aeue.2023.154707

[17] HU, C., CHEN, Y., WANG, Y., et al. A hybrid digital self interference cancellation method with attention-based TCN-GRU for full-duplex systems. *AEU - International Journal of Electronics and Communications*, 2024, vol. 176, p. 1–11. DOI: 10.1016/j.aeue.2024.155144

[18] KOLODZIEJ, K. E., PERRY, B. T., HERD, J. S. In-band full-duplex technology: Techniques and systems survey. *IEEE Transactions on Microwave Theory and Techniques*, 2019, vol. 67, no. 7, p. 3025–3041. DOI: 10.1109/TMTT.2019.2896561

[19] FUKUI, T., KOMATSU, K., MIYAJI, Y., et al. Analog self interference cancellation using auxiliary transmitter considering IQ imbalance and amplifier nonlinearity. *IEEE Transactions on Wireless Communications*, 2020, vol. 19, no. 11, p. 7439–7452. DOI: 10.1109/TWC.2020.3011467

[20] HONG, Z. H., ZHANG, L., LI, W., et al. Frequency-domain RF self interference cancellation for in-band full-duplex communications. *IEEE Transactions on Wireless Communications*, 2023, vol. 22, no. 4, p. 2352–2363. DOI: 10.1109/TWC.2022.3211196

[21] KOLODZIEJ, K. E., MCMICHAEL, J. G., PERRY, B. T. Multitap RF canceller for in-band full-duplex wireless communications. *IEEE Transactions on Wireless Communications*, 2016, vol. 15, no. 6, p. 4321–4334. DOI: 10.1109/TWC.2016.2539169

[22] BANELLI, P., CACOPARDI, S. Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels. *IEEE Transactions on Communications*, 2000, vol. 48, no. 3, p. 430–441. DOI: 10.1109/26.837046

[23] CABRAL, P. M., PEDRO, J. C., CARVALHO, N. B. Nonlinear device model of microwave power GaN HEMTs for high power-amplifier design. *IEEE Transactions on Microwave Theory and Techniques*, 2004, vol. 52, no. 11, p. 2585–2592. DOI: 10.1109/TMTT.2004.837196

[24] KWONG, R. H., JOHNSTON, E. W. A variable step size LMS algorithm. *IEEE Transactions on Signal Processing*, 1992, vol. 40, no. 7, p. 1633–1642. DOI: 10.1109/78.143435

[25] MORGAN, D. R., MA, Z., KIM, J., et al. A generalized memory polynomial model for digital predistortion of RF power amplifiers. *IEEE Transactions on Signal Processing*, 2006, vol. 54, no. 10, p. 3852–3860. DOI: 10.1109/TSP.2006.879264

[26] LUO, H., HOLM, M., RATNARAJAH, T. On the performance of active analog self interference cancellation techniques for beyond 5G systems. *China Communications*, 2021, vol. 18, no. 10, p. 158–168. DOI: 10.23919/JCC.2021.10.011

[27] ZHU, H. Y., WANG, C., YUAN, N. C., et al. Multi-domain blind source separation in-band full-duplex technique considering RF impairments. *Journal of Physics: Conference Series*, 2023, vol. 2625, no. 1, p. 1–12. DOI: 10.1088/1742-6596/2625/1/012060

[28] LIAO, W.-S., ZHAO, O., LI, K., et al. Implementation of in-band full-duplex using software defined radio with adaptive filter-based self interference cancellation. *Future Internet*, 2023, vol. 15, no. 11, p. 1–19. DOI: 10.3390/FI15110360

[29] LE, D. M., PATIL, O. S., NINO, C. F., et al. Accelerated gradient approach for deep neural network-based adaptive control of unknown nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2025, vol. 36, no. 4, p. 6299–6313. DOI: 10.1109/TNNLS.2024.3395064

[30] BALATSOUKAS-STIMMING, A. Non-linear digital self-interference cancellation for in-band full-duplex radios using neural networks. In *IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. Kalamata (Greece), 2018, p. 1–5. DOI: 10.1109/SPAWC.2018.8445987

[31] ELSAYED, M., EL-BANNA, A. A. A., DOBRE, O. A., et al. Low complexity neural network structures for self-interference cancellation in full-duplex radio. *IEEE Communications Letters*, 2021, vol. 25, no. 1, p. 181–185. DOI: 10.1109/LCOMM.2020.3024063

[32] ELSAYED, M., EL-BANNA, A. A. A., DOBRE, O. A., et al. Hybrid-layers neural network architectures for modeling the self-interference in full-duplex systems. *IEEE Transactions on Vehicular Technology*, 2022, vol. 71, no. 6, p. 6291–6302. DOI: 10.1109/TVT.2022.3159535

[33] LU, L., CHU, J., TANG, Y., et al. Driven non-linear digital self interference cancellation for in-band full duplex systems based on convolution long short-term memory deep neural network. *Journal of Electronics and Information Technology*, 2022, vol. 44, no. 11, p. 3874–3881. DOI: 10.11999/JEIT220110

[34] ZHAO, Z., XIA, S., GAO, L., et al. Digital nonlinear self-interference cancellation using SW-GRU-TPA. *AEU - International Journal of Electronics and Communications*, 2024, vol. 186, p. 1–10. DOI: 10.1016/j.aeue.2024.155463

[35] ZHAO, Z., NING, W., GAO, L., et al. Digital nonlinear self-interference cancellation for full-duplex OFDM systems using fast Kolmogorov–Arnold networks. *AEU - International Journal of Electronics and Communications*, 2026, vol. 213, p. 1–14. DOI: 10.1016/j.aeue.2026.156362

[36] ELSAYED, M., EL-BANNA, A. A. A., DOBRE, O. A., 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*, 2024, vol. 5, p. 21–46. DOI: 10.1109/OJVT.2023.3331185

[37] ZHU, H., HUANG, J., WANG, C., et al. Multi-domain self interference cancellation methods considering RF imperfections. *IEEE Transactions on Wireless Communications*, 2024, vol. 23, no. 8, p. 8668–8682. DOI: 10.1109/TWC.2024.3353306

[38] ZHANG, Z., XIAO, Y., MA, Z., et al. 6G wireless networks: Vision, requirements, architecture, and key technologies. *IEEE Vehicular Technology Magazine*, 2019, vol. 14, no. 3, p. 28–41. DOI: 10.1109/MVT.2019.2921208

[39] ROBERTS, I. P., ANDREWS, J. G., JAIN, H. B., et al. Millimeter-wave full duplex radios: New challenges and techniques. *IEEE Wireless Communications*, 2021, vol. 28, no. 1, p. 36–43. DOI: 10.1109/MWC.001.2000221

[40] AHMED, E., ELTAWIL, A. M. All-digital self interference cancellation technique for full-duplex systems. *IEEE Transactions on Wireless Communications*, 2015, vol. 14, no. 7, p. 3519–3532. DOI: 10.1109/TWC.2015.2407876

[41] CHOI, Y.-S., SHIRANI-MEHR, H. Simultaneous transmission and reception: Algorithm, design and system level performance. *IEEE Transactions on Wireless Communications*, 2013, vol. 12, no. 12, p. 5992–6010. DOI: 10.1109/TWC.2013.101713.121152

[42] DUARTE, M., DICK, C., SABHARWAL, A. Experiment-driven characterization of full-duplex wireless systems. *IEEE Transactions on Wireless Communications*, 2012, vol. 11, no. 12, p. 4296–4307. DOI: 10.1109/TWC.2012.102612.111278

[43] RAPP, C. Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting system. In *2nd European Conference on Satellite Communications (ECSC-2)*. Liege (Belgium), 1991, p. 179–186.

[44] KINGMA, D. P., BA, J. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations (ICLR)*. San Diego (USA), 2015, p. 1–15. DOI: 10.48550/arXiv.1412.6980

About the Authors . . .

HongYu ZHU was born in 1994. He received his Ph.D. degree in Engineering from the National University of Defense Technology (NUDT) in 2025. He is currently a postdoctoral researcher at NUDT. His main research interests include full-duplex self-interference cancellation and jamming waveform optimization.

WeiWei WU received her Ph.D. degree in Engineering from the National University of Defense Technology (NUDT). She is currently an associate professor and doctoral supervisor at NUDT. Her main research interests include intelligent metasurface design and electronic countermeasures in complex electromagnetic environments.

ZhiHao ZHENG completed both his undergraduate and postgraduate studies at the National University of Defense Technology (NUDT), China, and received his master's degree in 2026. His primary research interests lie in artificial intelligence and intelligent waveform design.

JingJian HUANG is with the College of Electronic Science and Technology, National University of Defense Technology (NUDT), Changsha, China. His research direction is electromagnetic fields and waves and complex antenna design. His main research interests include full-duplex self-interference cancellation, RF impairment modeling, and intelligent signal processing for wireless communication systems.

QingZhan SHI is an Associate Research Fellow with the College of Electronic Science and Technology, National University of Defense Technology (NUDT), Changsha, China. His research direction is intelligent signal processing and waveform design. His main research interests include deep learning for communication signal processing, nonlinear system identification, and self-interference cancellation in full-duplex systems.

XiaoFa ZHANG is with the College of Electronic Science and Technology, National University of Defense Technology (NUDT), Changsha, China. His research direction is complex system integration and RF system design, with specific interests in RF front-end impairment analysis, digital predistortion, and intelligent waveform design.

NaiChang YUAN is a Professor with the College of Electronic Science and Technology, National University of Defense Technology (NUDT), Changsha, China. His research interests include in-band full-duplex communication, electronic countermeasures, radar signal processing, system integration of phased-array systems, and electromagnetic fields and waves.

Appendix A: Derivation of the Phase-Noise Residual Power

Given the residual phase factor

$$P(l) = \exp\big(j\Phi((n+l_1)T_\mathrm{s})\big) - \exp\big(j\Phi((n+l_2)T_\mathrm{s}-\tau)\big) \quad \text{(A1)}$$

we aim to compute

$$\mathbb{E}_2(l) = \mathbb{E}\Big[\,\big|P(l)\big|^2\,\Big]. \quad \text{(A2)}$$

1. Phase noise statistics.

The LO phase noise $\Phi(t)$ is modeled as a zero-mean Gaussian process with a Lorentzian power spectral density. For such a process, the phase increment over a time interval Δt is a zero-mean Gaussian random variable whose variance scales linearly with $|\Delta t|$:

$$\mathbb{E}\Big[\,\big(\Phi(t+\Delta t) - \Phi(t)\big)^2\,\Big] = 2\pi\beta\,|\Delta t| \quad \text{(A3)}$$

where β (in Hz) is the single-sided 3-dB bandwidth of the LO phase noise spectrum, consistent with the definition in Sec. 2.

2. Time instants and their difference.

Define the two time instants appearing in (A1) as

$$t_a = (n+l_1)T_\mathrm{s}, \qquad t_b = (n+l_2)T_\mathrm{s} - \tau. \quad \text{(A4)}$$

Their difference is

$$\Delta t = t_a - t_b = (l_1 - l_2)T_\mathrm{s} + \tau = lT_\mathrm{s} + \tau \quad \text{(A5)}$$

where $l = l_1 - l_2$. Consequently, the phase difference

$$\Delta\Phi(l) = \Phi(t_a) - \Phi(t_b) \quad \text{(A6)}$$

which is a zero-mean Gaussian random variable with variance

$$\sigma^2 = \mathbb{E}\big[(\Delta\Phi(l))^2\big] = 2\pi\beta\,|lT_\mathrm{s} + \tau|. \quad \text{(A7)}$$

3. Expansion of the squared magnitude.

For any two angles Φ_a, Φ_b,

$$\begin{aligned}\big|e^{j\Phi_a} - e^{j\Phi_b}\big|^2 &= \big(e^{j\Phi_a} - e^{j\Phi_b}\big)\big(e^{-j\Phi_a} - e^{-j\Phi_b}\big)\\ &= 2 - \big(e^{j(\Phi_a-\Phi_b)} + e^{-j(\Phi_a-\Phi_b)}\big)\\ &= 2 - 2\cos(\Phi_a - \Phi_b). \end{aligned} \quad \text{(A8)}$$

Applying this to $\Phi_a = \Phi(t_a)$, $\Phi_b = \Phi(t_b)$, we obtain

$$\big|P(l)\big|^2 = 2 - 2\cos(\Delta\Phi(l)). \quad \text{(A9)}$$

4. Expectation of the cosine term.

Taking the expectation of (A8) yields

$$\mathbb{E}_2(l) = 2 - 2\,\mathbb{E}[\cos(\Delta\Phi(l))]\,. \quad \text{(A10)}$$

Since $\Delta\Phi(l) \sim \mathcal{N}(0, \sigma^2)$, its characteristic function is

$$\varphi(\omega) = \mathbb{E}[\exp(j\omega\Delta\Phi(l))] = \exp\Big(-\frac{1}{2}\omega^2\sigma^2\Big). \quad \text{(A11)}$$

Setting $\omega = 1$ gives

$$\mathbb{E}[\exp(j\Delta\Phi(l))] = \exp(-\sigma^2/2). \quad \text{(A12)}$$

The random variable $\Delta\Phi(l)$ has a symmetric (even) probability density because it is zero-mean Gaussian. Hence the imaginary part of $\mathbb{E}[e^{j\Delta\Phi(l)}]$, which equals $\mathbb{E}[\sin(\Delta\Phi(l))]$, vanishes, and we are left with

$$\mathbb{E}[\cos(\Delta\Phi(l))] = \mathrm{Re}\big(\mathbb{E}[\exp(j\Delta\Phi(l))]\big) = \exp(-\sigma^2/2). \quad \text{(A13)}$$

5. Final expression.

Substituting σ^2 from (A7) and the cosine expectation into $\mathbb{E}_2(l)$, we obtain

$$\begin{aligned}\mathbb{E}_2(l) &= 2 - 2\,\exp(-\sigma^2/2)\\ &= 2 - 2\,\exp\Big(-\frac{1}{2}\cdot 2\pi\beta\,|lT_\mathrm{s} + \tau|\Big)\\ &= 2 - 2\,\exp\Big(-\pi\beta\,|lT_\mathrm{s} + \tau|\Big). \end{aligned} \quad \text{(A14)}$$

The quantity $\mathbb{E}_2(l)$ is the average power (second moment) of the residual phase factor $P(l)$, not its variance. In the literature it is sometimes loosely referred to as "phase noise variance", but it actually characterizes the power of the interference caused by the phase noise difference between two time instants.