

Research Article

A Lazy Learning-Based Self-Interference Cancellation Approach for In-Band Full-Duplex Wireless Communication Systems

Ou Zhao , Wei-Shun Liao , Keren Li , Takeshi Matsumura , Fumihide Kojima , and Hiroshi Harada 

National Institute of Information and Communications Technology (NICT), 3-4 Hikarino-oka, Yokosuka 239-0847, Japan

Correspondence should be addressed to Ou Zhao; zhaoou@nict.go.jp

Received 25 March 2022; Revised 31 May 2022; Accepted 20 July 2022; Published 3 August 2022

Academic Editor: Lisheng Fan

Copyright © 2022 Ou Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a new lazy learning-based cancellation approach to improve spectral efficiency for current wireless communication systems, suppress self-interference (SI) sent from base stations, and enable in-band full-duplex (IBFD) transmissions in cellular networks. Our proposed approach consists of two phases based on traditional IBFD systems: an offline phase for database generation and an online phase for data transmission. In the offline phase, the output before a 0/1 decision is premeasured without the desired signal input and recorded in a database with self-defined feature vectors (FVs). In the online phase, a suitable result is sought from the generated database with the help of a learning method and FV for the same system architecture with the desired signal input. The result is then assigned an SI cancellation value. Regular and eager learning-based cancellation approaches are employed to evaluate the proposed method and simulate the transmission output. Computer simulation results indicated that the proposed cancellation methods could achieve about 134 dB SI suppression and achieve nearly the same transmission levels as methods with no SI effect, enabling the IBFD operations in wireless communication systems better than the regular and eager learning-based techniques.

1. Introduction

Wireless communication traffic has been rapidly increasing with the prevalence of smartphone applications and Internet of Things devices. Development of new spectrum resources and improvement of spectral efficiency (SE) to provide proliferating wireless traffic for future mobile communication systems beyond the 5th generation require time, effort, and money [1, 2].

This study concentrates on a potentially disruptive technology called in-band full-duplex (IBFD) to increase the SE in wireless communication systems. IBFD systems can double the SE under ideal conditions compared to traditional time-division duplex (TDD) and frequency-division duplex (FDD) systems. Independent transmitting and receiving are performed in TDD systems over a common frequency band (FB), and inversely simultaneous transmitting and

receiving are carried out in FDD systems using an independent FB. Consequently, the protocol for simultaneous transmission and reception over the same FB is adopted. However, the quality of the desired signal seriously deteriorates because of the self-interference (SI) effect caused by employing the same FB for transmitting and receiving. Thus, the SE either improves slightly or becomes worse [3].

Currently, various approaches have been proposed in the existing works to suppress SI and enhance SE, such as antenna, analog, and digital cancellation [4–11]. Generally, antenna cancellation is aimed at increasing the isolation between transmission and reception [5]. Analog cancellation is used to suppress the SI power by combining a reference SI signal in which the phase and amplitude are adjusted [6, 7, 11]. However, because of several physical constraints upon antenna design and inaccuracies in obtaining SI signals in analog circuits, residual SI remains powerful in the desired

signal [12]. Therefore, digital cancellation (DC) is introduced to construct a replica of the SI and subtract it from the received composite signals [9].

The desired portion of the received composite signal should be successfully demodulated if its power is tens of decibels larger than the SI power through a SI canceller. Unfortunately, in most common wireless communication systems, the power of the desired signal component is not much greater than system noise power, and the noise power is difficult to suppress. Therefore, the biggest challenge for the design of the SI canceller is to decrease the SI power down to the system noise level and double the SE. For instance, suppose that the considered IBFD operation is used in a common base station (BS) with an equivalent isotopically radiated power (EIRP) of more than 40 dBm and an average noise power of less than -100 dBm at the receiver component. The combined effect of the antenna, analog, and DCs is preferably encouraged to 140 dB or more to suppress the SI power down to the noise level [12]. In fact, a cross-polarization technique yielding antenna cancellation of 50 dB was presented in [13], and a combined analog and DC of 60 dB was proposed in [14]. These exciting results showed about 110 dB combined suppression effect for SI power can be achieved in experimental environments. Unfortunately, probably because of several hardware restrictions such as maximum output power of device and noise floor, currently, there have been few results related to SI cancellation approaches with suppression capabilities over 110 dB to the best of our knowledge.

Furthermore, most previous works such as [6, 7, 10, 15, 16] using analog and digital cancellers for IBFD systems are based on regular methods. These methods estimate the SI signal effects experienced in the time and frequency domains, including nonlinearity in the transmitter power amplifier (PA) and low noise amplifier (LNA), wireless channel propagation, and circuit delay. However, since the power gap between the SI and received signals is considerable, the nonlinearity of the PA and LNA, connector return and insertion losses, and other unknown losses are always present in practice. Therefore, a small error in the estimation process may cause large residual SI, and strict measurements are required in the whole system, even though that is challenging to ensure. Consequently, a tiny fracture grows into a chasm.

Several researchers such as in [7, 17–22] inspired by big data and machine learning (i.e., two of the most popular and useful technologies [23–25]) tried to cancel the effects of SI by (including but not limited to) predicting SI signals or estimating channel state information (CSI) to reconstruct SI waveforms with the help of the deep learning method, which is representative of eager learning (EL) in machine learning techniques. For example, the work of [18] proposed a real-time nonlinear SI cancellation solution using deep learning to realize IBFD wireless communication. In this solution, SI channel is modeled by a deep neural network (NN), and the NN is trained for cancellation of SI at wireless node. The results from their software-defined radio- (SDR-) based testbed showed a performance of 17 dB in DC and yielded an average of 8.5% bit error rate (BER) over many scenarios

and different modulation schemes. Similarly, to enable IBFD transmissions, the authors in [19] proposed a nonlinear DC approach by adapting support vector regression which is one of algorithms to solve regression problems in machine learning. Their tests were also performed on SDR-based platform and indicated that their proposal can provide more than 30 dB digital suppression for transmit power levels higher than 20 dBm.

Recently, a joint detection and nonlinear SI cancellation approach using EL with NN was proposed in [20]. In this work, the EL method is used to derive a function between output of desired binary data and received signal, and thus, the desired signal can be directly demodulated in the presence of SI. Although several questions need to be addressed in their future works, the preliminary experimental results showed that EL techniques can perform better comparing to conventional SI cancellation techniques. Moreover, hybrid beamforming design for IBFD millimeter wave systems using EL-based scheme was proposed in [21]. In this work, two frameworks based on extreme learning machine and convolutional NN were presented to design hybrid beamformers and further achieve SI cancellation. Their results showed that both learning-based schemes can provide more robust performance, improve spectral efficiency, and decrease computation time. In additional, an alternative application using EL for IBFD systems was proposed in [22]. The authors in this work introduced a use of EL with NN to accelerate tuning of multitap adaptive radio frequency (RF) cancellers by training the weights of in-phase and quadrature channels in adaptive cancellers. The results illustrated a fast convergence speed by using EL in RF cancellers and thereby enabled IBFD operation in dynamic interference environments.

Generally, in many cases, an eager learner abstracts away from the data during training and uses the trained model (abstraction) to make predictions. The most important benefit of using EL is solving the SI problem; time-insensitive parameters (e.g., PA and LNA nonlinearity as well as antenna return loss) in the entire system can be included in the trained model and do not have to be estimated. However, prediction errors caused by the trained model and estimated CSI for the SI calculation always exist and cannot be avoided because the model training is an essential operation [26]. Thus, errors can unexpectedly be amplified by powerful SI signals similar to the regular method described above.

Inspired by the previous discussion, it can be known that estimating or reconstructing SI signal waveforms for cancellation is unnecessary. However, all practical effects on the desired signal caused by the SI must be quantified and tagged with feature vectors (FVs), and these quantified values can be precisely found by using these FVs when required [27]. That highlights the difference with the model generation-based learning methods such as [17–22]. More specifically, the effects of the SI on wireless transmissions can be quantified by constellation values at the 0/1 decision in a signal demodulator; an FV can be defined as a set constructed with the estimated CSI between the interested user, antennas, and transmitted symbol from the BS. Contrary to EL, the lazy learning (LL) method [27, 28] does not require

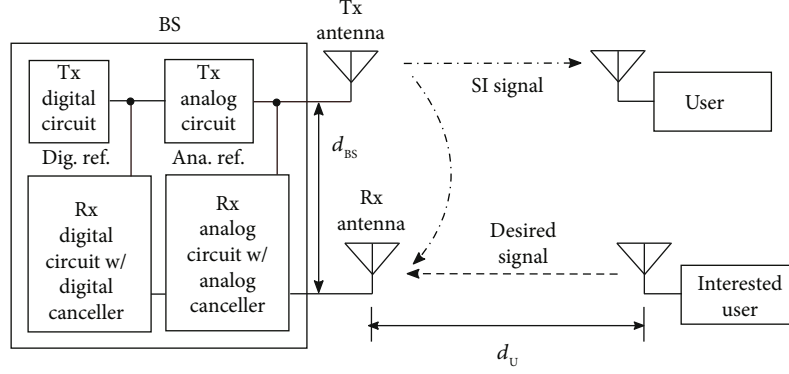


FIGURE 1: Unidirectional IBFD system illustrations.

model training, thereby the prediction error is not introduced in related operations for SI cancellation and may not be amplified by the powerful SI signals. Therefore, we propose an LL-based cancellation method to solve the SI problem.

We separately perform an offline phase for database generation and an online phase for data transmission. In the offline phase, the output before a 0/1 decision is premeasured without the desired signal input and recorded to a database with self-defined FVs. In the online phase, a suitable result is sought from the generated database with the help of a learning method and FV for the same system architecture with the desired signal input. The result is then assigned as the SI cancellation value. In other studies such as [22], the demodulation of composite signal, including desired signal, thermal noise, and potential residual SI, is independently performed after all of the cancellation processes are done (even the residual SI is still powerful). In our proposal, desired signals are directly demodulated in the presence of SI and thermal noise. We further provide system-level performance such as BER to evaluate the proposed approach in the IBFD systems, contrary to existing works [4–10, 12–16] where most studies only focused on the SI suppression capability.

The remainder of this paper is organized as follows. In Section 2, we describe the system architecture under consideration and formulate the problem. In Section 3, we explain three kinds of DC approaches in greater detail: the proposed technique, regular, and an EL-based method. Details regarding the employed channel models are provided in Section 4. In Section 5, we present and analyze simulated results and summarize the key findings. Discussion and concluding remarks are presented in Section 6 and Section 7, respectively.

2. System Model and Problem Formulation

2.1. System Model. We consider a unidirectional IBFD system in which an interested user equipped with a single antenna sends desired signals to a BS receiving antenna from a distance of d_U , while the BS simultaneously sends signals to other users via the same FB. The BS is equipped with a pair of transmitting and receiving antennas with an inter-antenna distance of d_{BS} , and the latter one is used to receive a

composite signal sent by the BS transmitting antenna (i.e., SI) and the interested user. Analog and digital cancellers are commonly employed in the IBFD BS structure for suppression of SI signals, and references from analog and digital transmitter circuits can provide supports for cancellation processes. An illustration of the considered unidirectional IBFD systems is shown in Figure 1.

In this study, we design a valid frame structure for transmitted signals, including the SI and desired signal, to evaluate the considered systems using the proposed cancellation approach. The structural design includes downlink and uplink aspects used for BS and an interested user, respectively. The downlink and uplink transmissions with N_{frame} frames indexed with i are performed, and there are N_{symbol} modulated symbols indexed with j in the data part in each frame. We assume that frame synchronization at receive side is perfect and pilot signals with N_{pilot} symbols are deployed at the head of the frame to obtain CSI at an arbitrary receiving terminal. An illustration of the frame structure is given in Figure 2.

2.2. Problem Formulation. In this study, the considered IBFD system is employed with the uses of antenna, analog, and DCs. An IBFD BS structure is shown in Figure 3. For j th data symbol in i th frame, the received waveform at BS which consists of desired signal, SI signal, and noise can be expressed as

$$\mathbf{y}(i, j) = G_{\text{ant,rx}}^{1/2} \mathbf{h}_U(i) \mathbf{x}_U(i, j) + \tilde{C}_{\text{ant}}^{1/2} G_{\text{ant,rx}}^{1/2} \mathbf{h}_{BS}(i) \mathbf{x}_{BS}(i, j) + \mathbf{w}_{\text{rx}}(i, j), \quad (1)$$

where $G_{\text{ant,rx}}$ is antenna gain of BS receive antenna. $\mathbf{h}_{\Psi}(i)$ for $\Psi = U$ and $\Psi = BS$ are defined as the channel gain coefficient of user-BS and the coefficient of BS's inter-antenna corresponding to the i th frame, respectively. We further use $\rho_{\Psi}(i)$ and $\phi_{\Psi}(i)$ to represent the amplitude coefficient and phase shift of the complex variable $\mathbf{h}_{\Psi}(i)$ so that

$$\mathbf{h}_{\Psi}(i) = \rho_{\Psi}(i) e^{j\phi_{\Psi}(i)}. \quad (2)$$

The length of frames is limited by the channel coherence time, and wireless channels are considered static

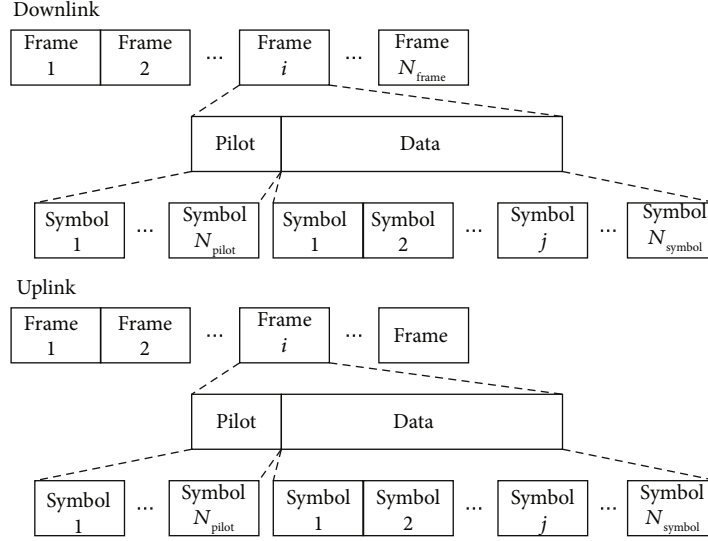


FIGURE 2: Frame structure illustration of IBFD systems.

during this coherence time [29]. Thus, effects due to the channel attenuation on all of the symbols in the same frame are identical and can be estimated using pilot signals. $\mathbf{w}_{\text{rx}}(i, j)$ is the complex additive white Gaussian noise (AWGN) at the receive antenna of BS with an average power of Ω_{rx} . \hat{C}_{ant} represents the antenna cancellation without path loss effect of interantenna at BS and further expressed as

$$\hat{C}_{\text{ant}} = C_{\text{ant}} I_{\text{BS}}^{-1}, \quad (3)$$

where L_{BS} represents the path loss between the antennas at BS and C_{ant} denotes the employed antenna cancellation, which can suppress the power of the SI signals.

$\mathbf{x}_{\Psi}(i, j)$ for $\Psi = \text{U}$ or $\Psi = \text{BS}$ is complex signal output from the transmit antenna of user or BS and can be expressed as

$$\mathbf{x}_{\Psi}(i, j) = G_{\text{ant,tx},\Psi}^{1/2} Q_{\text{PA},\Psi}(\mathbf{m}_{\Psi}(i, j) + \mathbf{w}_{\text{tx},\Psi}(i, j)) + G_{\text{ant,tx},\Psi}^{1/2} \mathbf{w}_{\text{PA},\Psi}(i, j), \quad (4)$$

where $G_{\text{ant,tx},\Psi}$ denotes antenna gain for user or BS. $Q_{\text{PA},\Psi}(z)$ is PA function for user or BS with third-order intermodulation distortion. According to [30], $Q_{\text{PA},\Psi}(z)$ can be expressed as

$$Q_{\text{PA},\Psi}(z) = G_{\text{PA},\Psi}^{1/2} z - \frac{4G_{\text{PA},\Psi}^{1/2}}{3O_{3,\Psi}} z^3, \quad (5)$$

where $G_{\text{PA},\Psi}$ denotes PA gain of user or BS, while $O_{3,\Psi}$ is the third-order intercept point (OIP3) of user or BS PA.

The terms of $\mathbf{w}_{\text{tx},\Psi}(i, j)$ are the complex AWGN existing in the user or BS modulator with an average power of $\Omega_{\text{tx},\Psi}$, while $\mathbf{w}_{\text{PA},\Psi}(i, j)$ is denoted as the additional output noises caused by PA of user or BS with an average power of $\Omega_{\text{PA},\Psi}$ that can be easily calculated by

$$\Omega_{\text{PA},\Psi} = (F_{\Psi} - 1) \Omega_{\text{tx},\Psi} G_{\text{PA},\Psi}^{1/2}, \quad (6)$$

where F_{Ψ} is noise factors in user or BS PA [31].

The terms of $\mathbf{m}_{\Psi}(i, j)$ in (4) represent the output complex signal from the user or BS modulator. For the use of a common modulation scheme, $\mathbf{m}_{\Psi}(i, j)$ can be written as

$$\mathbf{m}_{\Psi}(i, j) = A_{\Psi}(i, j) \cos(\theta_{\Psi}(i, j)) + j A_{\Psi}(i, j) \sin(\theta_{\Psi}(i, j)), \quad (7)$$

where $A_{\Psi}(i, j) = \sqrt{2E_{\text{symbol}}(i, j)f_{\text{symbol}}}$ and f_{symbol} is symbol rate for the considered modulator, while $E_{\text{symbol}}(i, j)$ denotes energy of j th symbol in i th frame. $\theta_{\Psi}(i, j)$ are the corresponding phase angles after symbol mapping processed on binary data which the user or BS needs to transmit.

Refer to Figure 3 with switches ON, after an analog cancellation C_{ana} for the SI signals in receive waveform $\mathbf{y}(i, j)$ is done, the CSI of antennas at BS and the CSI of user-BS part are estimated as

$$\tilde{\mathbf{h}}_{\Psi}(i) = \tilde{\rho}_{\Psi}(i) e^{j\tilde{\phi}_{\Psi}(i)}, \quad (8)$$

for $\Psi = \text{U}$ and $\Psi = \text{BS}$, respectively; then, a running of search algorithm for DC is employed even though it is not necessary in common wireless communication systems without IBFD operations. Thereafter, a signal detection method $\mathcal{D}(z)$ in the detector should be adopted and performed. The output of the detector as well as the input of demodulator is written as

$$\mathbf{y}_{\text{det}}(i, j) = \mathcal{D}\left(G_{\text{ant,rx}}^{1/2} \mathbf{h}_{\text{U}}(i) \mathbf{x}_{\text{U}}(i, j) + C_{\text{ana}}^{1/2} \hat{C}_{\text{ant}}^{1/2} G_{\text{ant,rx}}^{1/2} \mathbf{h}_{\text{BS}}(i) \mathbf{x}_{\text{BS}}(i, j) + \mathbf{w}_{\text{rx}}(i, j)\right). \quad (9)$$

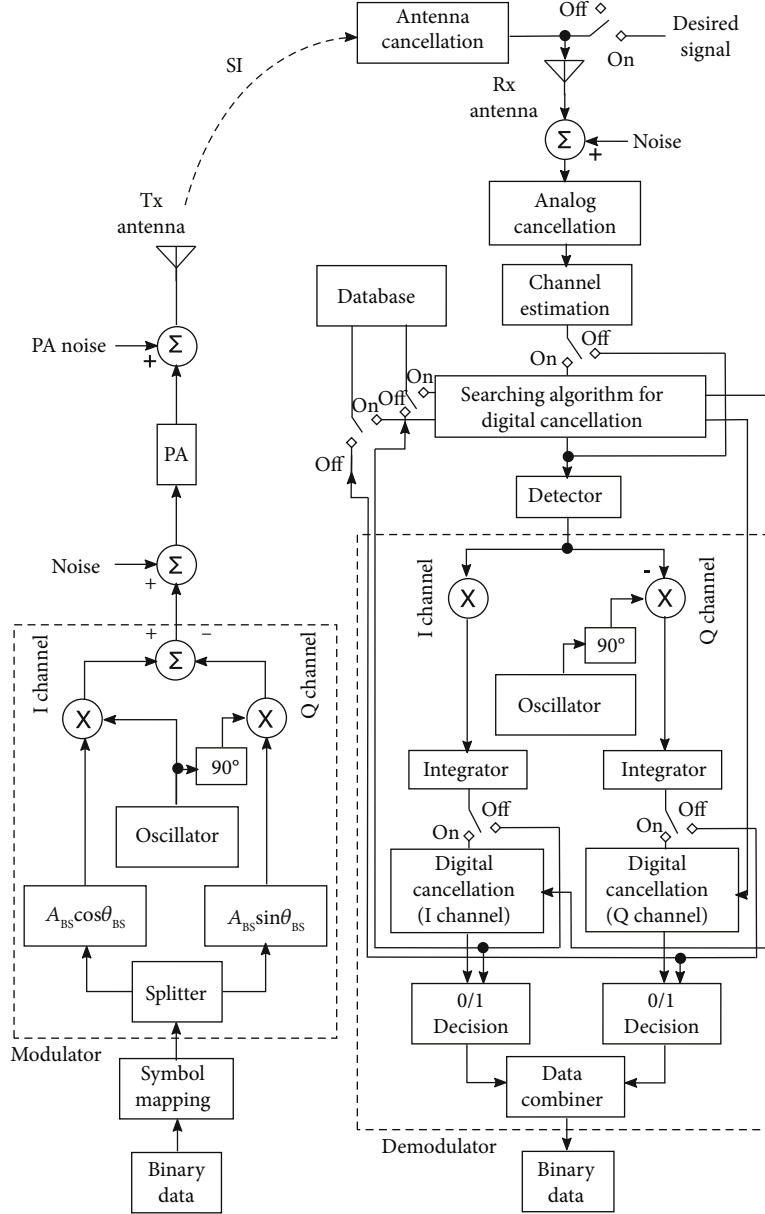


FIGURE 3: An IBFD BS structure with switches OFF for the offline phase and ON for the online phase. The offline phase is used for an offline-generated database in which constellation values of SI have been previously measured and recorded. The online phase is used for data transmission in which the desired signals are separated by subtracting the recorded SI from the received composite signals.

After a series of operations in demodulation with the uses of DCs C_{dig} which formed as a complex number, the inputs of 0/1 decision can be expressed as

$$\mathbf{y}_{\text{dec}}(i, j) = \mathbf{y}_{\text{det}}(i, j) - \mathbf{C}_{\text{dig}}(i, j, \mathbf{a}), \quad (10)$$

where the real part and imaginary part of $\mathbf{y}_{\text{dec}}(i, j)$, i.e., $\Re(\mathbf{y}_{\text{dec}}(i, j))$ and $\Im(\mathbf{y}_{\text{dec}}(i, j))$, are used for I and Q channels in 0/1 decision block, respectively. \mathbf{a} is an FV constructed by some input parameters for DC that is explained in the following section.

Substitute (9) and the related equations given in this section into (10), we rewrite (10) through the necessary mathe-

matical simplification as

$$\mathbf{y}_{\text{dec}}(i, j) = \mathbf{S}(i, j) + \mathbf{I}(i, j) - \mathbf{C}_{\text{dig}}(i, j, \mathbf{a}) + \mathbf{W}(i, j), \quad (11)$$

where $\mathbf{S}(i, j)$ and $\mathbf{I}(i, j)$ represent the complex constellation values of the desired and SI signals, respectively, and can be calculated by

$$\mathbf{S}(i, j) = G_{\text{ant,rx}}^{1/2} G_{\text{ant,tx,U}}^{1/2} G_{\text{PA,U}}^{1/2} \lambda^{-1} \mathbf{h}_{\text{U}}(i) \left(\mathbf{m}_{\text{U}}(i, j) - \frac{4}{3O_{3,\text{U}}} \mathbf{m}_{\text{U}}^3(i, j) \right), \quad (12)$$

$$\mathbf{I}(i, j) = C_{\text{ana}}^{1/2} \hat{C}_{\text{ant}}^{1/2} G_{\text{ant,rx}}^{1/2} G_{\text{ant,tx,BS}}^{1/2} G_{\text{PA,BS}}^{1/2} \lambda^{-1} \mathbf{h}_{\text{BS}}(i) \left(\mathbf{m}_{\text{BS}}(i, j) - \frac{4}{3O_{3,\text{BS}}} \mathbf{m}_{\text{BS}}^3(i, j) \right). \quad (13)$$

In the current study, a simple matched filter (MF) detection method [32] for $\mathcal{D}(\mathbf{z})$ is used and can roughly be expressed as

$$\mathcal{D}(\mathbf{z}) = \lambda^{-1} \mathbf{z}. \quad (14)$$

For the most common wireless communication systems, λ is assigned by $\tilde{\mathbf{h}}_{\text{U}}(i)$; thus, the desired signals can be detected because of $\tilde{\mathbf{h}}_{\text{U}}^{-1}(i) \mathbf{h}_{\text{U}}(i) \approx 1$ [32, 33]. The complex value $\mathbf{W}(i, j)$ caused by various noises can be further extended as

$$\mathbf{W}(i, j) = \mathbf{W}_{\text{tx,U}}(i, j) + \mathbf{W}_{\text{PA,U}}(i, j) + \mathbf{W}_{\text{tx,BS}}(i, j) + \mathbf{W}_{\text{PA,BS}}(i, j) + \mathbf{W}_{\text{rx}}(i, j), \quad (15)$$

where $\mathbf{W}_{\text{tx,U}}(i, j)$, $\mathbf{W}_{\text{PA,U}}(i, j)$, $\mathbf{W}_{\text{tx,BS}}(i, j)$, $\mathbf{W}_{\text{PA,BS}}(i, j)$, and $\mathbf{W}_{\text{rx}}(i, j)$ are defined as the complex values of noises generated from user's modulation and PA, BS's modulation and its PA, and the receive antenna at BS side, respectively, and can be easily calculated by

$$\begin{aligned} \mathbf{W}_{\text{tx,U}}(i, j) = & \left(G_{\text{ant,rx}}^{1/2} G_{\text{ant,tx,U}}^{1/2} G_{\text{PA,U}}^{1/2} \lambda^{-1} \mathbf{h}_{\text{U}}(i) \right. \\ & \cdot \left(1 - \frac{4}{O_{3,\text{U}}} \mathbf{m}_{\text{U}}^2(i, j) \right) \mathbf{w}_{\text{tx,U}}(i, j) \\ & \left. - \frac{4}{O_{3,\text{U}}} \mathbf{m}_{\text{U}}(i, j) \mathbf{w}_{\text{tx,U}}^2(i, j) - \frac{4}{3O_{3,\text{U}}} \mathbf{w}_{\text{tx,U}}^3(i, j) \right), \end{aligned} \quad (16)$$

$$\mathbf{W}_{\text{PA,U}}(i, j) = G_{\text{ant,rx}}^{1/2} G_{\text{ant,tx,U}}^{1/2} \lambda^{-1} \mathbf{h}_{\text{U}}(i) \mathbf{w}_{\text{PA,U}}(i, j), \quad (17)$$

$$\begin{aligned} \mathbf{W}_{\text{tx,BS}}(i, j) = & C_{\text{ana}}^{1/2} \hat{C}_{\text{ant}}^{1/2} G_{\text{ant,rx}}^{1/2} G_{\text{ant,tx,BS}}^{1/2} G_{\text{PA,BS}}^{1/2} \lambda^{-1} \mathbf{h}_{\text{BS}}(i) \\ & \cdot \left(\left(1 - \frac{4}{O_{3,\text{BS}}} \mathbf{m}_{\text{BS}}^2(i, j) \right) \times \mathbf{w}_{\text{tx,BS}}(i, j) \right. \\ & \left. - \frac{4}{O_{3,\text{BS}}} \mathbf{m}_{\text{BS}}(i, j) \mathbf{w}_{\text{tx,BS}}^2(i, j) - \frac{4}{3O_{3,\text{BS}}} \mathbf{w}_{\text{tx,BS}}^3(i, j) \right), \end{aligned} \quad (18)$$

$$\mathbf{W}_{\text{PA,BS}}(i, j) = C_{\text{ana}}^{1/2} \hat{C}_{\text{ant}}^{1/2} G_{\text{ant,rx}}^{1/2} G_{\text{ant,tx,BS}}^{1/2} \lambda^{-1} \mathbf{h}_{\text{BS}}(i) \mathbf{w}_{\text{PA,BS}}(i, j), \quad (19)$$

$$\mathbf{W}_{\text{rx}}(i, j) = \lambda^{-1} \mathbf{w}_{\text{rx}}(i, j), \quad (20)$$

respectively.

Actually, to get the desired binary data sent by the user successfully, we strongly hope that the input values of 0/1 decision, i.e., $\mathbf{y}_{\text{dec}}(i, j)$, are purely decided by the desired part $\mathbf{S}(i, j)$. Unfortunately, in IBFD systems, because the strength of SI signal at the BS's receive antenna is very huge compared to the desired signal, $\mathbf{y}_{\text{dec}}(i, j)$ presented in (11) is mainly dominated by $\mathbf{I}(i, j)$. That is the reason that we further propose the DC \mathbf{C}_{dig} to suppress the influence of SI signals. Define the residual SI as

$$\hat{\mathbf{I}}(i, j) = \mathbf{I}(i, j) - \mathbf{C}_{\text{dig}}(i, j, \boldsymbol{\alpha}), \quad (21)$$

naturally our target is to make $\hat{\mathbf{I}}(i, j)$ close to zero for arbitrary frame i and symbol j as much as possible.

3. Digital Cancellation Approaches

In this study, we propose an LL-based method to set DC and further enable the IBFD operations in the considered wireless communication systems. Several existing works such as [7, 16] have been studied on SI suppression with the help of deep learning, which is a typical representative of EL in machine learning techniques. Other works such as [6] employed a regular method in which the SI signal is approximately calculated using estimated CSI. In this section, we give more explanations for the DCs mentioned above.

3.1. Proposed LL-Based Cancellation Approach. Generally, LL is a learning method where generalization of the trained model is not necessary, and searching of the needed data is delayed until a query is made to the system, as opposed to in EL, where the system tries to generalize the model using training data before receiving queries [28]. The core of the proposed LL-based cancellation approach in our study is constructed by (1) an offline-generated database in which the constellation values of SI at the input of 0/1 decision block are premeasured and recorded and (2) an online data transmission in which desired signals are separated by subtracting the recorded SI from the received composite signals. The details are provided as follows.

3.1.1. Offline Database Generation. We first define a database \mathcal{B} which consists of an input space named \mathcal{B}_{in} and an output space named \mathcal{B}_{out} . According to (13), because SI is further affected by λ and λ is related to the coefficient of user channel for the considered MF detection method; in the input space \mathcal{B}_{in} , we define three independent parameters \bar{A}_{BS} , $\bar{\theta}_{\text{BS}}$, and $\bar{\mathbf{h}}_{\text{U}}$ and form an FV. The first and second parameters \bar{A}_{BS} and $\bar{\theta}_{\text{BS}}$ are assigned with amplitude and phase angle used in BS's modulation and can yield discrete values corresponding to the used modulation method. The last parameters $\bar{\mathbf{h}}_{\text{U}}$ are assigned with the channel coefficient \mathbf{h}_{U} between the interested user and BS, and in theory, its absolute value and phase can range from 0 to ∞ and 0 to 2π , respectively. The values in the input space \mathcal{B}_{in} are previously designed to take over all possible combinations of these parameters as much as possible under an acceptable level of computational cost.

To reduce the database size for our learning systems, in the output space \mathcal{B}_{out} , we define two parameters which are $\bar{\mathbf{h}}_{\text{BS}}$ and $\bar{\mathbf{I}}$. $\bar{\mathbf{h}}_{\text{BS}}$ is used to record the estimated values of channel coefficient \mathbf{h}_{BS} and can be considered as a tag of $\bar{\mathbf{I}}$ when we need to locate $\bar{\mathbf{I}}$ from the database \mathcal{B}_{out} . $\bar{\mathbf{I}}$ is used to record the measured constellation values of SI. To get the values for the output space \mathcal{B}_{out} , a series of provisional transmissions using the designed input space \mathcal{B}_{in} needs to be performed, where the provisional transmission corresponding to Figure 3 with switches OFF means

```

1 Input: Modulation, Possible channel gain for  $\mathbf{h}_U$ ;
2 Output:  $\mathcal{B}_{in}$ ,  $\mathcal{B}_{out}$ ;
3 Initialization:  $\mathcal{B}_{in} = \emptyset$ ,  $\mathcal{B}_{out} = \emptyset$ ;
4 %Generation of input space for database  $\mathcal{B}$ ;
5 foreach  $A_{BS} \in \text{adopted modulation}$  do
6   foreach  $\theta_{BS} \in \text{adopted modulation}$  do
7     foreach designed  $\mathbf{h}_U$  do
8        $\bar{A}_{BS}(b) = A_{BS}$ ,  $\bar{\theta}_{BS}(b) = \theta_{BS}$ ,  $\bar{\mathbf{h}}_U(b) = \mathbf{h}_U$ ;
9        $\mathcal{B}_{in}(b) = [\bar{A}_{BS}(b); \bar{\theta}_{BS}(b); \bar{\mathbf{h}}_U(b)]$ ;
10    end
11  end
12 end
13 foreach  $b$  do
14    $\mathcal{B}_{in} = \mathcal{B}_{in} \cup \mathcal{B}_{in}(b)$ ;
15 end
16 %Generation of output space for database  $\mathcal{B}$ ;
17 foreach  $\mathcal{B}_{in}(b)$  do
18   foreach  $b'$  do
19     Measure  $\mathbf{h}_{BS}$  as  $\tilde{\mathbf{h}}_{BS}$ 
20     Record the results to  $\tilde{\mathbf{h}}_{BS}(b, b')$ ;
21     Measure SI using (23) w/o desired signal input;
22     Record the results to  $\bar{\mathbf{I}}(b, b')$ ;
23      $\mathcal{B}_{out}(b, b') = [\tilde{\mathbf{h}}_{BS}(b, b'); \bar{\mathbf{I}}(b, b')]$ ;
24   end
25 end
26 foreach  $b$  and  $b'$  do
27    $\mathcal{B}_{out} = \mathcal{B}_{out} \cup \mathcal{B}_{out}(b, b')$ 
28 end

```

ALGORITHM 1: Offline database generation.

transmissions without desired signal inputs, i.e., $\mathbf{m}_U(i, j) = 0 \forall i, j$, and DC is not performed at BS. Note that the outputs of channel estimation block in Figure 3 are the estimated CSI of user to BS and the estimated CSI of interantenna of BS, and the value of the former one is replaced by the designed $\bar{\mathbf{h}}_U$ in the offline database generation.

Intuitively, we first chose a designed FV of input space as

$$\mathcal{B}_{in}(b) = [\bar{A}_{BS}(b); \bar{\theta}_{BS}(b); \bar{\mathbf{h}}_U(b)], \quad (22)$$

where $b = 1, \dots, |\mathcal{B}_{in}|$, and then use $\mathcal{B}_{in}(b)$ to perform the provisional transmission following the considered IBFD system. For this case of b , one of the possible channel gain $\bar{\mathbf{h}}_{BS}$ measured by the channel estimation block is recorded into $\tilde{\mathbf{h}}_{BS}(b, b')$, and at the same time, $\bar{\mathbf{I}}(b, b')$ is used to recorded the input of 0/1 decision, i.e., the constellation value of SI. According to (13) and (15), after averaging the noise effects by pilot signals, in mathematics, the recorded value of $\bar{\mathbf{I}}(b, b')$ by measurement can be written as

$$\bar{\mathbf{I}}(b, b') = \bar{\mathbf{I}}(b, b') \Big|_{\lambda = \bar{\mathbf{h}}_U(b), A_{BS}(i, j) = \bar{A}_{BS}(b), \theta_{BS}(i, j) = \bar{\theta}_{BS}(b)}. \quad (23)$$

A summary for this generation process can be found in Algorithm 1.

3.1.2. Online Data Transmission. Once the database \mathcal{B} is created, we can use it into online transmission for the considered IBFD systems as shown in Figure 3 with switches ON. When the BS prepares to demodulate the desired signals $\mathcal{S}(i, j)$ transmitted by the user under the effect of SI $\mathbf{I}(i, j)$ and noise $\mathbf{W}(i, j)$; in the block of searching algorithm for DC, the BS first searches an “optimal” FV from input space \mathcal{B}_{in} using estimated CSI of $\tilde{\mathbf{h}}_U(i)$ and the related modulation information $A_{BS}(i, j)$ and $\theta_{BS}(i, j)$ of the signals the BS sent.

The “optimal” FV is decided by the vector $[A_{BS}(i, j); \theta_{BS}(i, j); \tilde{\mathbf{h}}_U(i)]$ and written as

$$\mathcal{B}_{in}(b^*) = [\bar{A}_{BS}(b^*); \bar{\theta}_{BS}(b^*); \tilde{\mathbf{h}}_U(b^*)], \quad (24)$$

where

$$b^* = \left\{ \arg_{b \in \{1, \dots, |\mathcal{B}_{in}|\}} \bar{A}_{BS}(b) = A_{BS}(i, j) \right\} \cap \left\{ \arg_{b \in \{1, \dots, |\mathcal{B}_{in}|\}} \bar{\theta}_{BS}(b) = \theta_{BS}(i, j) \right\} \cap \left\{ \arg \min_{b \in \{1, \dots, |\mathcal{B}_{in}|\}} \|\bar{\mathbf{h}}_U(b) - \tilde{\mathbf{h}}_U(i)\| \right\}. \quad (25)$$

Based on the fact that $A_{BS}(i, j)$ and $\theta_{BS}(i, j)$ for all of i

and j are known at BS and have been decided by modulation method, the result of (25) is dominated by the difference between $\bar{\mathbf{h}}_{\text{U}}(b)$ and $\bar{\mathbf{h}}_{\text{U}}(i)$. In fact, the last part in (25) is the well-known optimization problem of nearest neighbor search (NNS) in LL. The simplest solution to the NNS problem is the so-called linear search which computes the Euclidean distance taking over all of candidates. Other methods such as space partitioning like k -dimensional tree [34] or greedy search [35] can give an exact or approximate solution at a lower computational cost. Some analysis and performance study for NNS problems can be found in [36].

After the “optimal” FV $\mathcal{B}_{\text{in}}(b^*)$ is found, the corresponding set of the potential SI effects in the output space \mathcal{B}_{out} is then located as $\mathcal{B}_{\text{out}}(b^*, b') = [\bar{\mathbf{h}}_{\text{BS}}(b^*, b'); \bar{\mathbf{I}}(b^*, b')]$. To assign suitable values to the DC $\mathbf{C}_{\text{dig}}(i, j, \alpha)$, we also consider the estimated CSI of $\tilde{\mathbf{h}}_{\text{BS}}(i)$. The value of $\bar{\mathbf{I}}(b^*, b'^*)$ is finally assigned to $\mathbf{C}_{\text{dig}}(i, j, \alpha)$ where b'^* is given by

$$b'^* = \arg \min_{b' \in \{1, \dots, |\mathcal{B}_{\text{out}}(b^*)|\}} \left\| \bar{\mathbf{h}}_{\text{BS}}(b^*, b') - \tilde{\mathbf{h}}_{\text{BS}}(i) \right\|, \quad (26)$$

where $|\mathcal{B}_{\text{out}}(b^*)|$ suggests the number of measured CSI between interantenna of BS. Notably, because effects from noise in $\bar{\mathbf{h}}_{\text{BS}}(b^*, b')$ and $\tilde{\mathbf{h}}_{\text{BS}}(i)$ have been averaged by using multiple pilot signals, the result of (26) ensured that $\bar{\mathbf{h}}_{\text{BS}}(b^*, b') \approx \tilde{\mathbf{h}}_{\text{BS}}(i)$. In addition, our proposed approaches work with an assumption that a channel realization between the BS antennas $\mathbf{h}_{\text{BS}}(i)$ can only be tagged by one estimated CSI $\tilde{\mathbf{h}}_{\text{BS}}(i)$; in other words, there has to be a one-to-one correspondence between $\mathbf{h}_{\text{BS}}(i)$ and $\tilde{\mathbf{h}}_{\text{BS}}(i)$. Generally, this one-to-one correspondence can be guaranteed. Thus, based on the above explanation, the parameter of α in the proposed DC $\mathbf{C}_{\text{dig}}(i, j, \alpha)$ is written as

$$\alpha = [A_{\text{BS}}(i, j); \theta_{\text{BS}}(i, j); \tilde{\mathbf{h}}_{\text{U}}(i); \tilde{\mathbf{h}}_{\text{BS}}(i)]. \quad (27)$$

Notably, according to (23), because the interference term $\bar{\mathbf{I}}(b, b')$ in the output space of database is premeasured using $\lambda = \bar{\mathbf{h}}_{\text{U}}(b)$ in offline database generation, to suppress the SI effects, in online transmission, the parameter λ in the detector process should be set to $\bar{\mathbf{h}}_{\text{U}}(b^*)$ rather than the common $\bar{\mathbf{h}}_{\text{U}}(i)$ in the proposed DC. So, corresponding to the parameters of $A_{\text{BS}}(i, j)$, $\theta_{\text{BS}}(i, j)$, $\mathbf{h}_{\text{U}}(i)$, and $\mathbf{h}_{\text{BS}}(i)$, the constellation value of SI, excluding the noise part after using the proposed method in online transmissions, is written as

$$\mathbf{I}(i, j) = \mathbf{I}(i, j) | \lambda = \bar{\mathbf{h}}_{\text{U}}(b^*), \quad (28)$$

tagged by $\tilde{\mathbf{h}}_{\text{BS}}(i)$, whereas the DC value $\mathbf{C}_{\text{dig}}(i, j, \alpha)$ for this case is decided by database \mathcal{B}_{out} and is written as

$$\begin{aligned} \mathbf{C}_{\text{dig}}(i, j, \alpha) &= \bar{\mathbf{I}}(b^*, b'^*) = \mathbf{I}(i, j) | \lambda = \bar{\mathbf{h}}_{\text{U}}(b^*), A_{\text{BS}}(i, j) \\ &= \bar{A}_{\text{BS}}(b^*), \theta_{\text{BS}}(i, j) = \bar{\theta}_{\text{BS}}(b^*), \end{aligned} \quad (29)$$

1 **Input:** $\mathcal{B}_{\text{in}}, \mathcal{B}_{\text{out}}, A_{\text{BS}}(i, j), \theta_{\text{BS}}(i, j), \tilde{\mathbf{h}}_{\text{U}}(i), \tilde{\mathbf{h}}_{\text{BS}}(i);$
 2 **Output:** $\mathbf{C}_{\text{dig}}(i, j, \alpha), \lambda;$
 3 **Initialization:** $\alpha = [A_{\text{BS}}(i, j); \theta_{\text{BS}}(i, j); \tilde{\mathbf{h}}_{\text{U}}(i); \tilde{\mathbf{h}}_{\text{BS}}(i)];$
 4 **Calculate** b^* using (25);
 5 **Calculate** b'^* using (26);
 6 $\mathbf{C}_{\text{dig}}(i, j, \alpha) = \bar{\mathbf{I}}(b^*, b'^*);$ %For DC blocks;
 7 $\lambda = \bar{\mathbf{h}}_{\text{U}}(b^*);$ %For detection block;

ALGORITHM 2: Value assignment for the proposed DC.

tagged by $\bar{\mathbf{h}}_{\text{BS}}(b^*, b'^*)$. Because (26) ensured that tags $\bar{\mathbf{h}}_{\text{BS}}(b^*, b'^*) \approx \tilde{\mathbf{h}}_{\text{BS}}(i)$ and $A_{\text{BS}}(i, j)$, $\theta_{\text{BS}}(i, j)$ are known at BS; further, the one-to-one correspondence between $\tilde{\mathbf{h}}_{\text{BS}}(i)$ and $\mathbf{h}_{\text{BS}}(i)$ can be generally guaranteed, and the results of (28) and (29) are almost equal. Consequently, the residual SI of nonnoise part $\bar{\mathbf{I}}(i, j)$ can be substantially suppressed. A summary of this process can be found in Algorithm 2.

Moreover, according to (12), the desired signal part in the online transmissions thereby becomes to

$$\mathbf{S}(i, j) = \mathbf{S}(i, j) | \lambda = \bar{\mathbf{h}}_{\text{U}}(b^*), \quad (30)$$

and thus introduces an error caused by $|\bar{\mathbf{h}}_{\text{U}}(b^*) - \tilde{\mathbf{h}}_{\text{U}}(i)|$ in the demodulation process because λ is commonly assigned by $\tilde{\mathbf{h}}_{\text{U}}(i)$. Fortunately, the differences between $\bar{\mathbf{h}}_{\text{U}}(b^*)$ and $\tilde{\mathbf{h}}_{\text{U}}(i)$ decrease with an increase in database size. Naturally, the residual signal-to-interference-plus-noise ratio (RSINR) for the i th frame and the j th symbol after using the proposed DC is written as

$$\begin{aligned} \eta_{\text{pro}}(i, j) &\approx G_{\text{PA,U}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(i) \mathbf{h}_{\text{U}}(i) \mathbf{m}_{\text{U}}(i, j) \right\|^2 \\ &\quad \cdot \left(G_{\text{PA,U}} \left\| \left(\bar{\mathbf{h}}_{\text{U}}^{-1}(b^*) - \tilde{\mathbf{h}}_{\text{U}}^{-1}(i) \right) \times \mathbf{h}_{\text{U}}(i) \mathbf{m}_{\text{U}}(i, j) \right\|^2 \right. \\ &\quad + C_{\text{ana}} \bar{C}_{\text{ant}} \left(G_{\text{PA,BS}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{h}_{\text{BS}}(i) \mathbf{w}_{\text{tx,BS}}(i, j) \right\|^2 \right. \\ &\quad \left. \left. + \left\| \bar{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{h}_{\text{BS}}(i) \mathbf{w}_{\text{PA,BS}}(i, j) \right\|^2 \right) + \left\| \bar{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{w}_{\text{rx}}(i, j) \right\|^2 \right)^{-1}, \end{aligned} \quad (31)$$

where we considered the results of (28) and (29) as approximate, assumed gains of all of antennas are 0 dBi, all of PAs are linear, and ignored noise effects from the interested user for simplicity of expressions.

The expression of (31) suggests that the SI effect has been suppressed to the first and second parts in the denominator of (31) by the proposed DC. The first part has to be left behind if $\bar{\mathbf{h}}_{\text{U}}(b^*) \neq \tilde{\mathbf{h}}_{\text{U}}(i)$ and results to a limitation in RSINR for high desired signal power. The second part caused by the noises in BS's modulation and PA can only be decreased depending on the antenna and analog cancellations (AACs). Our simulation results can provide some evidence for the analysis.

3.2. EL-Based Cancellation Approach. As a comparison object to evaluate our proposed approach, the EL-based cancellation approach should be taken into consideration. The EL used approach needs to train a model using a prepared database, such as the one generated following the description in Section 3.1.1, to predict the constellation value of SI in the 0/1 decision process. Assuming that the generation of database \mathcal{B} is complete, we can construct a training set using parameters in \mathcal{B} to train the model.

The training set includes a part of FVs \mathcal{V} and a part of target values \mathcal{T} which are designed by

$$\mathcal{V} = \left\{ \left[\bar{A}_{\text{BS}}(z); \bar{\theta}_{\text{BS}}(z); \bar{\mathbf{h}}_{\text{U}}(z); \bar{\mathbf{h}}_{\text{BS}}(z) \right] \right\}_{z=1}^{|\mathcal{B}_{\text{in}}|} \sum_{b=1}^{|\mathcal{B}_{\text{out}}(b)|}, \quad (32)$$

$$\mathcal{T} = \left\{ \bar{\mathbf{I}}(z) \right\}_{z=1}^{|\mathcal{B}_{\text{in}}|} \sum_{b=1}^{|\mathcal{B}_{\text{out}}(b)|}, \quad (33)$$

respectively. Afterward, a suitable model (or function) \mathbb{M}^* can be trained by substituting \mathcal{V} and \mathcal{T} into a given network architecture and performing learning algorithms. Without loss of generality and noting that \mathcal{T} are numerical output, we employ deep NN architecture and adopt a regression learning algorithm where the former is one of the most popular networks in EL methods, and the latter is used to train the model \mathbb{M}^* [37]. Mathematically, \mathbb{M}^* can be written as

$$\mathbb{M}^* = \arg \min_{\mathbb{M}} \mathcal{L}(\mathcal{T} - \mathbb{M}(\mathcal{V})), \quad (34)$$

where $\mathcal{L}(z)$ denotes loss function for training process in EL methods.

In fact, the trained model \mathbb{M}^* is used to estimate constellation values of SI in 0/1 decision, and obviously, the outputs of trained model should be assigned to DC for SI suppression. Therefore, the DC $\mathbf{C}_{\text{dig}}(i, j, \mathbf{a})$ is calculated by

$$\mathbf{C}_{\text{dig}}(i, j, \mathbf{a}) = \mathbb{M}^*(\mathbf{a}), \quad (35)$$

where the same \mathbf{a} in (27) in the proposed LL-based cancellation approach is used.

According to (35), the introduction of error between the real and predicted SI in the 0/1 decision cannot be avoided since both real CSI of user-BS and interantenna of BS cannot be acquired in practice, and there are always errors in the model training process [26]. Moreover, as shown in

$$\begin{aligned} \eta_{\text{EL}}(i, j) = & G_{\text{PA,U}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(i) \mathbf{h}_{\text{U}}(i) \mathbf{m}_{\text{U}}(i, j) \right\|^2 \\ & \cdot \left(\left\| \mathbf{C}_{\text{ana}}^{1/2} \hat{\mathbf{C}}_{\text{ant}}^{1/2} G_{\text{PA,BS}}^{-1/2} \tilde{\mathbf{h}}_{\text{U}}^{-1}(i) \mathbf{h}_{\text{BS}}(i) \mathbf{m}_{\text{BS}}(i, j) - \mathbb{M}^* \right\|^2 \right. \\ & \cdot \left(\left[A_{\text{BS}}(i, j); \theta_{\text{BS}}(i, j); \tilde{\mathbf{h}}_{\text{U}}(i); \tilde{\mathbf{h}}_{\text{BS}}(i) \right] \right\|^2 \\ & + C_{\text{ana}} \hat{\mathbf{C}}_{\text{ant}} \left(G_{\text{PA,BS}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(b^*) \times \mathbf{h}_{\text{BS}}(i) \mathbf{w}_{\text{tx,BS}}(i, j) \right\|^2 \right. \\ & \left. \left. + \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{h}_{\text{BS}}(i) \mathbf{w}_{\text{PA,BS}}(i, j) \right\|^2 \right) + \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{w}_{\text{rx}}(i, j) \right\|^2 \right)^{-1}, \end{aligned} \quad (36)$$

for RSINR of the EL methods with the same conditions as the proposed method, suppose the residual SI, i.e., the first component in the denominator, still significantly dominates the composite signal power because of amplification effect from $G_{\text{PA,BS}}$. In that case, the demodulation for the desired signals is not optimistic. In the last section, we perform more simulations to evaluate this cancellation approach.

3.3. Regular Cancellation Approach. Since pilot signals are commonly used in each frames sent by users and BS in modern wireless communication system; the real CSI between the interested user and BS $\mathbf{h}_{\text{U}}(i)$ and the real CSI between the interantenna of BS $\mathbf{h}_{\text{BS}}(i)$ are estimated by $\tilde{\mathbf{h}}_{\text{U}}(i)$ and $\tilde{\mathbf{h}}_{\text{BS}}(i)$, respectively. According to the SI expression in (13), an estimation-based regular method for assignment of DC $\mathbf{C}_{\text{dig}}(i, j, \mathbf{a})$ is directly calculating the constellation value of SI $\mathbf{I}(i, j)$ using the estimated CSI $\tilde{\mathbf{h}}_{\text{U}}(i)$ and $\tilde{\mathbf{h}}_{\text{BS}}(i)$. Mathematically, $\mathbf{C}_{\text{dig}}(i, j, \mathbf{a})$ in this regular cancellation approach is assigned as

$$\mathbf{C}_{\text{dig}}(i, j, \mathbf{a}) = \mathbf{I}(i, j) \lambda = \tilde{\mathbf{h}}_{\text{U}}(i), \mathbf{h}_{\text{BS}}(i) = \tilde{\mathbf{h}}_{\text{BS}}(i), \quad (37)$$

where \mathbf{a} is decided by (27).

The RSINR of the regular approach can be expressed as

$$\begin{aligned} \eta_{\text{reg}}(i, j) = & G_{\text{PA,U}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(i) \mathbf{h}_{\text{U}}(i) \mathbf{m}_{\text{U}}(i, j) \right\|^2 \\ & \cdot \left(C_{\text{ana}} \hat{\mathbf{C}}_{\text{ant}} G_{\text{PA,BS}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(i) (\mathbf{h}_{\text{BS}}(i) - \tilde{\mathbf{h}}_{\text{BS}}(i)) \mathbf{m}_{\text{BS}}(i, j) \right\|^2 \right. \\ & + C_{\text{ana}} \hat{\mathbf{C}}_{\text{ant}} \left(G_{\text{PA,BS}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{h}_{\text{BS}}(i) \mathbf{w}_{\text{tx,BS}}(i, j) \right\|^2 \right. \\ & \left. \left. + \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{h}_{\text{BS}}(i) \mathbf{w}_{\text{PA,BS}}(i, j) \right\|^2 \right) + \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(b^*) \mathbf{w}_{\text{rx}}(i, j) \right\|^2 \right)^{-1}, \end{aligned} \quad (38)$$

based on the same conditions as the proposed approach. This expression indicates that the SI effect, i.e., the first part in denominator, cannot be completely removed because an error of $\mathbf{h}_{\text{BS}}(i) - \tilde{\mathbf{h}}_{\text{BS}}(i) \neq 0$ has to be introduced, and this error is further multiplied by $G_{\text{PA,BS}} \left\| \tilde{\mathbf{h}}_{\text{U}}^{-1}(i) \right\|^2$ times. This fact suggests that, basically, only an excellent performance for estimation (e.g., estimation of wireless channel and PA nonlinearity) and good channel condition for the desired signal (i.e., bigger $\mathbf{h}_{\text{U}}(i)$) can make sure that IBFD systems

with the regular DC work. Surely, we perform more simulations to evaluate this cancellation approach.

4. Channel Model

Without loss of generality and for simplicity, in this study, we consider composite fading channels with path loss and Rayleigh fading to simulate the attenuation between the interested user and BS, i.e., $\mathbf{h}_U(i) = \rho_U(i)e^{j\phi_U(i)}\forall i$. The phase shift $\phi_U(i)\forall i$ is modeled as independent and identically distributed (i.i.d.) random variable (RV) and follows uniform distribution between 0 and 2π radians. The amplitude gain $\rho_U(i)$ for i th frame can be expressed as

$$\rho_U(i) = \frac{c_o}{4\pi f_0} \sqrt{d_U^{-\zeta_U}(i)} r_U(i), \quad (39)$$

where c_o is the velocity of light, $d_U(i)$ denotes the distance between the interested user and BS for the transmitting of i th frame, and ζ_U is the path loss exponent around user [38, 39]. The term of $r_U(i)\forall i$ is also modeled as i.i.d. RV and follows the Rayleigh distribution with the same cumulative distribution function (CDF) expressed as

$$\text{CDF}_{r_U(i)}(z) = 1 - \exp\left(-\frac{z^2}{2\sigma^2}\right), \forall i, \quad (40)$$

where σ is the scale parameter of the Rayleigh distribution [32].

For the modeling of channel attenuation between the BS's antennas $\mathbf{h}_{BS}(i)$, considering two facts of not complicated surroundings and low building density around BS, further, since a centralized antenna deployment with a limited interantenna distance [40] is usually employed on the traditional BS, the channel varying between transceiver antennas can be assumed to be static. Based on the above assumption and description, the channel gains of $\mathbf{h}_{BS}(i)$ for all of frames are thus modeled by the path loss model and expressed as

$$\mathbf{h}_{BS}(i) = \frac{c_o}{4\pi f_0} \sqrt{d_{BS}^{-\zeta_{BS}}} \forall i, \quad (41)$$

where d_{BS} is the distance between the BS's antennas and ζ_{BS} denotes the path loss exponent around BS.

5. Computer Simulations

In this section, we evaluate and analyze the proposed LL-based cancellation approach for the IBFD systems with the help of computer simulations. To do that, we first explain how to generate the offline database based on the designed format, which is introduced in Section 3.1. Then, the use of the generated database for online transmission in the considered IBFD systems is described. At last, we present and analyze simulated results with comparisons of the EL-based cancellation approach described in Section 3.2 and the regular cancellation approach described in Section 3.3 and finally summarize the key findings.

5.1. Offline Database Generation for Simulations. In simulations, since quadrature phase shift keying (QPSK) modulations are planned to be used into both user and BS, the dimension for each FV of input space \mathcal{B}_{in} in database \mathcal{B} can be reduced to $\bar{\theta}_{BS}(b)$ and $\mathbf{h}_U(b) = \bar{\rho}_U(b)e^{j\bar{\phi}_U(b)}$, where $\bar{\theta}_{BS}(b)$ can be assigned as four possible values in the common QPSK modulation and is written as

$$\bar{\theta}_{BS}(b) \in \left\{ \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4} \right\}. \quad (42)$$

The term of $\bar{\rho}_U(b)$ is the amplitude gain of the channel between the user and BS, and it can range from 0 to ∞ caused by the Rayleigh fading. In the simulations, a variable range for $\bar{\rho}_U(b)$ is assumed as $[\bar{\rho}_{min}, \bar{\rho}_{max}]$, and it is calculated by

$$\bar{\rho}_{min} = \frac{c_o}{4\pi f_0} \sqrt{-2\sigma^2 \ln(1-\delta)} d_U^{-\zeta_U}, \quad (43)$$

$$\bar{\rho}_{max} = \frac{c_o}{4\pi f_0} \sqrt{-2\sigma^2 \ln \delta} d_U^{-\zeta_U}, \quad (44)$$

according to Rayleigh CDF, where δ denotes a given probability in the CDF of Rayleigh distribution and can range from 0 to 1. Thereafter, we define Δ_{am} as the resolution of amplitude coefficient $\bar{\rho}_U(b)$, and thus, there are $\lceil (\bar{\rho}_{max} - \bar{\rho}_{min})/\Delta_{am} \rceil$ possible values to be assigned, and

$$\bar{\rho}_U(b) \in \left\{ \bar{\rho}_{min}, \bar{\rho}_{min} + \Delta_{am}, \bar{\rho}_{min} + 2\Delta_{am}, \dots, \bar{\rho}_{min} + \left(\left\lceil \frac{\bar{\rho}_{max} - \bar{\rho}_{min}}{\Delta_{am}} \right\rceil - 1 \right) \Delta_{am} \right\}. \quad (45)$$

The term of $\bar{\phi}_U(b)$ is the phase shift caused by the Rayleigh fading, and it can range from 0 to 2π radians. For the database generation, we define Δ_{ph} as the resolution of phase shift $\bar{\phi}_U(b)$, and hence, there are $\lceil 2\pi/\Delta_{ph} \rceil$ possible values to be assigned. Mathematically, $\bar{\phi}_U(b)$ can be written as

$$\bar{\phi}_U(b) \in \left\{ 0, \Delta_{ph}, 2\Delta_{ph}, \dots, \left(\left\lceil \frac{2\pi}{\Delta_{ph}} \right\rceil - 1 \right) \Delta_{ph} \right\}. \quad (46)$$

Based on the above configurations, an input space \mathcal{B}_{in} with

$$|\mathcal{B}_{in}| = 4 \times \left\lceil \frac{\bar{\rho}_{max} - \bar{\rho}_{min}}{\Delta_{am}} \right\rceil \times \left\lceil \frac{2\pi}{\Delta_{ph}} \right\rceil, \quad (47)$$

unduplicated FVs can be generated for our simulations. For measurement and record of the output space \mathcal{B}_{out} , based on our channel assumptions, the output space \mathcal{B}_{out} is, thus, reduced to $\bar{\mathbf{I}}$, and can create a one-to-one correspondence with input space \mathcal{B}_{in} . The size of \mathcal{B}_{in} as well as \mathcal{B}_{out} becomes larger with decreases of Δ_{am} , Δ_{ph} , and δ , and thus, more storage for saving and higher compute capability for searching are needed. Certainly, the parameters of $\bar{\rho}_{min}$,

$\bar{\rho}_{\max}$, Δ_{am} , and Δ_{ph} are not restricted by the considered channel models. Actually, these parameters can be set freely and independently to fit various hard- or software environments. In addition, in the database generation process, we generate the same database $N_{\mathcal{B}}$ times and average them to eliminate the effects caused by the various noises as much as possible.

5.2. Model Training for EL-Based Cancellation Approach. According to Section 3.2, once the above database is generated, we can format a training set which consists of FV

$$\mathcal{V} = \left\{ \left[\bar{\theta}_{\text{BS}}(z); \bar{\mathbf{h}}_{\text{U}}(z) \right] \right\}_{z=1}^{|\mathcal{B}_{\text{in}}|} \sum_{b=1}^{|\mathcal{B}_{\text{out}}|} |\mathcal{B}_{\text{out}}(b)|, \quad (48)$$

and the corresponding target value

$$\mathcal{T} = \left\{ \bar{\mathbf{I}}(z) \right\}_{z=1}^{|\mathcal{B}_{\text{in}}|} \sum_{b=1}^{|\mathcal{B}_{\text{out}}|} |\mathcal{B}_{\text{out}}(b)|, \quad (49)$$

for model training in the EL-based cancellation approach. A deep NN architecture is employed to train model \mathbb{M}^* and is shown in Figure 4 in which the input layer and the output layer are used to accept the FV \mathcal{V} and target value \mathcal{T} , respectively. A Levenberg-Marquardt back-propagation algorithm [41] is used to train node coefficients of all of layers in our feedforward NNs, and an introduction study of deep learning for wireless physical layer can be found in [37]. With a consideration of limited computational cost, we configure a maximum value of epoch for model training in which an epoch can be described as one complete cycle through the entire training dataset. Some essential parameters for model training in our current simulations are listed in Table 1.

5.3. Online Data Transmission for Simulations. Considering 4.6~4.9 GHz is assigned to a 5G-based private network called local 5G in Japan [11], we employ 4.6 GHz as carrier frequency in our simulations. After the database \mathcal{B} and model \mathbb{M}^* are generated, they can be used in online transmissions for the considered IBFD systems. When the desired symbol is received under the effects of SI and noises, the BS first attenuates the received composite signal waveform excluding the desired portion using the AACs C_{ant} and C_{ana} . Then, the searching algorithm is implemented for the proposed cancellation approach following the instructions described in Section 3.1.2 with the help of the generated database \mathcal{B} . The SI can also be estimated through the trained model \mathbb{M}^* for the EL-based cancellation approach, or the estimated CSI can predict the SI for the regular cancellation approach. Subsequently, constellation values of SI found from the database \mathcal{B} , predicted by the model \mathbb{M}^* , or estimated by the regular approach are taken out and subtracted independently before the 0/1 decisions in BS's demodulator. Finally, binary data sent by the user is recovered, and some transmission performances are evaluated and analyzed. Simulation parameters are listed in Table 2.

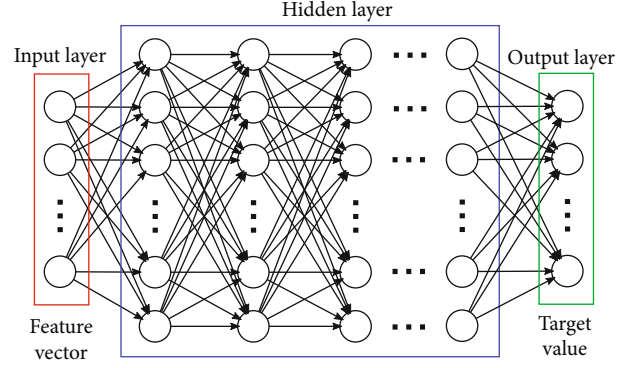


FIGURE 4: An illustration of fully connected deep NN architecture employed in EL-based cancellation approach for the IBFD systems.

5.4. Simulation Results. In this subsection, we first list three factors which may affect the performances of the above-mentioned DC approaches. They are (1) PA additional noise at BS side, (2) errors in channel estimation, and (3) database size. The additional noise is assumed as generated noise in all PAs, and its existence results noise figures are not one. The errors in channel estimation process occurred if the BS cannot obtain perfect CSI. In our simulations, the estimated CSI $\tilde{\mathbf{h}}_{\Psi}(i)$ for $\Psi = \text{U}$ or $\Psi = \text{BS}$ is generated by

$$\tilde{\mathbf{h}}_{\Psi}(i) = \mathbf{h}_{\Psi}(i) + \Delta_{\text{err},\Psi} |\mathbf{h}_{\Psi}(i)| e^{j\tau}, \quad (50)$$

where $\mathbf{h}_{\Psi}(i)$ denotes the real CSI of the interested user to BS or interantenna at BS side. $\Delta_{\text{err},\Psi}$ is defined as channel estimation error and can range from 0 to 1. τ is a random phase ranged from 0 to 2π and follows uniform distribution. The database size can be calculated by (47). Thereafter, we demonstrate simulation results under the effects of these factors, respectively. Finally, we evaluate transmission performances with all of the mentioned factors for the different DCs and show some key findings.

Notably, in our current systems, although all of the signals passing through the PAs are somewhat distorted owing to the PA nonlinearity, estimation errors on PA nonlinearity were not considered in this study. Inaccurate estimation of the PA nonlinearity may substantially degrade transmission performances for the estimation-based regular cancellation approach. However, considering the time-invariant property in all PAs, it seems not to be a major issue for the learning- and EL-based cancellation approaches. Nevertheless, our future work may concentrate on estimating PA nonlinearity.

In Figure 5, we demonstrate some comparisons of BER simulated by no SI, the proposed, the regular, and the EL-based DC approaches with variable AACs representing by $C_{\text{ant}}C_{\text{ana}}$ and fixed average receive signal-to-noise ratio (SNR). The average receive SNR is defined as the ratio of the average power of the desired signal to the average power of noise at the receiver antenna and is mainly dominated by the large-scale fading between the interested user and BS. The channel between the interested user and BS is assumed to be static to evaluate the effect caused by the factor (1), and channel estimation errors for all channels are ignored.

TABLE 1: Parameters for model training.

Parameters	Values	Parameters	Values
Number of hidden layers	3	Loss functions	Mean squared error
Number of hidden layer nodes	16, 8, 4	Train algorithm	Levenberg-Marquardt
Activation functions	Sigmoid (i.e., $(1 + e^{-z})^{-1}$)	Max epoch	1000

TABLE 2: Simulation parameters.

Parameters	Values	Parameters	Values
Num. of frames, N_{frame}	1000	Carrier frequency, f_0	4.6 GHz
Num. of generated databases, $N_{\mathcal{B}}$	20	Modulation method for BS and user	QPSK
Num. of symbols in data part, N_{symbol}	8	Energy per symbol, E_{symbol}	2×10^{-6} mJ
Num. of symbols in pilot part, N_{pilot}	4	Symbol rate, f_{symbol}	5×10^6 /s
Antenna gain at BS receive side, $G_{\text{ant,rx}}$	0 dBi	Antenna cancellations, C_{ant}	Variable
Antenna gain at BS transmit side, $G_{\text{ant,tx,BS}}$	0 dBi	Analog cancellations, C_{ana}	Variable
Antenna gain at user transmit side, $G_{\text{ant,tx,U}}$	0 dBi	PA gain for user, $G_{\text{PA,U}}$	13 dB
PA gain for BS, $G_{\text{PA,BS}}$	30 dB	PA OIP3 for user, $O_{3,U}$	18 dBm
PA OIP3 for BS, $O_{3,BS}$	35 dBm	Noise factor in user's PA, F_U	2.5
Noise factor in BS's PA, F_{BS}	2.5	Inter-antenna distance at BS, d_{BS}	0.2 m
Detection method, $\mathcal{D}(\mathbf{z})$	MF	Path loss exponent around BS, ζ_{BS}	2
Distance of user-BS, $d_U(i)\forall i$	Variable	Phase shift resolution, Δ_{ph}	Variable
Path loss exponent around user, ζ_U	4	Amplitude resolution, Δ_{am}	$\bar{\rho}_{\text{min}}/10^2$
Probability in the CDF of Rayleigh distribution, δ	0.01	Noise power at BS receive antenna, Ω_{rx}	-100 dBm
Scale parameter of Rayleigh distribution, σ	$1/\sqrt{2}$	Noise power at user's modulator, $\Omega_{\text{tx,U}}$	-100 dBm
Noise power at BS's modulator, $\Omega_{\text{tx,BS}}$	-100 dBm	Estimation error for user-BS, $\Delta_{\text{err,U}}$	Variable
Estimation error for BS's antennas, $\Delta_{\text{err,BS}}$	Variable		

This figure indicates that the DC mentioned above approaches achieved similar BER performances over the entire range of $C_{\text{ant}}C_{\text{ana}}$ for a given SNR. The results clarify that factor (1) is not the major reason for the difference among the proposed, the regular, and the EL-based DC approaches. Moreover, there are gaps between BER performances of no SI and that of using DCs when $C_{\text{ant}}C_{\text{ana}} > -50$ dB. This phenomenon occurs because various noises at the BS side generated by PA and modulator are added into SI signals. Because these noises are random and there is nothing to do about them by the DC approaches, AACs had to be used to limit their power strength that can be confirmed by the second term in the denominator in (31), (36), and (38). Consequently, using a larger AACs with $C_{\text{ant}}C_{\text{ana}} \leq -50$ dB decreased the effects from the noises and resulted the similar BER performances as that of no SI effect. In fact, numerous authors have reported that they achieved more than 45 dB SI suppression capability using antenna cancellation (i.e., $C_{\text{ant}} < -45$ dB) [5, 13] and more than 30 dB suppression capability using analog cancellation (i.e., $C_{\text{ana}} < -30$ dB) [6, 11]. Considering feasibility, in our further simulations, we refer to the results of previous works and conservatively set $C_{\text{ant}} = -45$ dB and $C_{\text{ana}} = -15$ dB to suppress the effects of these noises.

In Figure 6(a), we present comparisons of RSINR versus SNR to demonstrate how factor (2) affects the mentioned DC approaches. The comparisons were obtained employing the proposed, regular, and EL-based DC approaches with variable channel estimation errors, assumption of static channels, and 60 dB of AACs, which suppressed the effect from factor (1). The BER performances of the mentioned DC approaches corresponding to Figure 6(a), with the same simulation conditions, are shown in Figure 6(b) to further verify the effect caused by the factor (2).

Figure 6(a) indicates that the RSINR of the regular and EL-based DCs is significantly decreased with increasing of channel estimation error from 10^{-5} to 10^{-3} over the entire range of SNR; on the contrary, the RSINR of the proposed approach is not affected by the channel estimation error. One of the major reasons for this result is whether the estimated CSI $\hat{\mathbf{h}}_{\text{BS}}(i)$ is directly adopted in DC approaches or not. For the proposed DC, $\hat{\mathbf{h}}_{\text{BS}}(i)$ works as a tag of SI and is used to help record quantized SI effect into database in off-line phase or is used to help search target SI value from database in online phase. In other words, $\hat{\mathbf{h}}_{\text{BS}}(i)$, actually, is not designed to participate in any SI estimation-related computations in the proposal. Naturally, it does not appear in the

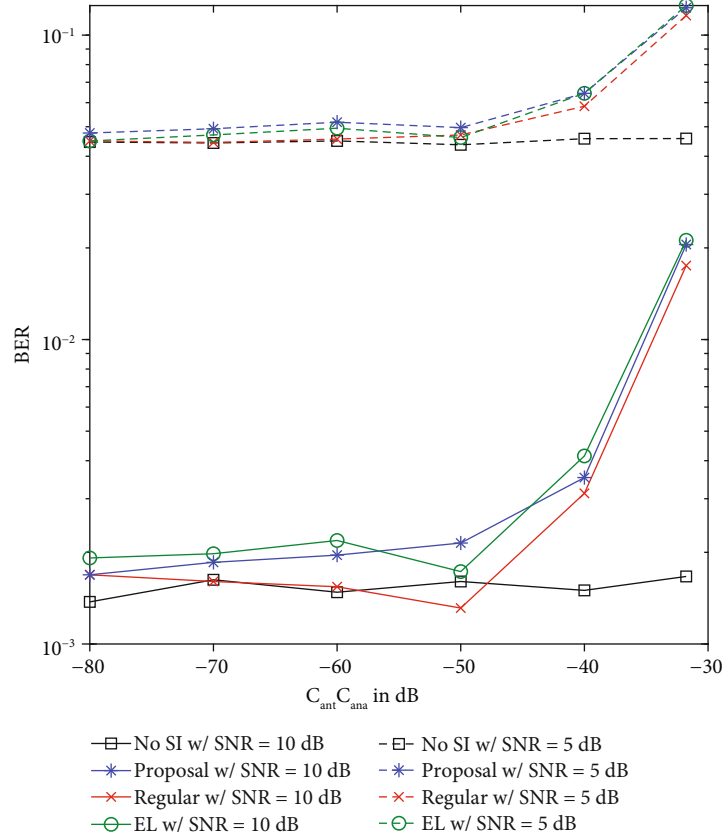


FIGURE 5: Analysis for factor (1): comparisons of BER simulated by no SI, the proposed, the regular, and the EL-based DC approaches with variable $C_{\text{ant}} C_{\text{ana}}$ and fixed average receive SNR. Channels are assumed to be static, and channel estimation errors are ignored.

expression of RSINR (31) and hardly affects the transmission performances such as BER. This also highlights the difference from other DC approaches. Conversely, in the other two DC approaches, $\tilde{h}_{\text{BS}}(i)$ is a necessary variable to calculate the estimated value of SI and thereby directly affects the RSINR values in expressions (36) and (38). In practice, 10^{-3} of channel estimation error is a typical setting and acceptable. However, the RSINR of the regular and EL-based DCs with 10^{-3} estimation error cannot meet the general communication needs, and consequently, their BER performance shown in Figure 6(b) is terrible, unless improving measurement accuracy on channel estimation process to 10^{-5} .

Figures 7(a) and 7(b) exhibit RSINR and BER versus SNR plots simulated by no SI and the three DCs mentioned above under the effects of factor (3) database size. Rayleigh fading is considered in the channel between the interested user and BS to evaluate the effect caused by the factor (3), and perfect channel estimation and 60 dB AACs are assumed. Facing the limited computational cost and storage, we only vary the resolution of phase shift Δ_{ph} to control the database size. Note that since the regular DC approach does not require a database, there is only one curve for the regular approach in figures.

Both Figures 7(a) and 7(b) demonstrate that transmission performance (RSINR and BER) when using the proposed approach and EL-based DC approaches can be improved by increasing the database size. The results can

be well explained by the big data technology because more potential SI values can be stored in a larger database and thus the probability of finding out an estimated SI value that is more closer to the real value of SI becomes higher. Interestingly, for a given database size, the RSINR of the proposed approach gradually approaches a fixed value with increasing SNR and finally results in a floor in BER. In fact, (31) also indicated that in a large SNR range, for example, a large $G_{\text{PA,U}}$, RSINR of the proposed approach is not linear growth with SNR unless $\tilde{h}_{\text{U}}(b^*) = \tilde{h}_{\text{U}}(i)$. However, the RSINR and BER performances are much better than the EL-based DC approach because the latter introduced an unwanted error caused by the model training, and this error was amplified by BS's PA according to the first term in the denominator in (36). In addition, the estimation-based regular approach shows better performances than the proposed approach and the EL-based DCs because channel estimation errors, i.e., factor (2), in these simulations are not considered.

Finally, to evaluate a mixed effect of factors (1), (2), and (3), in Figures 8(a) and 8(b), we present some comparisons of RSINR and BER simulated by the proposed, the regular, and the EL-based DC approaches over fading channel with a common channel estimation error $\Delta_{\text{err,BS}} = \Delta_{\text{err,U}} = 10^{-3}$, a given database size $\Delta_{\text{ph}} = 0.1\pi$, and 60 dB of AACs. Under the consideration of the mixed effects of factors (1), (2), and (3), in Figure 8(a), for the case of SNR = 30 dB (i.e., a

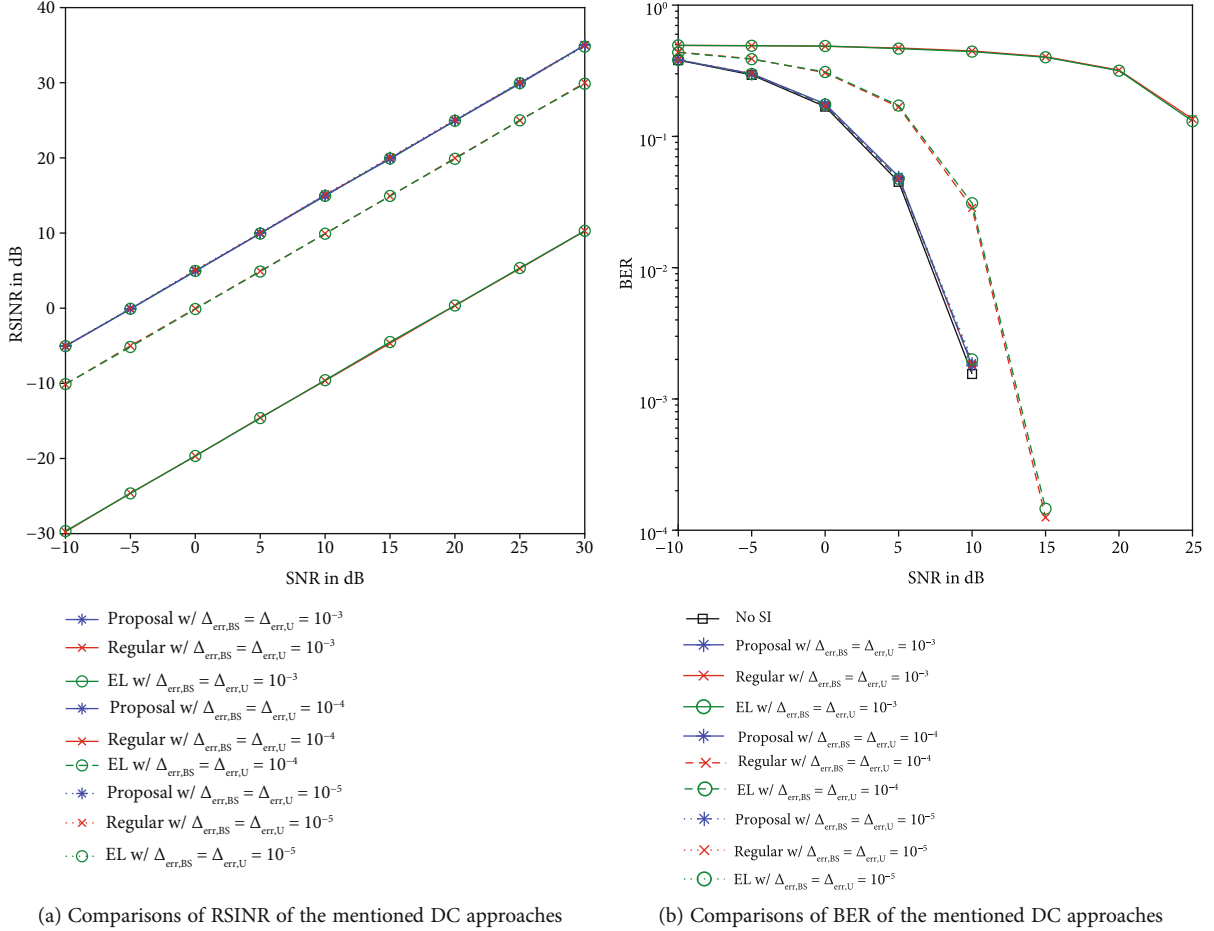


FIGURE 6: Analysis for factor (2) with variable channel estimation errors, assumption of static channels, and 60 dB of AACs.

desired signal power of -70 dBm compared to noise power of -100 dBm), the proposed approach with $\Delta_{ph} = 0.1\pi$ approximately improved the RSINR to 24 dB with -94 dBm of residual SI power. Considering 40 dBm of EIRP at the BS, the figure clearly indicates that about 134 dB of SI suppression capability is achieved by the proposed approach with 60 dB of AACs, and Figure 8(b) thereby claims that our proposed approach can enable IBFD operations in the considered wireless communication systems and, consequently, can double the SE approximately.

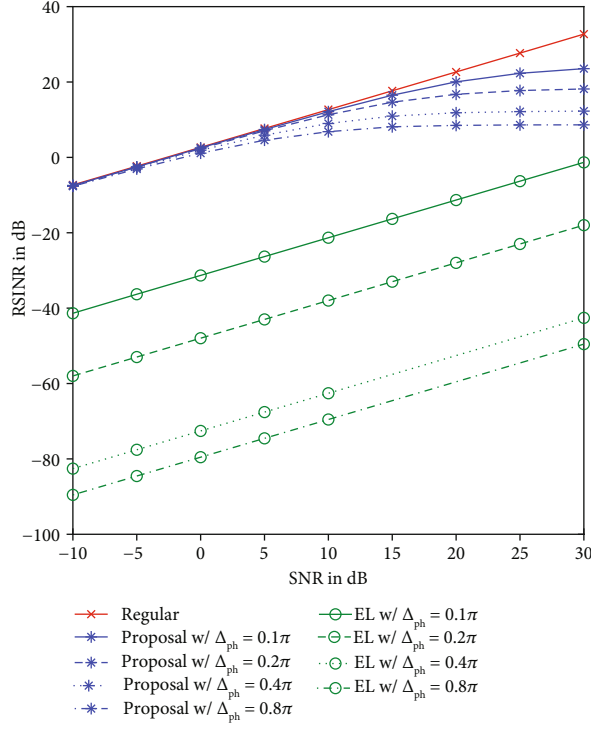
Conversely, there are considerable RSINR and BER gaps between no SI and the regular or EL-based DC approach. The main reasons may be rough channel estimation and insufficient number of hidden layers and nodes in the network architecture. In fact, enhanced transmission performance for the regular DC can be achieved through strong AACs with highly accurate and comprehensive estimation operations. Similarly, applying strong AACs with highly extensive databases or nodes and layers in the NN architecture may improve EL-based DC performance. Therefore, although the performances of the regular and EL-based DC are inadequate under the current simulation setting, we do not deny the two potential DC approaches for IBFD systems. Instead, we strongly encourage exploring them, especially, for the EL-based DC approach, by, for example, using more

faster learning algorithms, reducing dimension of the database, and upgrading hardware devices. Because the proposed LL-based DC approach for IBFD wireless systems can also benefit a lot from technological advance in machine learning-related research fields. Certainly, any type of DC approach should be combined with the outstanding AAC techniques; thereby, improving the existing AACs and solving related issues for the above-mentioned three DC approaches should be studied.

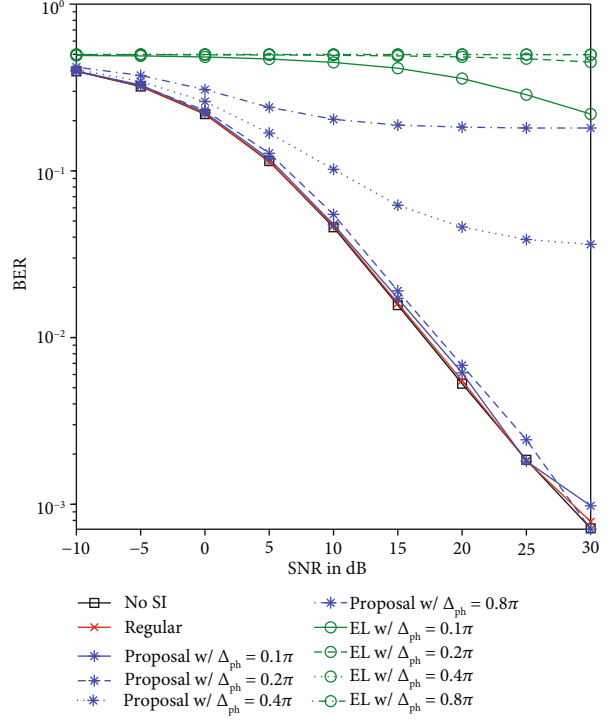
6. Discussion

Although our simulation results for SI cancellation are promising, the proposed LL-based DC approach is still needs to be evaluated using more hard- and/or software environments for the practical applications, and there are several works that need to be addressed.

For example, the proposed approach in the current simulations is compatible with a signal carrier transmission system. For the popular orthogonal frequency division multiplexing (OFDM) which is adopted in most modern communications standards, multiple databases should be independently generated for each subcarrier because of frequency selective characteristics in wireless channels. Since creating and searching operations on a large number of

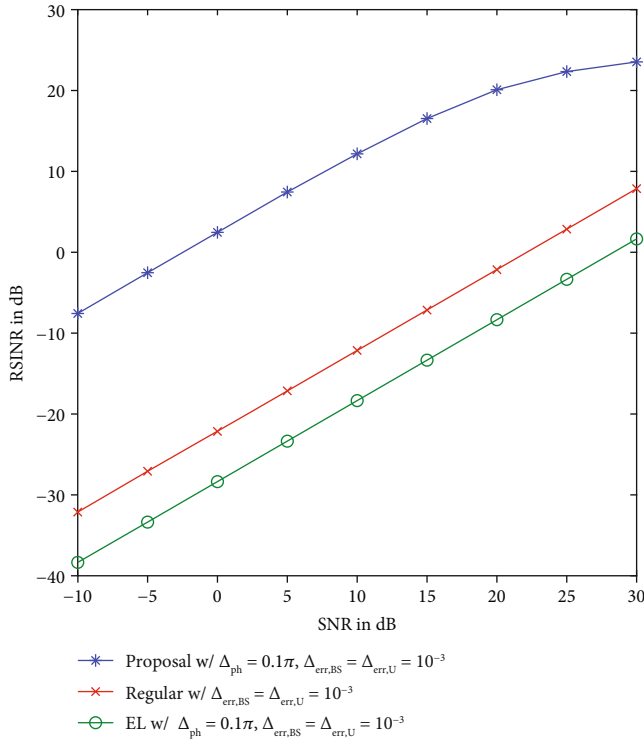


(a) Comparisons of RSINR of the mentioned DC approaches

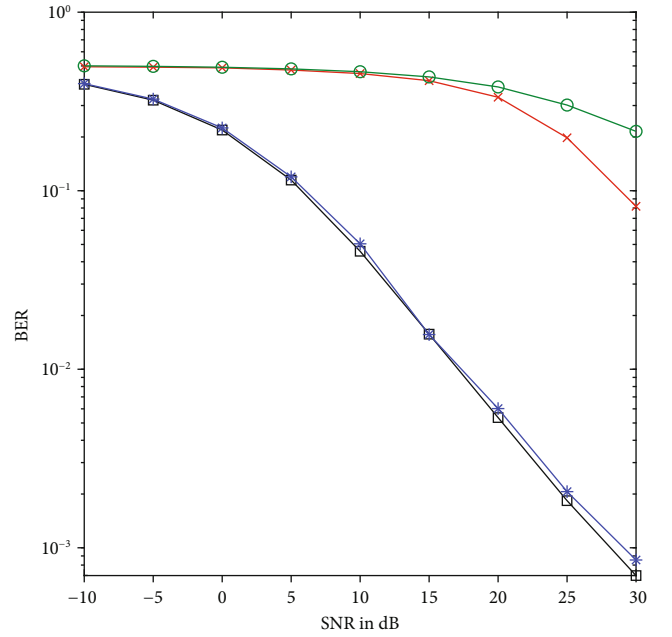


(b) Comparisons of BER of the mentioned DC approaches

FIGURE 7: Analysis for factor (3) over fading channel with perfect channel estimation and 60 dB of AACs.



(a) Comparisons of RSINR of the mentioned DC approaches



(b) Comparisons of BER of the mentioned DC approaches

FIGURE 8: Analysis for mixed effect of factors (1), (2), and (3) over fading channel with a given channel estimation error, a given database size, and 60 dB of AACs.

databases must be considered, a lots of challenges on learning algorithm and calculation ability still need to be overcome before widespread use of the proposal.

Moreover, it seems straightforward to extend our proposal to the high-order constellations such as 256 quadrature amplitude modulation (QAM) by completely using the designed amplitude term \bar{A}_{BS} and phase term $\bar{\theta}_{BS}$ in database (actually, \bar{A}_{BS} in the database is not used for the current QPSK modulation-based simulations). However, more advanced modulation means more combinations of amplitude and phase. Hence, a very large database also introduced a huge challenges on learning algorithm and calculation ability in this case.

Finally, in the present simulations, we assumed static channel between transceiver antennas on BS side. In fact, in a real system, the channel is time-varying. Similarly, the proposed approach can work for this situation by activating \bar{h}_{BS} in the designed database (actually, \bar{h}_{BS} is not used in the current simulations because of static channel). However, a large database caused by the dimension increase is still one of the unavoidable problems, and it is unclear how the complexity will increase in this case.

Based on the above description, the biggest challenge on the learning-based proposal is how to handle the huge amounts of data generated from the systems with OFDM operations, high-order constellations, and complex propagation environment. In the future, one of our major works is to solve this problem by introducing more advanced machine learning-related algorithms and technologies. Furthermore, we focus on conducting multiple evaluations for the proposal based on hard- and software environments, for example, using a SDR platform.

7. Conclusion

In this study, we proposed an LL-based DC approach for SI suppression and enabled IBFD transmissions to improve SE for current wireless communication systems. An offline and online phases for database generation and data transmission, respectively, are performed separately. In the offline phase, the output before the O/1 decision is previously measured without the desired signal input and is recorded to a database with self-defined FV. In the online phase, a suitable result is located from the generated database with the help of the learning method and the FV usage for the same system architecture with desired signal input; the result is then assigned as the value of SI cancellation. An estimation-based regular and an EL-based DC approach are employed to simulate the transmission performance and evaluate the proposed approach. The simulation results signify that the proposed method could achieve about 134 dB SI suppression capability, BER performance comparable to zero SI, and enabled IBFD operations in wireless communication systems, superior to the aforementioned approaches.

Data Availability

The data including simulation configurations, parameters, and results used to support the findings of this study are included within the article.

Disclosure

Partial content of this study has been previously presented in [27] as conference.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was partly conducted under a contract of the R&D for Expansion of Radio Wave Resources (JPJ00254), organized by the Ministry of Internal Affairs and Communications (MIC), Japan.

References

- [1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.
- [2] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.
- [3] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: challenges and opportunities," *IEEE Transactions on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, 2014.
- [4] D. Bharadia, E. McMillin, and S. Katti, "Full duplex radios," in *In The 13th ACM Special Interest Group on Data Communication (SIGCOMM)*, pp. 1–12, Hong Kong, China, August 2013.
- [5] J.-H. Xun, L.-F. Shi, W.-R. Liu, and D.-J. Xin, "A self-interference suppression structure for collinear dipoles," *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 10, pp. 2100–2104, 2019.
- [6] Y. Liu, W. Ma, X. Quan, W. Pan, K. Kang, and Y. Tang, "An architecture for capturing the nonlinear distortion of analog self-interference cancellers in full-duplex radios," *IEEE Microwave and Wireless Components Letters*, vol. 27, no. 9, pp. 845–847, 2017.
- [7] H. Guo, J. Xu, S. Zhu, and S. Wu, "Realtime software defined self-interference cancellation based on machine learning for in-band full duplex wireless communications," in *In 2018 International Conference on Computing, Networking and Communications (ICNC)*, pp. 1–5, Maui, HI, USA, July 2018.
- [8] L. Shen, B. Henson, Y. Zakharov, and P. Mitchell, "Digital self-interference cancellation for full-duplex underwater acoustic systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 1, pp. 192–196, 2020.
- [9] C. Zhang and X. Luo, "Adaptive digital self-interference cancellation for millimeter-wave full-duplex backhaul systems," *IEEE Access*, vol. 7, pp. 175542–175553, 2019.
- [10] K. Komatsu, Y. Miyaji, and H. Uehara, "Basis function selection of frequency-domain Hammerstein self-interference canceller for in-band full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3768–3780, 2018.
- [11] T. Matsumura and F. Kojima, "Prototype of analog self-interference cancellation based on super-heterodyne architecture for in-band full-duplex cellular system," in *In 2020 23rd International Symposium on Wireless Personal Multimedia*

- Communications (WPMC)*, pp. 1–5, Okayama, Japan, October 2020.
- [12] C. D. Nwankwo, L. Zhang, A. Quddus, M. A. Imran, and R. Tafazolli, “A survey of self-interference management techniques for single frequency full duplex systems,” *IEEE Access*, vol. 6, pp. 30242–30268, 2018.
 - [13] C. Psomas, C. Skouroumounis, I. Krikidis, A. Kalis, Z. Theodosiou, and A. Kounoudes, “Performance gains from directional antennas in full-duplex systems,” in *In 2015 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS)*, pp. 1–5, Tel Aviv, Israel, November 2015.
 - [14] A. K. Khandani, “Two-way (true full-duplex) wireless,” in *In 2013 13th Canadian Workshop on Information Theory*, pp. 33–38, Toronto, ON, Canada, October 2013.
 - [15] J. Mirza, G. Zheng, S. Saleem, and K.-K. Wong, “Optimization of uplink CSI training for full-duplex multiuser MIMO systems,” *IEEE Communications Letters*, vol. 23, no. 12, pp. 2325–2329, 2019.
 - [16] W. Zhang, J. Yin, D. Wu, G. Guo, and Z. Lai, “A self-interference cancellation method based on deep learning for beyond 5G full-duplex system,” in *2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pp. 1–5, Qingdao, China, December 2018.
 - [17] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, “Multi-user full duplex transceiver design for mmWave systems using learning-aided channel prediction,” *IEEE Access*, vol. 7, pp. 66068–66083, 2019.
 - [18] H. Guo, S. Wu, H. Wang, and M. Daneshmand, “DSIC: deep learning based self-interference cancellation for in-band full duplex wireless,” in *In 2019 IEEE global communications conference (GLOBECOM)*, pp. 1–6, Waikoloa, HI, USA, December 2019.
 - [19] M. Erdem, H. Ozkan, and O. Gurbuz, “Nonlinear digital self-interference cancellation with SVR for full duplex communication,” in *In Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, Seoul, Korea (South), May 2020.
 - [20] A. Balatsoukas-Stimming, “Joint detection and self-interference cancellation in full-duplex systems using machine learning,” in *In 2021 55th Asilomar conference on signals, Systems, and Computers*, pp. 989–992, Pacific Grove, CA, USA, November 2021.
 - [21] S. Huang, Y. Ye, and M. Xiao, “Learning-based hybrid beam-forming design for full-duplex millimeter wave systems,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 120–132, 2021.
 - [22] K. E. Kolodziej, A. U. Cookson, and B. T. Perry, “RF canceller tuning acceleration using neural network machine learning for in-band full-duplex systems,” *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1158–1170, 2021.
 - [23] L. Shan, O. Zhao, K. Temma, K. Hattori, F. Kojima, and F. Adachi, “Evaluation of machine learnable bandwidth allocation strategy for user cooperative traffic forwarding,” *IEEE Access*, vol. 7, pp. 85213–85225, 2019.
 - [24] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, “Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks,” *IEEE Access*, vol. 6, pp. 32328–32338, 2018.
 - [25] W.-S. Liao, O. Zhao, M. G. Kibria, G. P. Villardi, K. Ishizu, and F. Kojima, “Machine learning based signal detection for comp downlink in ultra-dense small cell networks,” *IEEE Access*, vol. 8, pp. 17454–17463, 2020.
 - [26] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
 - [27] O. Zhao, W.-S. Liao, K. Li, T. Matsumura, F. Kojima, and H. Harada, “Lazy learning-based self-interference cancellation for wireless communication systems with in-band full-duplex operations,” in *In 2021 IEEE 32th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1589–1594, Helsinki, Finland, October 2021.
 - [28] E. K. Garcia, S. Feldman, M. R. Gupta, and S. Srivastava, “Completely lazy learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 9, pp. 1274–1285, 2010.
 - [29] F. A. P. de Figueiredo, F. A. C. M. Cardoso, I. Moerman, and G. Fraidenraich, “Channel estimation for massive MIMO TDD systems assuming pilot contamination and flat fading,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, 11 pages, 2018.
 - [30] J. P. Dunsmore, *Handbook of Microwave Component Measurements*, John Wiley & Sons, 2012.
 - [31] Agilent, “Fundamentals of RF and microwave noise figure measurements,” *Application Note*, pp. 1–34, Keysight Technologies, 2010.
 - [32] A. Goldsmith, *Wireless Communications*, Stanford University, 2012.
 - [33] O. Zhao, L. Shan, W. S. Liao et al., “A device-centric clustering approach for large-scale distributed antenna systems using user cooperation,” *IEICE Transactions on Communications*, vol. 7, no. 2, pp. 359–372, 2019.
 - [34] J. L. Bentley, “Multidimensional binary search trees used for associative searching,” *Communications of the ACM*, vol. 18, no. 9, pp. 509–517, 1975.
 - [35] Y. A. Malkov and D. A. Yashunin, “Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence, Early Access Article*, vol. 42, no. 4, pp. 824–836, 2018.
 - [36] R. Weber, H. J. Schek, and S. Blott, “A quantitative analysis and performance study for similarity search methods in high dimensional spaces,” in *In VLDB’98 Proceedings of the 24th International Conference on Very Large Data Bases*, pp. 194–205, New York City, New York, USA, 1998.
 - [37] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, “Deep learning for wireless physical layer: opportunities and challenges,” *China Communications*, vol. 14, pp. 92–111, 2017.
 - [38] O. Zhao and H. Murata, “A study on dynamic clustering for large-scale multi-user MIMO distributed antenna systems with spatial correlation,” *IEICE Transactions on Communications*, vol. 99, no. 4, pp. 928–938, 2016.
 - [39] O. Zhao, W.-S. Liao, K. Ishizu, and F. Kojima, “Dynamic and non-centric networking approach using virtual gateway platforms for low power wide area systems,” *IEEE Access*, vol. 7, pp. 186078–186090, 2019.
 - [40] O. Zhao and H. Murata, “Sum-rate analysis for centralized and distributed antenna systems with spatial correlation and inter-cell interference,” *IEICE Transactions on Communications*, vol. E98.B, no. 3, pp. 449–455, 2015.
 - [41] H. P. Gavin, “The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems,” *Department of Civil and Environmental Engineering, Duke University*, vol. 19, 2019.