Extreme Learning Machine-Based Feature Refinment for Channel Estimation in RIS-ISAC Systems

Alice Faisal, *Graduate Student Member, IEEE*, Ibrahim Al-Nahhal, *Senior Member, IEEE*, Octavia A. Dobre, *Fellow, IEEE*, and Hyundong Shin, *Fellow, IEEE*

***Abstract*—Integrated sensing and communication (ISAC) systems have emerged as a key enabler for future wireless networks, aiming to optimize spectral resource utilization for both sensing and communication tasks. The incorporation of reconfigurable intelligent surfaces (RIS) with ISAC enables more efficient utilization of resources, improving the quality of communication and the accuracy of sensing. A critical aspect of deploying such systems reliably is accurate channel estimation. Traditional deep learning methods, though effective, often struggle with the intricate characteristics of communication channel matricies. This work introduces an innovative two-stage channel estimation approach for RIS-ISAC systems. The first stage focuses on feature refinement using an extreme learning machine framework to process the received signals and extract essential channel features. The second stage employs these refined features to estimate the desired channels through dedicated neural networks specifically designed for sensing and communication tasks. The numerical simulations demonstrate that the proposed two-stage approach significantly outperforms the existing techniques across various system configurations. Moreover, the proposed method achieves a remarkable computational complexity reduction as compared to the state-of-the-art works. The results prove the robustness and efficiency of the proposed approach, facilitating more robust RIS-assisted ISAC deployments.**

***Index Terms*—Reconfigurable intelligent surface (RIS), integrated sensing and communication (ISAC), channel estimation, deep learning, extreme learning machine (ELM).**

I. Introduction

As THE demand for higher data rates and ubiquitous connectivity continues to escalate, future-generation wireless communication systems face unprecedented challenges. Traditional communication systems, which often operate independently of sensing functions, are suffering in meeting the requirements of modern applications. To this end, there is a critical need for innovative solutions that can address the demands for enhanced spectral efficiency and reliable connectivity [1], [2], [3]. Integrated sensing and communication (ISAC) has emerged as a promising concept that leverages simultaneous communication and sensing functionalities [4], [5], [6], [7], [8], thus, enhancing spectral efficiency and reducing hardware costs. The integration of these two functionalities brings many opportunities to wireless systems. In particular, ISAC allows for adaptive networks that can dynamically adjust to changing conditions, which results in improved reliability in various scenarios. Moreover, the dual-use of signals for communication and sensing reduces latency and enhances the overall communication by providing situational awareness. Consequently, ISAC systems are envisioned to revolutionize the landscape of the next-generation wireless networks by providing a unified solution that meets the increasing demands for both high-speed connectivity and precise sensing capabilities [4], [9], [10], [11], [12], [13].

In addition to ISAC, the reconfigurable intelligent surfaces (RIS) technology has emerged as a key enabler for future-generation wireless communication systems [14], [15], [16]. Consisting of a large array of passive reflecting elements, RIS can dynamically manipulate the electromagnetic waves to improve signal propagation and help mitigate adverse effects, such as signal scattering, refraction, and interference. By intelligently tuning the phase, polarization, and amplitude of reflected signals, RIS can optimize the wireless conditions, leading to enhanced data rates, extended coverage, and improved energy efficiency [17], [18]. The integration of RIS in communication systems offers promising opportunities for network optimization, particularly in challenging environments. Furthermore, RIS can be incorporated into existing infrastructure as a cost-effective solution for upgrading current networks. The potential of the RIS extends beyond communication, as it can also facilitate integrated sensing functions, thereby playing a crucial role in the development of advanced ISAC systems [19], [20].

In particular, the integration of RIS with ISAC systems marks a notable development in the design and optimization of next-generation wireless networks [21], [22], [23], [24]. RIS technology, with its ability to intelligently manipulate the propagation environment, complements the dual-functionality of ISAC systems, which simultaneously support communication and sensing tasks. Since both tasks operate over shared spectral and hardware resources, they inherently interfere with each other. By incorporating RIS into ISAC frameworks, the propagation environment can be dynamically reconfigured to enhance signal quality, mitigate interference between communication and sensing tasks, and improve the overall system performance. This integration allows

Alice Faisal and Ibrahim Al-Nahhal are with the Memorial University St. John's: NL A1C 5S7, Canada (e-mail: afaisal@mun.ca; ioalnahhal@mun.ca).

Octavia A. Dobre is with the Memorial University St. John's: NL A1C 5S7, Canada, and also with the Kyung Hee University, Gyeonggi-do 17104, South Korea (e-mail: odobre@mun.ca).

Hyundong Shin is with the Kyung Hee University, Gyeonggi-do 17104, South Korea (e-mail: hshin@khu.ac.kr).

for precise control over the signal propagation paths, thereby enabling more accurate sensing and reliable communication. The combined capabilities of the RIS and ISAC have been recently explored in different areas such as autonomous driving, smart environments, and industrial automation, where real-time situational awareness and robust connectivity are paramount [9], [10], [25], [26], [27]. It has been showed that the nature of RIS can be leveraged to address the inherent challenges of spectrum sharing in ISAC systems, thus, improving the overall system performance significantly.

Most of the work in the literature focused on the resource allocation, beamforming design, and deployment of RIS-assisted ISAC across various scenarios, often assuming perfect channel state information (CSI) [23], [24], [28], [29], [30]. However, this assumption does not always hold in practice. In real-world systems, CSI acquisition is limited by estimation errors, pilot contamination, hardware impairments, and feedback constraints [31], [32], [33], [34]. These imperfections can significantly degrade the communication and sensing performance, especially in RIS-assisted ISAC systems where both functions share spectrum and hardware resources. Therefore, developing accurate and robust channel estimation frameworks is essential for the practical realization of such systems. The difficulty of channel estimation in RIS-assisted ISAC systems arises from the inherent interference between communication and sensing functions and the increased number of channel links. Unlike traditional systems, RIS introduces additional links and high-dimensional channel matrices, making traditional estimation techniques computationally challenging. Moreover, the dual functionality introduced by ISAC systems complicates the channel estimation process and requires advanced techniques to accurately differentiate between communication and sensing signals. A few works investigated the challenge of channel estimation in RIS-assisted ISAC systems [35], [36], [37]. Deep learning (DL)-based methods were considered due to their ability in handling high-dimensional data and learn complex patterns, thereby potentially offering more accurate and efficient solutions compared to traditional techniques. The study in [35] examined the channel estimation issue in a single-user RIS-assisted ISAC system. A convolutional neural network (CNN) framework was proposed to accurately estimate the channels in three stages. Additionally, the study in [36] investigated a multi-user RIS-assisted ISAC system, where a two-stage deep NN (DNN)-based framework was proposed to estimate the channels. In particular, the initial stage assumed that the RIS was deactivated to coherently estimate the sensing channel, and the final stage assumed that the base station (BS) was off to estimate the communication channel accurately. The work in [37] investigated an uplink and downlink multi-user RIS-assisted ISAC system, in which an extreme learning machine (ELM) approach was proposed to estimate the direct and reflected links in two stages, respectively. Finally, the work in [38] further proposed the use of conditional generative adversarial networks (CGANs) to tackle the channel estimation challenge in multi-user RIS-assisted ISAC systems.

The majority of the aforementioned studies have evaluated the performance using the least squares (LS) estimation method as a baseline. While commonly employed, LS suffers from several limitations, including its susceptibility to noise and challenges in handling non-linear estimation problems. These drawbacks result in inaccurate and inefficient estimates [39]. This highlights the need for novel methods that do not only improve the estimation accuracy but are also computationally efficient. Within this context, this work proposes a novel estimation approach designed to address the limitations of existing methods. The proposed approach consists of two key stages: The first stage is feature refinement, which utilizes an ELM to extract essential features from noisy signals. The second stage utilizes the refined features to estimate the desired channels. This two-stage approach is applied to both sensing and communication links, with sensing channels estimated at the BS and communication links at the user side. The main contributions are outlined as follows:

- A novel two-stage estimation framework is developed, incorporating feature refinement and estimation stages.
- An ELM model is specifically designed to extract essential features from noisy received signals, leveraging its advantages of faster training times and better generalization capabilities compared to traditional learning methods.
- Two distinctive models are designed to estimate sensing and communication channels, where one is deployed at the BS to approximate sensing channels, and the other at the user side to estimate the communication channels.
- The proposed approach demonstrates outstanding performance compared to existing works. This is mainly accomplished by the feature refinement stage design, which ensures more accurate and reliable channel estimation in RIS-assisted ISAC systems. The design of this step is crucial as it aims to avoid computationally expensive networks, while also carefully preserving the inherent characteristics of the channel.
- The computational complexity is evaluated to assess its efficiency and feasibility, demonstrating a considerable reduction in the complexity compared to the literature estimators. This is accomplished while simultaneously delivering significant enhancements in estimation performance.
- Comprehensive simulations are conducted to assess the estimation performance of the proposed method, with numerical results indicating significant improvements over existing methods in terms of accuracy and robustness.

The rest of the paper is structured as follows: Section II outlines the system model and problem formulation. The proposed estimation approach and its complexity analysis are detailed in Sections III and IV, respectively. Lastly, Sections V and VI present the simulation results and conclusions.

II. System Model

Consider a scenario involving an RIS-assisted multiple-input single-output multi-user ISAC system, consisting of user equipment (UEs), a BS, an RIS, and a target as shown in Fig. 1(a). The full-duplex (FD) BS facilitates communication and sensing tasks simultaneously by transmitting and receiving signals to and from the UEs and target. The RIS consists of N passive reflecting elements that are dynamically controlled to manipulate the

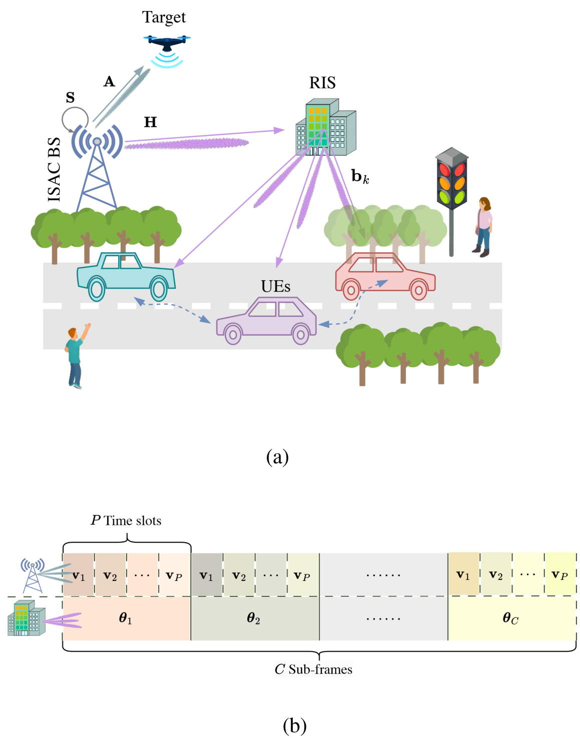


Fig. 1. RIS-assisted ISAC system: (a) System architecture, (b) Pilot transmission protocol.

signal propagation environment. The RIS reflects signals from the BS towards K UEs, facilitating improved communication performance. The BS is equipped with M transmit antennas and both BS and UE are equipped with a single receive antenna. The BS-target-BS channel, where the BS transmits radar signals to the target and receives the echo signals, denoted by $\mathbf{A} \in \mathbb{C}^{M\times M}$. The direct link between the BS and UE is presumed to be hindered by obstacles, and the RIS is utilized to enhance communication. To this end, the channel coefficients corresponding to the link between the BS and RIS, the RIS and k-th user, are denoted by $\mathbf{H} \in \mathbb{C}^{M\times N}$ and $\mathbf{b}_k \in \mathbb{C}^{N\times 1}$, respectively. As the ISAC BS functions in FD mode, it simultaneously transmits and receives signals, resulting in self-interference (SI), denoted by $\mathbf{S} \in \mathbb{C}^{M\times M}$. Considering the narrowband transmission, all channels are presumed to exhibit flat-fading characteristics. To address the inherent interference between communication and sensing tasks, a pilot transmission scheme is developed to estimate the RIS-assisted ISAC system channels, as depicted in Fig. 1(b).

The BS sends pilot sequences across multiple sub-frames, denoted by C. Each sub-frame is segmented into several time slots, represented by P. Within the p-th time slot, the pilot signal is expressed as $\mathbf{v}_p \in \mathbb{C}^{M\times 1}$. Collectively, the pilot signals for the c-th sub-frame form the pilot matrix $\mathbf{V}_c = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_P] \in \mathbb{C}^{M\times P}$. To ensure uniformity across sub-frames, the ISAC BS sends the same pilot sequences, represented by $\mathbf{V}$. During a single sub-frame, the RIS phase-shifts stay unchanged and are represented by $\boldsymbol{\theta}_c \in \mathbb{C}^{N\times 1}$, where $\boldsymbol{\theta}_c = \left(e^{j\varphi_1}, \ldots, e^{j\varphi_n}, \ldots, e^{j\varphi_N}\right) \in \mathbb{C}^{N\times 1}$ and $\varphi_n \in [0, 2\pi)$ is the phase shift introduced by the n-th reflecting element. The phase shift matrix for the entire set of sub-frames is denoted as $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_C] \in \mathbb{C}^{N\times C}$.

To minimize the pilot overhead, it is assumed that $P = M$ and $C = N$. This choice is based on the principle that each antenna requires at least one distinct pilot time slot to capture the required channel information. Such a design strikes a balance between estimation accuracy and pilot efficiency, effectively capturing the full channel characteristics without introducing redundancy or incurring unnecessary pilot overhead.

Moreover, both the pilot matrix $\mathbf{V}$ and the phase shift matrix $\boldsymbol{\Theta}$ are structured using the discrete Fourier transform (DFT) matrix. The DFT matrix $\mathbf{F}_N \in \mathbb{C}^{N\times N}$ is defined as

$$\mathbf{F}_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j\frac{2\pi}{N}} & \cdots & e^{-j\frac{2\pi(N-1)}{N}} \\ 1 & e^{-j\frac{4\pi}{N}} & \cdots & e^{-j\frac{4\pi(N-1)}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\frac{2\pi(N-1)}{N}} & \cdots & e^{-j\frac{2\pi(N-1)(N-1)}{N}} \end{bmatrix}, \quad (1)$$

where N is the size of the matrix. Thus, the pilot matrix $\mathbf{V}$ and the phase shift matrix $\boldsymbol{\Theta}$ are given by

$$\mathbf{V} = \mathbf{F}_M, \quad (2)$$

and

$$\boldsymbol{\Theta} = \mathbf{F}_N. \quad (3)$$

This modeling helps mitigate interference between sensing and communication signals while enabling the separation of multi-user transmissions. It ensures that the pilot signals and phase shifts leverage the orthogonality and frequency domain properties of the DFT, leading to efficient signal processing and accurate channel estimation in the RIS-assisted multi-user ISAC systems [40].

Accordingly, the signal received by the k-th UE within c and p, is represented as

$$y^{\mathrm{UE}}_{k,c,p} = \mathbf{b}_k^H \mathrm{diag}\{\boldsymbol{\theta}_c^H\}\mathbf{H}^H\mathbf{v}_p + n_{k,c,p}, \quad (4)$$

where $n_{k,c,p} \sim \mathcal{CN}(0, \sigma^2)$ represents the additive white complex Gaussian noise with zero-mean and variance of σ^2. The received signal in (4) can be algebraically rearranged to isolate the cascaded communication channel by using the identity $\mathbf{b}_k^H \mathrm{diag}(\boldsymbol{\theta}^H) = \mathrm{diag}(\mathbf{b}_k^H)\boldsymbol{\theta}^H$. Applying this, the expression becomes

$$y^{\mathrm{UE}}_{k,c,p} = \left(\mathbf{b}_k^H \mathrm{diag}\{\boldsymbol{\theta}_c^H\}\mathbf{H}^H\right)\mathbf{v}_p + n_{k,c,p} \quad (5)$$

$$= \left(\mathrm{diag}\{\mathbf{b}_k^H\}\boldsymbol{\theta}_c^H\mathbf{H}^H\right)\mathbf{v}_p + n_{k,c,p}. \quad (6)$$

Let $\mathbf{R}_k = \mathbf{H}\mathrm{diag}\{\mathbf{b}_k\}$, then, (4) becomes

$$y^{\mathrm{UE}}_{k,c,p} = \boldsymbol{\theta}_c^H \mathbf{R}_k^H \mathbf{v}_p + n_{k,c,p}. \quad (7)$$

This matrix abstraction facilitates the estimation of the effective cascaded channel $\mathbf{R}_k$, which jointly captures the links between the BS-RIS and RIS-k-th user.

In contrast, the signal received by the BS is given as

$$\mathbf{y}^{\mathrm{BS}}_{c,p} = \mathbf{A}^H\mathbf{v}_p + \mathbf{S}^H\mathbf{v}_p + \mathbf{n}_{c,p}, \quad (8)$$

where $\mathbf{A}^H\mathbf{v}_p$ represents the sensing signal, $\mathbf{S}^H\mathbf{v}_p$ accounts for the residual SI, and $\mathbf{n}_{c,p}$ represents the noise. The noise term is modeled as $\mathbf{n}_{c,p} \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_M)$, denoting the additive white complex Gaussian noise with zero-mean and variance of σ^2, where $\mathbf{I}_M$ is an identity matrix of size M. The SI channel, $\mathbf{S}$, can be estimated to mitigate any residual SI in (8) before performing channel estimation. This is feasible because the propagation environment between the transmitting and receiving antennas of the ISAC BS is considered stable [41]. Consequently, estimating the communication and sensing channels in (4) and (8) is necessary. The pilot symbol duration is assumed to be shorter than the channel coherence time but long enough to exceed the propagation delays of transmitted signals [42], [43]. This assumption guarantees that the SAC channels remain constant throughout the estimation phase, while also enabling synchronization of the received SAC and residual SI signals. In the subsequent sections, we introduce an innovative method for estimating both $\mathbf{R}_k$ and $\mathbf{A}$ utilizing the pilot protocol depicted in Fig. 1(b).

III. Proposed Estimation Approach

In channel estimation problems, the primary information available at the receiver side is the noisy received signal, which can be utilized without additional processing. Although using this information directly to estimate channels can yield good performance, it is often limited by the effects of noise, cascaded channels, and interference. To overcome these limitations, the proposed approach divides the estimation process into two distinct stages: input feature refinement and channel estimation. The first stage is dedicated to extracting features to enhance the quality of the noisy received signal. In the second stage, the refined output from the first stage is used to accurately estimate the channels. The proposed framework is designed to perform channel estimation locally at the BS for sensing and at the UE for communication, thereby minimizing the need for real-time feedback. Within each of these stages, the sensing and communication signals are processed separately to ensure optimal NN design for each type of signal. This approach aims to maximize the accuracy and reliability of the channel estimation process. In what follows, the proposed approach is explained in detail.

A. Stage 1: Feature Refinement

In this subsection, a feature refinement stage is designed for both sensing and communication links. An ELM-based NN is deployed to take the noisy received signal as input and produce a refined version, preserving critical channel information while mitigating adverse effects of noise and interference. This approach sets a solid foundation for accurate channel estimation in the second stage.

1) ELM: The ELM is structured as a feedforward NN (FNN) consisting of input, single hidden, and output layers. Unlike traditional FNNs that use gradient-based learning algorithms, ELMs offer a more efficient training method by randomly initializing the weights and biases of the hidden layer neurons and analytically determining the output weights [44], [45]. Consider an NN with η_I input neurons, η_h hidden neurons, and η_O output neurons. Let $\mathbf{Z} \in \mathbb{R}^{N_{tr}\times\eta_I}$ and $\mathbf{Y} \in \mathbb{R}^{N_{tr}\times\eta_O}$ denote the input and output data matrices, respectively, where N_{tr} is the number of training samples. The process of training an ELM can be described as follows: The input weights $\mathbf{W} \in \mathbb{R}^{\eta_h\times\eta_I}$ and biases $\mathbf{b} \in \mathbb{R}^{\eta_h\times 1}$ for the hidden layer neurons are randomly generated. The input data $\mathbf{Z}$ is composed of individual input vectors ζ. The output of the hidden layer $\mathbf{U} \in \mathbb{R}^{N_{tr}\times\eta_h}$ is computed as follows:

$$\mathbf{U}=\begin{bmatrix} f(\boldsymbol{\omega}_1\zeta(1,1)+b_1) & \cdots & f(\boldsymbol{\omega}_{\eta_h}\zeta(1,1)+b_{\eta_h}) \\ \vdots & \ddots & \vdots \\ f(\boldsymbol{\omega}_1\zeta(N_{tr},\eta_I)+b_1) & \cdots & f(\boldsymbol{\omega}_{\eta_h}\zeta(N_{tr},\eta_I)+b_{\eta_h}) \end{bmatrix}. \tag{9}$$

Here, $\boldsymbol{\omega}_{\eta_h}$ represents the h-th row of the weight matrix $\mathbf{W}$, and b_{η_h} represents the h-th element of the bias vector $\mathbf{b}$. Specifically, the matrix form is expressed as

$$\mathbf{U} = f(\mathbf{Z}\mathbf{W}^T + \mathbf{b}), \tag{10}$$

where $f(\cdot)$ denotes the activation function applied to the output.

The output weights $\boldsymbol{\Lambda}_{\text{ELM}} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_{\eta_h}]^{\text{T}} \in \mathbb{R}^{\eta_h\times\eta_O}$, are determined by solving the linear system as

$$\mathbf{Y} = \mathbf{U}\boldsymbol{\Lambda}_{\text{ELM}}. \tag{11}$$

It has been demonstrated in [44] that if the activation function $f(\cdot)$ is differentiable, $\boldsymbol{\omega}$ and b in $\mathbf{H}$ do not require to be updated during the training process. This finding allows (11) to be treated as a linear system. To this end, the ELM model updates $\boldsymbol{\Lambda}$ by minimizing the loss function as

$$\hat{\boldsymbol{\Lambda}}_{\text{ELM}} = \min_{\Lambda} \|\mathbf{U}\boldsymbol{\Lambda}_{\text{ELM}} - \mathbf{Y}\|_2. \tag{12}$$

The solution is obtained using the Moore-Penrose pseudoinverse of $\mathbf{U}$ as

$$\hat{\boldsymbol{\Lambda}}_{\text{ELM}} = \mathbf{U}^{\dagger}\mathbf{Y}. \tag{13}$$

According to the above training scheme, ELM has several advantages over traditional FNN. In particular, ELM significantly accelerates the learning process while minimizing the number of trainable parameters compared to traditional gradient-based learning methods [46]. Specifically, the linear relationship in (11) holds as $\mathbf{U}$ remains constant during the training process. To this end, unlike the typical approach where all parameters are updated through backpropagation, the ELM only requires the determination of the weights $\boldsymbol{\Lambda}$ by performing matrix inversion. This implies that the output weights are computed analytically, rather than through iterative updates, and therefore ELM ensures deterministic behavior for a given input. Through its simplicity and by avoiding the iterative tuning of weights, ELM ensures minimal training error and provides robust generalization capabilities. This is a significant advantage over many traditional learning algorithms, which might encounter local minima, overfitting issues, or difficulties in choosing the correct learning rate [44], [45]. Furthermore, the ELM-based feature refinement facilitates a structured design in the proposed framework. By efficiently denoising the received signal, the ELM stage simplifies the learning requirements for the subsequent estimation networks. This enables the use of computationally efficient NNs in the second stage, dedicated to estimating the sensing and communication channels from the denoised features. To this end, the combination of ELM and task-specific NNs enables improved generalization, reduced complexity, and strong estimation performance in practical RIS-assisted ISAC scenarios.

2) Input-Output Design: Consider the pilot design adopted in Section II to generate the input-output pairs for the feature refinement stage. In the c-th sub-frame, P received signals at UE_k are stacked, simplifying (7) to

$$\mathbf{y}_{k,c}^{\mathrm{UE}} = \boldsymbol{\theta}_c^H \mathbf{R}_k^H \mathbf{V} + \mathbf{n}_{k,c}. \tag{14}$$

Here, $\mathbf{y}_{k,c}^{\mathrm{UE}} = [y_{k,c,1}^{\mathrm{UE}}, y_{k,c,2}^{\mathrm{UE}}, \dots, y_{k,c,P}^{\mathrm{UE}}] \in \mathbb{C}^{1\times P}$ and $\mathbf{n}_{k,c} = [n_{k,c,1}, n_{k,c,2}, \dots, n_{k,c,P}] \in \mathbb{C}^{1\times P}$. The data generation is designed based on the received signal, and is given as

$$\begin{aligned}\zeta_{S_1}^{\mathrm{UE}_k} = \Big[&\mathfrak{Re}\{\mathbf{y}_{k,1}^{\mathrm{UE}}, \mathbf{y}_{k,2}^{\mathrm{UE}}, \dots, \mathbf{y}_{k,C}^{\mathrm{UE}}\},\\ &\mathfrak{Im}\{\mathbf{y}_{k,1}^{\mathrm{UE}}, \mathbf{y}_{k,2}^{\mathrm{UE}}, \dots, \mathbf{y}_{k,C}^{\mathrm{UE}}\}\Big]^T,\end{aligned} \tag{15}$$

where $\mathfrak{Re}\{\cdot\}$ and $\mathfrak{Im}\{\cdot\}$ represent the extraction of the real and imaginary components of a complex vector, respectively. The corresponding output is constructed to simulate the ideal signal representation, which theoretically corresponds to the received signal without the noise factor. This can be mathematically expressed by isolating the term $\boldsymbol{\theta}_c^H \mathbf{R}_k^H \mathbf{V}$ from (14), thus focusing on channel characteristics and pilot sequences. Specifically, the desired output vector, $\gamma_{S_1}^{\mathrm{UE}_k}$, is given by

$$\gamma_{S_1}^{\mathrm{UE}_k} = \left[\mathfrak{Re}\{\boldsymbol{\theta}_c^H \mathbf{R}_k^H \mathbf{V}\}, \mathfrak{Im}\{\boldsymbol{\theta}_c^H \mathbf{R}_k^H \mathbf{V}\}\right]^T \in \mathbb{C}^{2C\times P}. \tag{16}$$

This output representation serves as a benchmark for training the feature refinement stage, ensuring that the network is capable of accurately identifying and preserving the fundamental signal attributes despite the imperfections introduced by the noisy environment.

Conversely, to construct the input-output mappings for $\mathbf{A}$, refer to the received signal in (8). In a similar manner, the P received signals at the ISAC BS are aggregated across the c-th sub-frame, resulting in

$$\mathbf{Y}_c^{\mathrm{BS}} = \mathbf{A}^H \mathbf{V} + \mathbf{N}_c. \tag{17}$$

Here, $\mathbf{Y}_c^{\mathrm{BS}} = [\mathbf{y}_{c,1}^{\mathrm{BS}}, \mathbf{y}_{c,2}^{\mathrm{BS}}, \dots, \mathbf{y}_{c,P}^{\mathrm{BS}}] \in \mathbb{C}^{M\times P}$ and $\mathbf{N}_c = [\mathbf{n}_{c,1}, \mathbf{n}_{c,2}, \dots, \mathbf{n}_{c,P}] \in \mathbb{C}^{M\times P}$. Therefore, the data generation is given as

$$\begin{aligned}\zeta_{S_1}^{\mathrm{BS}} = \Big[&\mathfrak{Re}\{\mathbf{Y}_1^{\mathrm{BS}}, \mathbf{Y}_2^{\mathrm{BS}}, \dots, \mathbf{Y}_C^{\mathrm{BS}}\},\\ &\mathfrak{Im}\{\mathbf{Y}_1^{\mathrm{BS}}, \mathbf{Y}_2^{\mathrm{BS}}, \dots, \mathbf{Y}_C^{\mathrm{BS}}\}\Big]^T.\end{aligned} \tag{18}$$

The corresponding output, which represents the ideal signal, can be expressed as

$$\gamma_{S1}^{\mathrm{BS}} = \left[\mathfrak{Re}\{\mathbf{A}^H \mathbf{V}\}, \mathfrak{Im}\{\mathbf{A}^H \mathbf{V}\}\right]^T \in \mathbb{C}^{2C\times P}. \tag{19}$$

Given the input-output design for both sensing and communication links, the following sub-section details the proposed networks for the feature refinement stage.

TABLE I
ARCHITECTURE OF S-ELM AND C-ELM

Model	Layer	Size	Activation Function
S-ELM	Input	$2MP$	-
	Hidden	200	*Sigmoid*
	Output	$2MP$	-
C-ELM	Input	$2PC$	-
	Hidden	900	*Sigmoid*
	Output	$2PC$	-

3) Training and Testing: The proposed NN framework incorporates two distinct ELM structures. The first, known as the sensing estimation ELM (S-ELM), is utilized for estimating direct sensing channels, while the second, termed the communication estimation ELM (C-ELM), is designed for estimating reflected communication channels. The configuration of the S-ELM and C-ELM hyperparameters is carefully tailored for their respective estimation tasks. In particular, the hidden layer of the S-ELM includes $\eta_h = 200$ neurons. Furthermore, given the increased complexity of reflected channel estimation, the hidden layer of the C-ELM is expanded to $\eta_h = 900$ neurons. Both utilize *sigmoid* activation functions to enhance the feature refinement capabilities.

To enhance the training process, a normalization is performed on the input data. It can be mathematically represented as

$$\zeta_{\mathrm{std}}^{\mathrm{BS/UE}_k}(q,x) = \frac{\zeta^{\mathrm{BS/UE}_k}(q,x) - \mathbb{E}\{\zeta^{\mathrm{BS/UE}_k}(q,x)\}}{\mathbb{D}\{\zeta^{\mathrm{BS/UE}_k}(q,x)\}}, \quad x \in \mathcal{N}_1^X, q \in \mathcal{N}_1^Q. \tag{20}$$

Here, $\mathbb{E}\{\zeta^{\mathrm{BS/UE}_k}(q,x)\}$ and $\mathbb{D}\{\zeta^{\mathrm{BS/UE}_k}(q,x)\}$ represent the mean and standard deviation of the input, respectively.

In the testing phase, the input is initially normalized as detailed in (20). By obtaining $\mathbf{H}$ according to (9), $\hat{\mathbf{\Lambda}}$ can be obtained as in (13). To this end, the trained ELM produces an output, denoted by $\tilde{\mathbf{Y}}$, and is expressed as

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{H}}\hat{\mathbf{\Lambda}}. \tag{21}$$

The refined received signals for sensing and communication links are estimated as $\bar{\mathbf{Y}}_c^{\mathrm{BS}}$ and $\bar{\mathbf{y}}_c^{\mathrm{UE}_k}$, respectively. Table I summarizes the proposed ELM-based networks architectures.

B. Stage 2: Channel Estimation

The second stage of the channel estimation process utilizes the purified data produced in the first stage (i.e., feature refinement stage). The refined features serve as the foundation for accurately estimating the actual channels, where it better reflects on the channel characteristics within a less noisy environment. To achieve this, two specialized NNs are designed and optimized for the distinct tasks of sensing and communication channel estimation. The NNs are carefully tailored to handle the specific characteristics and requirements of their respective channels, ensuring precise and reliable channel estimation. By leveraging the enhanced quality of the input data from the first stage, this approach aims to deliver superior estimation performance, thereby facilitating more robust and efficient communication and sensing operations in multi-user RIS-assisted ISAC systems.

1) Input-Output Design: According to the proposed design of the first stage in Section III-A, the input to the channel estimation stage is based on refined received signal (i.e., which is obtained in the first stage). To this end, the communication link design is represented as

$$\zeta_{S_2}^{\mathrm{UE}_k} = \big[\mathfrak{Re}\{\bar{\mathbf{y}}_{k,1}^{\mathrm{UE}}, \bar{\mathbf{y}}_{k,2}^{\mathrm{UE}}, \dots, \bar{\mathbf{y}}_{k,C}^{\mathrm{UE}}\}, \\ \mathfrak{Im}\{\bar{\mathbf{y}}_{k,1}^{\mathrm{UE}}, \bar{\mathbf{y}}_{k,2}^{\mathrm{UE}}, \dots, \bar{\mathbf{y}}_{k,C}^{\mathrm{UE}}\}\big]^T. \tag{22}$$

The output, representing the target, corresponds to the ground truth channel, $\mathbf{R}_k$, represented as

$$\gamma_{S_2}^{\mathrm{UE}_k} = [\mathfrak{Re}\{\mathrm{vec}[\mathbf{R}_k]\}.\mathfrak{Im}\{\mathrm{vec}[\mathbf{R}_k]\}]^T. \tag{23}$$

Similarly, the sensing link design is formulated as

$$\zeta_{S_2}^{\mathrm{BS}} = \big[\mathfrak{Re}\{\bar{\mathbf{Y}}_1^{\mathrm{BS}}, \mathbf{Y}_2^{\mathrm{BS}}, \dots, \bar{\mathbf{Y}}_C^{\mathrm{BS}}\}, \\ \mathfrak{Im}\{\bar{\mathbf{Y}}_1^{\mathrm{BS}}, \bar{\mathbf{Y}}_2^{\mathrm{BS}} \cdots, \bar{\mathbf{Y}}_C^{\mathrm{BS}}\}\big]^T, \tag{24}$$

where the output is the ground truth channel, $\mathbf{A}$, given as

$$\gamma_{S_2}^{\mathrm{BS}} = [\mathfrak{Re}\{\mathrm{vec}[\mathbf{A}]\}, \mathfrak{Im}\{\mathrm{vec}[\mathbf{A}]\}]^T. \tag{25}$$

2) Training and Testing: In the channel estimation stage, two distinct NNs are developed for separate functionalities: sensing and communication. These are designated as S-EST and C-EST, respectively. The S-EST network, which serves the sensing link, is structured as FNN with input, hidden, and output layers. The hidden layer is equipped with 100 neurons and utilizes a *tanh* activation function. The use of an FNN is motivated by the relatively lower dimensionality and structural simplicity of the sensing estimation task, where a fully connected architecture is sufficient to learn the required input-output mapping. By segmenting the estimation process into two phases, the design allows the second NN, C-EST, to remain relatively simple while still achieving effective performance with only 100 neurons. However, due to the higher dimensionality and spatial structure of the cascaded communication channel, the C-EST network is modeled as a CNN to better capture the complex relationships inherent in this application. It consists of an input, two hidden, and an output layer. The first hidden layer is modeled as a convolutional layer equipped with 132 filters of size 4, using a *tanh* activation function and followed by a flatten layer. The last hidden layer is structured as FF layer with 900 neurons employing a linear activation function. This task-specific architecture improves computational efficiency and enhances the model approximation under various noisy conditions. The Adam optimizer is utilized, where the learning rate set at 2×10^{-4} to tune the CNN parameters. Minibatch transitions, each containing 200 samples, are randomly selected to minimize training sample correlation. Additionally, a stopping criterion is implemented during training: the process halts if validation accuracy fails to improve for 5 sequential epochs, or if the maximum number of epochs is reached, $T_\epsilon = 200$.

Given the input-output design of the estimation stage, the (q, v)-th pair in the dataset is first pre-processed according to (20). Additionally, the target label is scaled to enhance the estimation accurancy, expressed as

$$\gamma_s^{\mathrm{BS/UE}_k}(q, x) = \rho\gamma^{\mathrm{BS/UE}_k}(q, x), \tag{26}$$

TABLE II
ARCHITECTURE OF S-EST AND C-EST

Model	Layers	Size	Kernel	Activation Function
S-EST	Input	$2MP$	-	-
	FFL	100	-	*tanh*
	Output	$2M^2$	-	-
C-EST	Input	$2PC$	-	-
	CL	132	4	*tanh*
	FFL	900	-	linear
	Output	$2MN$	-	-

where ρ is the scaling factor. To this end, the input-output pair of the second stage is represented as $(\zeta_{\mathrm{std}}^{\mathrm{BS/UE}_k}(q,x), \gamma_s^{\mathrm{BS/UE}_k}(x))$. To this end, the proposed CNN estimates $\gamma^{\mathrm{BS/UE}_k}$ during training as

$$\hat{\gamma}^{\mathrm{BS/UE}_k}(x) \approx f\left(\zeta_{\mathrm{std}}^{\mathrm{BS/UE}_k}(q,x), \Lambda_{\mathrm{CNN}}^{\mathrm{BS/UE}_k}\right), \quad x \in N_X^1, q \in N_Q^1, \tag{27}$$

where $f(\cdot; \Lambda_{\mathrm{CNN}}^{\mathrm{BS/UE}_k})$ denote the CNN associated with the estimation task (i.e., sensing or communication). The proposed CNN optimizes the loss function, $\mathcal{L}$, by minimizing the mean square error (MSE) to update Λ, as

$$\mathcal{L} = \frac{1}{XQ} \sum_{x \in N_X^1, q \in N_Q^1} \\ \left(f\left(\zeta_{\mathrm{std}}^{\mathrm{BS/UE}_k}(q,x), \Lambda_{\mathrm{CNN}}^{\mathrm{BS/UE}_k}\right) - \hat{\gamma}^{\mathrm{BS/UE}_k}(x)\right)^2. \tag{28}$$

In the testing phase, the input is firstly standardized as detailed in (20). Subsequently, the output of the trained CNN is produced as

$$\hat{\gamma}^{\mathrm{BS/UE}_k}(x) = \frac{1}{\rho} f\left(\zeta_{\mathrm{std}}^{\mathrm{BS/UE}_k}, \hat{\Lambda}_{\mathrm{CNN}}^{\mathrm{BS/UE}_k}\right), \tag{29}$$

where $\hat{\Lambda}^{\mathrm{BS/UE}_k}$ represents the trained NN parameters. By performing simple operations on the output of the network, such as combining the real and imaginary parts, the estimated channels can be obtained. Specifically, the estimated channels for the sensing and communication links are represented as $\hat{\mathbf{A}}$ and $\hat{\mathbf{R}}_k$, respectively. The detailed working principles of both estimation stages are provided in Algorithm 1, while Fig. 2 summarizes the overall approach and demonstrates how both stages are integrated together. Table II summarizes the proposed networks architectures for the channel estimation task of the second stage.

C. Data Generation

The training samples are generated using Q received signals (i.e., containing the communication and sensing signals in (14) and (17), respectively). To enhance the dataset and improve model generalization, data augmentation is applied by creating $X - 1$ additional variations of signal q. These augmented samples are obtained by adding synthetic additive white Gaussian noise to the original channel, following the signal-to-noise ratio (SNR) relation $\mathrm{SNR}_{\mathbf{A}/\mathbf{R}_k} = \frac{p_{\mathbf{A}/\mathbf{R}_k}}{\sigma^2_{\mathbf{A}/\mathbf{R}_k}}$, where $p_{\mathbf{A}/\mathbf{R}_k}$ and

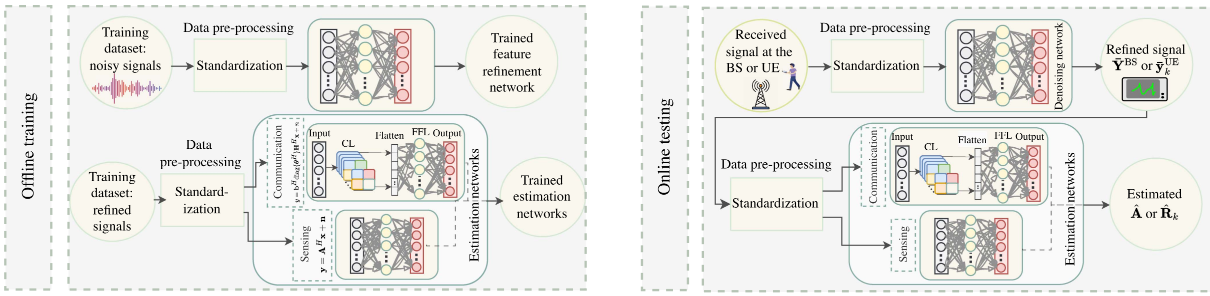


Fig. 2. Proposed estimation approach.

Algorithm 1: Proposed Estimation Approach.

Offline Training:

Stage 1:

Generate $(\mathcal{Z}_{S_1}^{\text{UE}_k}, \mathcal{L}_{S_1}^{\text{UE}_k})$ at the downlink UE and $(\mathcal{Z}_{S_1}^{BS}, \mathcal{L}_{S_1}^{BS})$ at the ISAC BS according to (30) and (31);

Pre-process $(\mathcal{Z}_{S_1}^{BS}, \mathcal{L}_{S_1}^{BS})$ and $(\mathcal{Z}_{S_1}^{\text{UE}_k}, \mathcal{L}_{S_1}^{\text{UE}_k})$ according to (20);

Initialize $\{\boldsymbol{\omega}_h, b_h\}, h \in \mathbb{N}_1^{N_h}$ randomly;

Compute $\mathbf{U}$ according to (9);

Compute $\hat{\boldsymbol{\Lambda}}$ according to (13);

Output the trained networks with $\hat{\boldsymbol{\Lambda}}_{\text{ELM}}^{\text{BS}}$ and $\hat{\boldsymbol{\Lambda}}_{\text{ELM}}^{\text{UE}_k}$ at the ISAC BS and downlink UE$_k$, respectively.

Stage 2:

Input the received signals to the trained ELM network in stage 1, obtaining $\bar{\mathbf{Y}}_c^{\text{BS}}$ and $\bar{\mathbf{y}}_c^{\text{UE}_k}$;

Use the received signals to formulate $(\mathcal{Z}_{S_2}^{\text{UE}_k}, \mathcal{L}_{S_2}^{\text{UE}_k})$ at the downlink UE and $(\mathcal{Z}_{S_2}^{BS}, \mathcal{L}_{S_2}^{BS})$ at the ISAC BS according to (30) and (31);

Pre-process $(\mathcal{Z}_{S_2}^{BS}, \mathcal{L}_{S_2}^{BS})$ and $(\mathcal{Z}_{S_2}^{\text{UE}_k}, \mathcal{L}_{S_2}^{\text{UE}_k})$ according to (20) and (26);

while $t \leq T_\epsilon$ or the validation accuracy increases over 5 sequential epochs **do**

  Input the samples to the designed CNN $f(\cdot; \Lambda_{\text{CNN}}^{\text{BS/UE}_k})$;

  Update Λ_{CNN} by minimizing $\mathcal{L}$ in (28);

  Update $t \leftarrow t + 1$;

end while

Output the trained CNN $f(\cdot; \hat{\Lambda}_{\text{CNN}}^{\text{BS/UE}_k})$;

Online Testing:

Input: Received signals $\mathbf{y}_{k,c}^{\text{UE}}$ and $\mathbf{Y}_c^{\text{BS}}$ that are used to generate the testing dataset $\mathcal{Z}^{BS}$ and $\mathcal{L}^{\text{UE}_k}$;

Refine the received signals using the trained ELM network in stage 1, obtaining $\bar{\mathbf{Y}}_c^{\text{BS}}$ and $\bar{\mathbf{y}}_c^{\text{UE}_k}$;

Use the refined signals to estimate $\mathbf{A}$ and $\mathbf{R}_k$ by using the trained S-CNN and C-CNN, respectively;

Output: Estimated channels $\hat{\mathbf{A}}, \hat{\mathbf{R}}_k$.

$\sigma^2_{\mathbf{A}/\mathbf{R}_k}$ represent the channel power and synthetic noise, respectively [36]. The addition of noise during training reflects practical deployment scenarios, where ground-truth labels are not perfectly available. This strategy not only enhances robustness but also encourages the model to learn generalized representations that remain effective under varying noise conditions. Through this augmentation strategy, the dataset can be formulated as

$$(\mathcal{Z}_{S_i}^{\text{UE}_k}, \mathcal{L}_{S_i}^{\text{UE}_k}) = \Big\{ \Big(\zeta_{S_i(1,1)}^{\text{UE}_k}, \gamma_{S_i(1)}^{\text{UE}_k}\Big), \Big(\zeta_{S_i(1,2)}^{\text{UE}_k}, \gamma_{S_i(1)}^{\text{UE}_k}\Big), \ldots, \Big(\zeta_{S_i(1,X)}^{\text{UE}_k}, \gamma_{S_i(1)}^{\text{UE}_k}\Big), \Big(\zeta_{S_i(2,1)}^{\text{UE}_k}, \gamma_{S_i(2)}^{\text{UE}_k}\Big), \ldots, \Big(\zeta_{S_i(Q,X)}^{\text{UE}_k}, \gamma_{S_i(Q)}^{\text{UE}_k}\Big) \Big\}, \tag{30}$$

and

$$(\mathcal{Z}_{S_i}^{\text{BS}}, \mathcal{L}_{S_i}^{\text{BS}}) = \Big\{ \Big(\zeta_{S_i(1,1)}^{\text{BS}}, \gamma_{S_i(1)}^{\text{BS}}\Big), \Big(\zeta_{S_i(1,2)}^{\text{BS}}, \gamma_{S_i(1)}^{\text{BS}}\Big), \ldots, \Big(\zeta_{S_i(1,X)}^{\text{BS}}, \gamma_{S_i(1)}^{\text{BS}}\Big), \Big(\zeta_{S_i(2,1)}^{\text{BS}}, \gamma_{S_i(2)}^{\text{BS}}\Big), \ldots, \Big(\zeta_{S_i(Q,X)}^{\text{BS}}, \gamma_{S_i(Q)}^{\text{BS}}\Big) \Big\}, \tag{31}$$

where $i \in \{1, 2\}$ denotes the stage number.

IV. Complexity Analysis

This section analyzes the computational complexity of the proposed two-stage estimation approach, which consists of a feature refinement stage followed by a channel estimation stage. The complexity of each stage is evaluated separately by considering the operations performed by the NNs during the estimation process. The analysis is derived based on the number of real multiplications, $C_{\mathcal{M}}$, and additions, $C_{\mathcal{A}}$, required by the NN to approximate sensing and communication channels. In the first stage, the ELM-based network is used to refine the received signal by extracting essential features. As ELM only contains one hidden layer, let η_I, η_h, and η_O denote the number of neurons of the input, hidden, and output layers, respectively. Therefore, the number of real multiplications and additions is respectively expressed as

$$C_{\mathcal{M}}^{\chi-\text{ELM}} = \eta_h^\chi(\eta_I^\chi + \eta_O^\chi), \tag{32}$$

$$C_{\mathcal{A}}^{\chi-\text{ELM}} = \eta_h^\chi(\eta_I^\chi + \eta_O^\chi + 1) + \eta_O^\chi, \tag{33}$$

where $\chi \in \{\text{S}, \text{C}\}$ denotes the network task (i.e., sensing and communication) and is based on the parameters summarized in Table I.

In the second stage, the refined signal is used to estimate the desired channels using dedicated NNs for sensing and communication links. Different models are designed based on the estimation task as detailed in Section III-B, where a fully connected network is designed for the sensing task and a CNN-based network is designed for the communication task. This stage involves additional layers of processing to accurately estimate the channel parameters. To this end, based on the parameters provided in Table II, the computational complexity associated with the first FF layer (FFL) in the S-EST network is $\eta_1\eta_2$ real multiplications and $\eta_1\eta_2 + \eta_2$ real additions, where η_i represents the number of neurons of the i-th layer. Accordingly, the required number of multiplications and additions by the S-EST network are respectively given as

$$C_{\mathcal{M}}^{\text{S-EST}} = \sum_{i=1}^{2} \eta_i^{\text{S}} \eta_{i+1}^{\text{S}}, \tag{34}$$

$$C_{\mathcal{A}}^{\text{S-EST}} = \sum_{i=1}^{2} \eta_i^{\text{S}} \eta_{i+1}^{\text{S}} + \sum_{i=1}^{2} \eta_{i+1}^{\text{S}}. \tag{35}$$

In contrast, the C-EST network includes a convolutional layer (CL), FFL, input and output layers. The output size of the first CL is represented as

$$\begin{aligned} \eta_F &= \left\lfloor \frac{\eta_1^{\text{C}} - F_z}{F_s} + 1 \right\rfloor \\ &= \left\lfloor \frac{2PC - F_z}{F_s} + 1 \right\rfloor, \end{aligned} \tag{36}$$

where $\lfloor \cdot \rfloor$ is the floor operation, and F_z and F_s denote the filter size and stride, respectively. As the flatten layer between the CL and FFL does not involve addition or multiplication operations, the computational complexity of C-EST is calculated by the contribution of all the other layers. The number of required multiplications is formulated as $F_z\eta_F F_n$, $\eta_3\eta_F F_n$, and $\eta_3\eta_4$, respectively, where F_n is the number of filters. Moreover, the number of additions is expressed as $(F_z + 1)\eta_F F_n$, $\eta_3(\eta_F F_n + 1)$, and $(\eta_4 + 1)\eta_3$, respectively. Therefore, the overall computational complexity of the C-EST framework is presented as

$$C_{\mathcal{M}}^{\text{C-EST}} = (F_z + \eta_3^{\text{C}})\eta_F F_n + \eta_3^{\text{C}}\eta_4^{\text{C}} \tag{37}$$

$$C_{\mathcal{A}}^{\text{C-EST}} = (F_z + \eta_3^{\text{C}} + 1)\eta_F F_n + (\eta_4^{\text{C}} + 1)\eta_3^{\text{C}} + \eta_4^{\text{C}}. \tag{38}$$

Consequently, the total number of multiplication and addition operations of the proposed framework for the sensing task is given as

$$\begin{aligned} C_{\mathcal{M}}^{\text{Total - S}} &= C_{\mathcal{M}}^{\text{S-ELM}} + C_{\mathcal{M}}^{\text{S-EST}} \\ &= \eta_h^{\text{S}}(\eta_I^{\text{S}} + \eta_O^{\text{S}}) + \sum_{i=1}^{2} \eta_i^{\text{S}}\eta_{i+1}^{\text{S}} \\ &= 4\eta_h^{\text{S}} MP + \eta_2^{\text{S}}(2MP + 2M^2), \end{aligned} \tag{39}$$

and

$$\begin{aligned} C_{\mathcal{A}}^{\text{Total - S}} &= C_{\mathcal{A}}^{\text{S-ELM}} + C_{\mathcal{A}}^{\text{S-EST}} \\ &= \eta_h^{\text{S}}(\eta_I^{\text{S}} + \eta_O^{\text{S}} + 1) + \eta_O^{\text{S}} + \sum_{i=1}^{2} \eta_i^{\text{S}}\eta_{i+1}^{\text{S}} + \sum_{i=1}^{2} \eta_{i+1}^{\text{S}} \\ &= \eta_h^{\text{S}}(4MP + 1) + 2MP + \eta_2^{\text{S}}(2MP + 2M^2) \\ &\quad + \eta_2^{\text{S}} + 2M^2. \end{aligned} \tag{40}$$

Similarly, the total number of multiplications and additions required by the proposed framework for the communication task is given as

$$\begin{aligned} C_{\mathcal{M}}^{\text{Total - C}} &= C_{\mathcal{M}}^{\text{C-ELM}} + C_{\mathcal{M}}^{\text{C-EST}} \\ &= \eta_h^{\text{C}}(\eta_I^{\text{C}} + \eta_O^{\text{C}}) + (F_z + \eta_3^{\text{C}})\eta_F F_n + \eta_3^{\text{C}}\eta_4^{\text{C}} \\ &= 4\eta_h^{\text{C}} PC + (F_z + \eta_3^{\text{C}})\eta_F F_n + \eta_3^{\text{C}} 2MN, \end{aligned} \tag{41}$$

and

$$\begin{aligned} C_{\mathcal{A}}^{\text{Total - C}} &= C_{\mathcal{A}}^{\text{C-ELM}} + C_{\mathcal{A}}^{\text{C-EST}} \\ &= \eta_h^{\text{C}}(\eta_I^{\text{C}} + \eta_O^{\text{C}} + 1) + \eta_O^{\text{C}} + (F_z + \eta_3^{\text{C}} + 1)\eta_F F_n \\ &\quad + (\eta_4^{\text{C}} + 1)\eta_3^{\text{C}} + \eta_4^{\text{C}} \\ &= \eta_h^{\text{C}}(4PC + 1) + 2PC + (F_z + \eta_3^{\text{C}} + 1)\eta_F F_n \\ &\quad + (2MN + 1)\eta_3^{\text{C}} + 2MN. \end{aligned} \tag{42}$$

V. Simulation Results

This section presents comprehensive simulation results to assess the performance of the proposed two-stage estimation approach. The simulations are designed to evaluate the effectiveness of both the signal refinement and channel estimation stages under various conditions. The performance of the proposed method is compared to the literature approaches to demonstrate its superiority in terms of accuracy and computational efficiency. The simulation parameters, evaluation metrics, and test cases are detailed in the following subsections.

A. Simulation Parameters

In the following simulation results, consider $K = 3$, $N = 30$, and $M = 4$ unless otherwise stated. The sensing channel is modeled as in [47], [48] referring to the radar channel model, expressed as

$$\mathbf{A} = \alpha \mathbf{a}(\theta)\mathbf{a}(\theta)^H, \tag{43}$$

where α represents the reflection coefficient related to the target, with phase shifts uniformly distributed within the range $[0, 2\pi)$. The corresponding steering vector, $\mathbf{a}(\theta)$, is defined as

$$\mathbf{a}(\theta) = \left[1, e^{j\frac{2\pi d}{\lambda}\sin(\theta)}, \ldots, e^{j\frac{2\pi d(M-1)}{\lambda}\sin(\theta)}\right]^T. \tag{44}$$

In this context, θ, d, and λ represent the azimuth angle, the antenna spacing at the BS, and the signal wavelength, respectively, with θ set to $-\frac{2\pi}{3}$. Meanwhile, the communication channels follow a Rician distribution, presented as

$$\mathbf{c} = \sqrt{PL}\left(\frac{K_1}{K_1 + 1}\bar{\mathbf{c}} + \frac{1}{K_1 + 1}\tilde{\mathbf{c}}\right), \tag{45}$$

where $\mathbf{c}$ denotes the communication channels, including $\mathbf{H}$ and $\mathbf{b}_k$, while K_1 indicates the Rician factor. The line-of-sight (LoS) components are represented by $\bar{\mathbf{h}}$, and defined as $\bar{\mathbf{c}} = \mathbf{a}(\dot{\theta})\mathbf{a}(\bar{\theta})^H$, resembling the structure of (44). In this case, $\mathbf{a}(\dot{\theta})$ and $\mathbf{a}(\bar{\theta})$ correspond to the departure and arrival angles, respectively, both set to $\frac{\pi}{3}$. The non-LoS components, $\tilde{\mathbf{c}}$, consist of independent and identically distributed elements following $\mathcal{CN}(0, 1)$. The Rician factor is set to $K_1 = 10$ for $\mathbf{H}$ and $K_1 = 0$ for $\mathbf{b}_k$, the latter representing Rayleigh fading [40].

The path loss (PL) model is defined by the expression $\text{PL} = \text{PL}_r(\frac{d_j}{d_r})^{-\epsilon_j}$, where d_j represents the link distance, PL_r represents the reference path loss at distance d_r, and ϵ_j indicates the path loss exponent. For simulation purposes, the reference path loss is set to $\text{PL}_r = -30\,\text{dBm}$ at $d_r = 1\,\text{m}$. The PL exponents for the BS-target-BS, BS-RIS, and RIS-UE$_k$ links are set to $\epsilon_1 = 3$, $\epsilon_2 = 2.3$, and $\epsilon_3 = 2$, respectively. The corresponding distances are defined as $d_1 = 140\,\text{m}$, $d_2 = 50\,\text{m}$, and $d_3 = 2\,\text{m}$. Following the approach in [36], [49], the ISAC BS transmit power is configured as 20 dBm.

The dataset size is configured to $Q \times X = 10^4$ for each SNR level, with $Q = 1000$ and $X = 10$. We allocate 90% of the data for training and 10% for testing. During the training phase, the SNR values of the generated data are set as $\text{SNR} = 10 : 5 : 20\,\text{dB}$, while the testing phase uses an extended rage of SNR values $\text{SNR} = -10 : 2.5 : 20\,\text{dB}$. This selection of the SNR range ensures that the model encounters not only unseen samples but also conditions outside its training experience. This approach demonstrates the model generalization capabilities, eliminating the necessity to estimate the SNR prior to channel estimation stage in practical scenarios.

To assess the performance of the proposed approach, the normalized MSE (NMSE) is utilized as the primary evaluation metric, defined as

$$\text{NMSE} = \mathbb{E}\left\{ \frac{\left\| \mathbf{c}_{\text{Estimated}} - \mathbf{c}_{\text{True}} \right\|_F^2}{\left\| \mathbf{c}_{\text{True}} \right\|_F^2} \right\}, \tag{46}$$

where $\|\cdot\|_F$ is the F-norm of a matrix. The works in [36] and [38] are adopted as benchmarks to effectively evaluate the performance of the proposed algorithm in the subsequent subsections. This comparative analysis provides a comprehensive understanding of the advantages and potential improvements offered by the proposed estimation approach.

B. Impact of Varying SNR

Fig. 3 shows the estimation performance under various SNR conditions, where Fig. 3(a) illustrates the performance of the feature refinement stage and Fig. 3(b) demonstrates the overall estimation performance for both sensing and communication links. In Fig. 3(a), the focus is solely on refining the received signal without performing actual channel estimation. The baseline in this context assumes the received signal itself, without any denoising process, while the ELM-based approach applies feature refinement to suppress noise. The NMSE here indicates the difference between the refined signal and the ideal signal, providing a measure of how closely the processed signal approximates the received signal. As expected, the performance of both the baseline and the ELM-based approach improves with higher SNRs, since the noise is reduced. However, it is evident that the ELM-based approach achieves superior performance compared to the baseline. The ELM-based refinement effectively enhances the signal quality, as shown by the lower NMSE values across all SNRs. This indicates that the proposed method can extract essential features and mitigate noise, setting a solid foundation for the subsequent channel estimation stage.

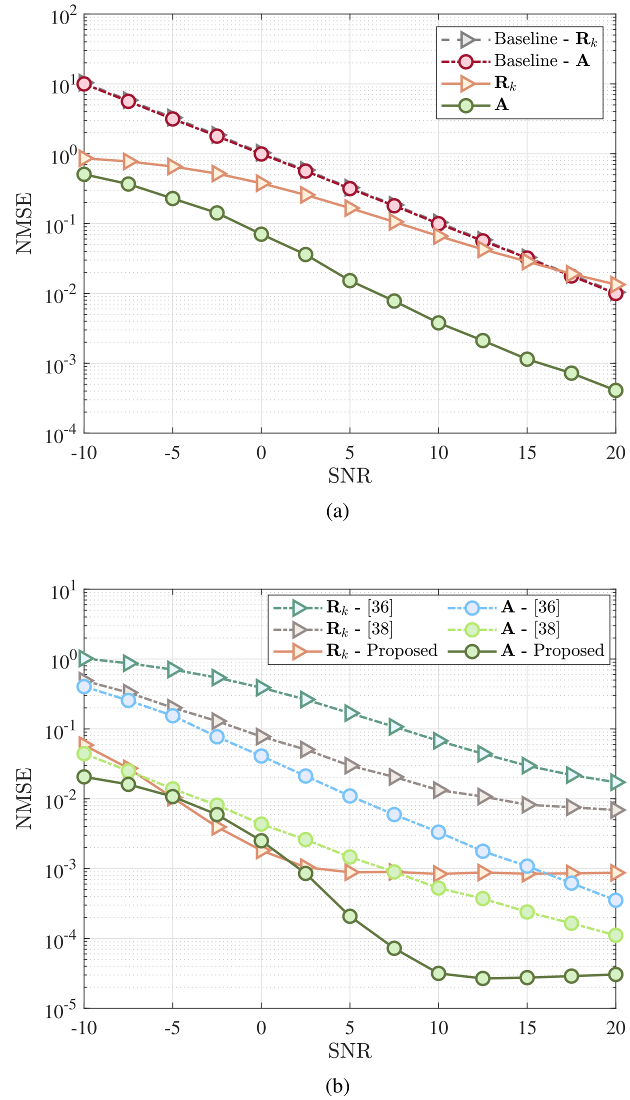


Fig. 3. Estimation performance versus SNR: (a) Stage 1, (b) Stage 2.

Fig. 3(b) shows that the proposed algorithm demonstrates a significant improvement compared to the baseline estimation approach presented in [36] and [38]. For the sensing link estimation, the performance of the proposed method is comparable to that of the benchmark approach in [38] in the low SNR range (i.e., for $\text{SNR} = -10 : 0\,\text{dB}$). However, as the SNR increases, the proposed method exhibits substantial improvements. This performance gain can be partially attributed to the fact that the SNR range used in training was between 10 and 20 dB, enabling the model to perform well at higher SNR values. Additionally, the augmented performance can be attributed to the effective feature extraction capability of the first stage, which significantly enhances the signal quality and provides a refined input for the second stage. It is worth noting that the communication channel estimation, $\mathbf{R}_k$, is inherently more complex due to the cascaded relationship illustrated in (14). However, the proposed algorithm achieves a remarkable performance, outperforming the benchmark approaches across all SNR levels. This demonstrates the efficacy of the two-stage estimation approach in handling

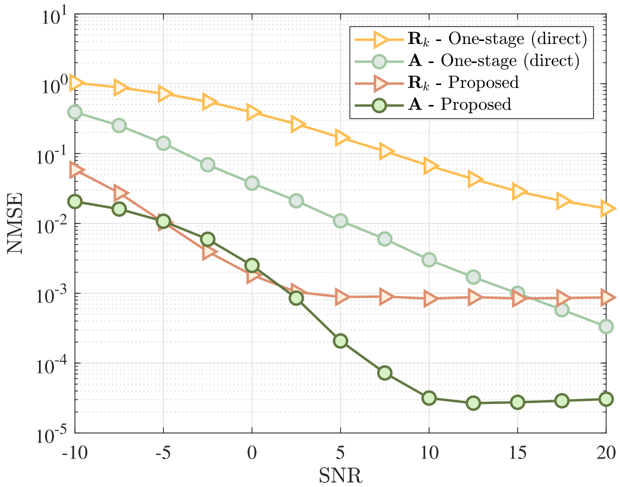


Fig. 4. Estimation performance of the proposed two-stage framework versus a direct one-stage approach for $\mathbf{A}$ and $\mathbf{R}_k$.

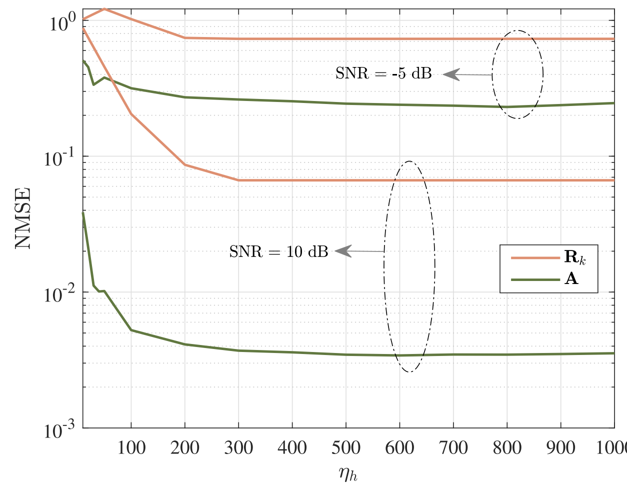


Fig. 5. NMSE performance for the first stage versus η_h.

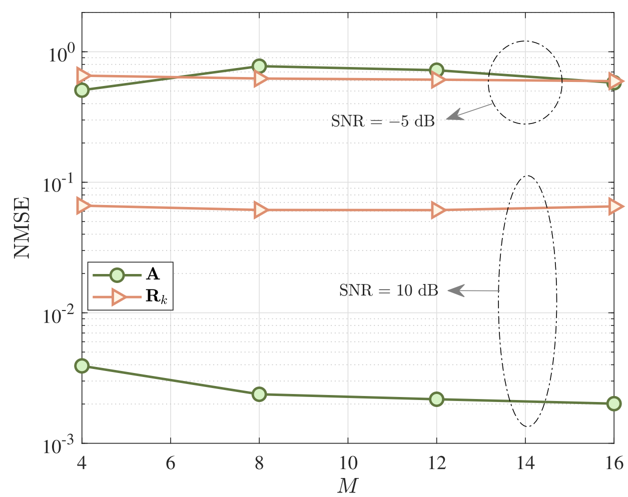


Fig. 6. NMSE performance for the first stage versus M at $N = 30$.

the complexity of cascaded channels. The figure further highlights the generalization performance of the proposed approach, demonstrating its superior performance compared to benchmark methods, even within SNR ranges excluded from the training phase.

To further assess the contribution of the feature refinement stage, Fig. 4 presents a comparison between the proposed two-stage framework and a direct one-stage estimation approach. In the latter, the same networks used in the second stage of the proposed method (i.e., S-EST and C-EST) are directly applied to estimate the channels from the noisy input signals, without the benefit of the ELM-based denoising stage. All hyperparameters and training settings are kept identical to ensure a fair comparison. As can be seen, the proposed two-stage approach achieves consistently superior performance across all SNR levels for both the sensing channel, $\mathbf{A}$, and the communication channel, $\mathbf{R}_k$. This result confirms that the two-stage design not only improves estimation accuracy but also enhances robustness under practical noise conditions.

C. Impact of Varying η_h

Fig. 5 shows the NMSE performance of the ELM-based feature refinement stage as a function of the number of hidden neurons, η_h, for both the cascaded communication channel, $\mathbf{R}_k$, and the sensing channel, $\mathbf{A}$. The results are shown at -5 dB and 10 dB. As η_h increases, the ELM approximation capability improves, resulting in progressively lower NMSE values. In the low neuron range, fluctuations in NMSE are observed due to limited model capacity. As the number of hidden neurons increases, the performance stabilizes and eventually saturates, indicating convergence. For the sensing channel, $\mathbf{A}$, convergence is observed around $\eta_h = 200$, whereas for the more complex communication channel, $\mathbf{R}_k$, convergence occurs at a higher value, approximately $\eta_h = 400$. This behavior is expected, as estimating the cascaded channel structure in the communication scenario is inherently more challenging and requires a higher-capacity model for effective denoising. Moreover, it is evident from the figure that for both sensing and communication channels, the convergence is slower at lower SNR levels. For instance, at -5 dB, the performance saturates more gradually compared to the higher SNR case (e.g., 10 dB). This is consistent with expectations, since higher noise levels require a larger number of neurons to extract meaningful features and suppress noise effectively. Based on these observations, the hidden layer sizes for the ELM modules are selected as $\eta_h^{\mathrm{S}} = 200$ and $\eta_h^{\mathrm{C}} = 900$ for the for the S-ELM and C-ELM, respectively. The larger setting for C-ELM is chosen to accommodate larger-scale scenarios (i.e., larger configurations of M and N). These configurations ensure that the ELM modules operate in the convergence regime, thereby enabling stable feature refinement performance, as further evidenced in the subsequent simulation results.

D. Impact of Varying M

Fig. 6 examines the performance of the feature refinement as M changes at -5 dB and 10 dB SNR, respectively. As illustrated, the NMSE performance remains relatively stable across

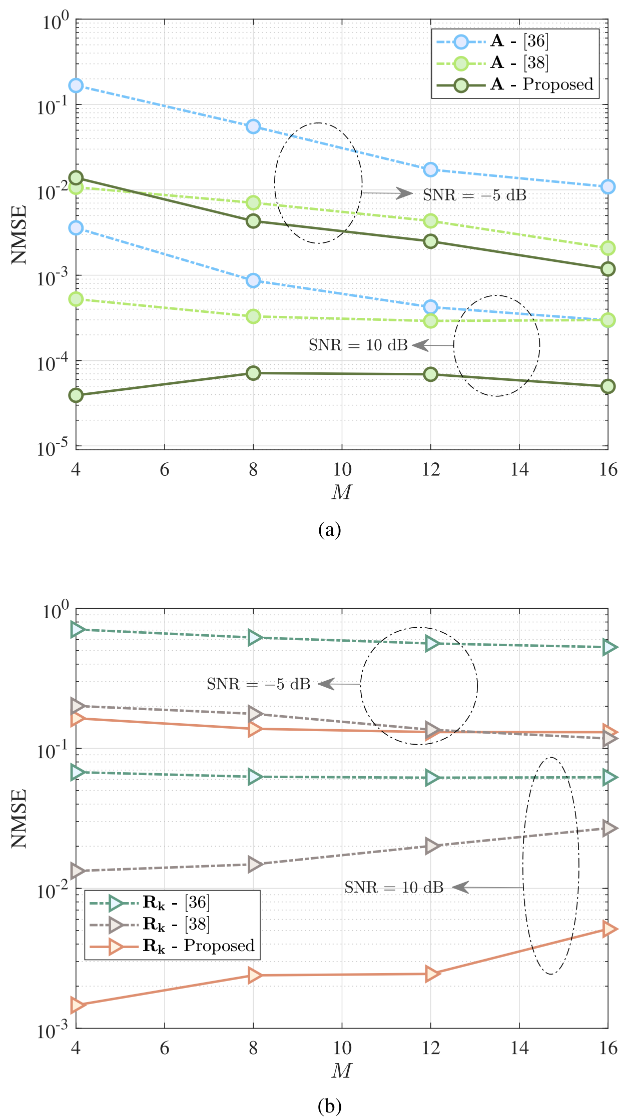


Fig. 7. NMSE performance for the second stage versus M at $N = 30$: (a) Sensing channel, (b) Communication channel.

different values of M, indicating that the feature refinement stage effectively mitigates noise even at a larger scale. This stability shows the robustness of the proposed ELM-based approach in handling noise and maintaining good performance regardless of the number of antennas. At SNR $= 10\,$dB, the performance of the proposed method is further enhanced. The NMSE decreases as the number of antennas increases, indicating that the ELM-based model benefits from the additional antennas by providing more accurate signal enhancement.

Fig. 7 analyzes the performance of the channel estimation stage for both sensing and communication links as M increases under the same SNR conditions as in Fig. 6. Fig. 7(a) demonstrates the performance of the proposed approach in estimating the sensing channel, $\mathbf{A}$. At $-5\,$dB SNR, the NMSE of the proposed approach decreases as the number of antennas increases and maintains a stable performance. While the proposed approach outperformes the performance presented in [36], it has a comparable performance to the benchmark in [38]. This was also shown in Fig. 3(b), where the proposed algorithm achieves similar performance to the CGAN-based approach in [38] in the low SNR region, up to SNR $= 0\,$dB. However, at SNR $= 10\,$dB, the proposed approach significantly enhances the performance when compared to both benchmark methods. In particular, at $M = 16$, the proposed approach achieves an NMSE of approximately 5×10^{-5}, while the benchmark methods in [36] and [38] achieve NMSE values of around 2.9×10^{-4}, respectively. This trend proves that the proposed method consistently outperforms the benchmark methods and remains stable even when increasing the channel dimensions. This is attributed to the effectiveness of the feature refinement and estimation stages in enhancing the quality of received signals and estimation accuracy.

Fig. 7(b) studies the performance of the proposed framework in estimating the communication channel, $\mathbf{R}_k$, while varying M under two SNR values. As can be seen, at SNR $= -5\,$dB, the proposed algorithm outperforms the estimation performance in [36], but exhibits comparable performance to the benchmark in [38]. This feature highlights the robustness of the proposed model, showing its capability to handle noisy measurements as effectively as the CGAN network proposed in [38]. While CGAN works intelligently to generate data and manage noise, regular NNs typically struggle in such conditions. However, our two-stage approach, incorporating feature refinement followed by channel estimation, enables similar high performance by effectively denoising and accurately estimating channels, thus matching the strengths of [38] without the added complexity of a discriminator network and the intricate training process. It is also worth noting that the estimation performance enhances as M increases, showcasing its ability in generalizing to larger-scale systems. On the other hand, at SNR $= 10\,$dB, the proposed estimation approach significantly outperforms both benchmark methods. For instance, at $M = 12$, the proposed approach achieves an NMSE of approximately 2.4×10^{-3}, while the benchmark methods in [36] and [38] achieve NMSE values of around 6.1×10^{-2} and 2×10^{-2}, respectively. There is a gradual increase in NMSE of the proposed algorithm as M increases. This phenomenon can be attributed to the complexities associated with capturing high-dimensional dependencies in the data. This is particularly evident as the network architecture and optimization have been primarily tuned for small channel dimensions, where the number of antennas and RIS elements are set to $M = 4$ and $N = 30$, respectively. Despite this, the overall performance of the proposed method consistently surpasses that of both benchmark methods. The significant improvement observed, even with the slight increase in NMSE at higher SNRs, highlights the robustness and effectiveness of our two-stage approach in denoising and accurately estimating channels across various SNR levels.

E. Impact of Varying N

Fig. 8 illustrates the performance of the proposed approach in estimating the communication channel, $\mathbf{R}_k$, as the number of reflecting elements, N, varies, under the same SNR conditions as in Fig. 6. At an SNR of $-5\,$dB, the proposed algorithm

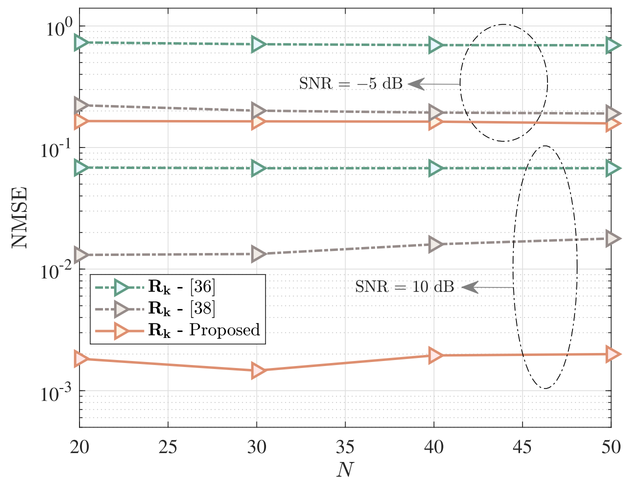
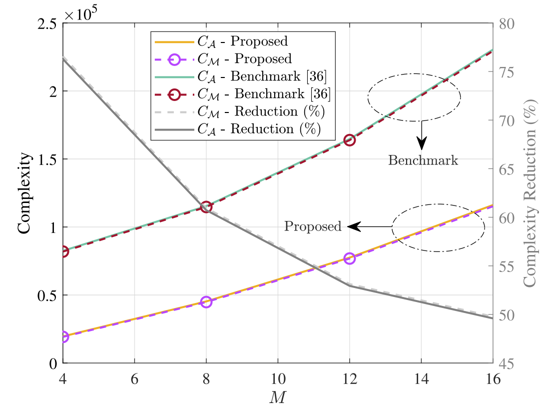
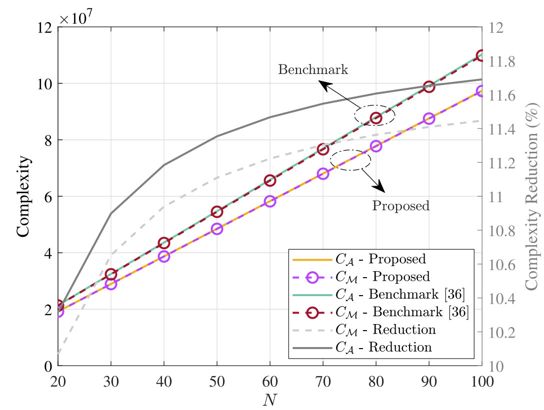


Fig. 8. NMSE performance for the communication channel estimation versus N at $M = 4$.



(a)



(b)

Fig. 9. Computational complexity evaluation: (a) Sensing link, (b) Communication link.

shows slight enhancement over the benchmark method in [38]. The NMSE for the proposed algorithm remains stable even as the number of RIS elements increases. This demonstrates the robustness of the proposed method in low SNR conditions, effectively mitigating the noise and providing accurate channel estimates. It is worth noting that one of the main advantages of the proposed CGAN approach in [38] is its ability to generate accurate estimates from noisy data (i.e., at low SNR) by leveraging the generative process. However, our approach mimics this beneficial behavior while avoiding the added complexity of the discriminator network, thus offering a more efficient solution. Alternatively, at an SNR of 10 dB, the proposed algorithm shows significant improvement over the benchmark methods. This indicates that it can leverage the higher SNR to extract essential features, resulting in much lower NMSE values in the estimation stage across various SNR ranges. Overall, the proposed algorithm performance proves its capability in handling complex relationships and larger state spaces, showcasing its robustness and effectiveness across varying numbers of RIS reflecting elements.

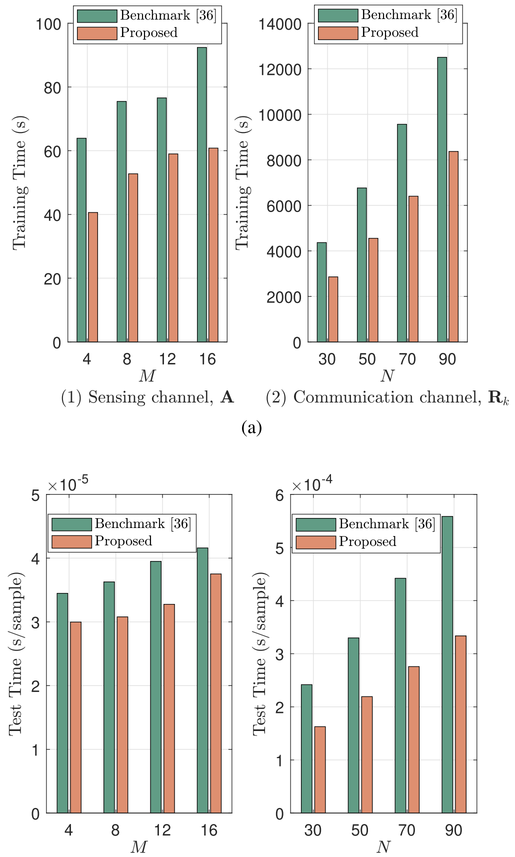
F. Complexity Evaluation

Fig. 9(a) and (b) compare the computational complexity of the proposed method with the state-of-the-art approach in [36]. It is worth noting that the work in [38] achieves a similar complexity to [36]. Therefore, for clarity of presentation, the results focus on a comparison with [36]. The analysis is based on the number of real additions and multiplications needed for each method. Furthermore, these figures shows the complexity reduction of using the proposed two-stage algorithm over the benchmark scheme, which is expressed as

$$\text{Reduction} = \frac{C_{\chi}^{[36]} - C_{\chi}^{\text{Proposed}}}{C_{\chi}^{[36]}}, \ \chi \in \{\mathcal{A}, \mathcal{M}\}. \tag{47}$$

Fig. 9(a) shows the computational complexity and reduction percentage of approximating $\mathbf{A}$ as M increases. It can be seen that the proposed approach shows a significant reduction compared to the benchmark. In particular, the proposed method achieves a reduction ranging from approximately 75% to 50% as the number of antennas increases. This considerable reduction in complexity, particularly maintaining a reduction close to 50% even for larger antenna arrays, highlights the efficiency and scalability of the proposed approach. It is worth noting the while the proposed algorithm realizes a substantial reduction in complexity, it achieves a significant improvement in performance as shown in Fig. 3.

Furthermore, Fig. 9(b) illustrates the computational complexity and reduction percentage of estimating $\mathbf{R}_k$ as N increases. It can also be seen here that the proposed method consistently reduces the computational complexity compared to the benchmark. Specifically, the complexity reduction for real additions ranges from approximately 10.5% to 11.8%, while for multiplications, it ranges from around 10.2% to 11.6%. It can be observed that the complexity reduction in communication channel estimation is moderate. This is due to the inherent challenges of estimating communication channels, which typically require more sophisticated models. Despite these challenges, the proposed approach manages to maintain a complexity comparable to the benchmark while considerably outperforming it in terms of estimation performance, proving its effectiveness and robustness under various practical system configurations.



(1) Sensing channel, $\mathbf{A}$ (2) Communication channel, $\mathbf{R}_k$

(a)

(1) Sensing channel, $\mathbf{A}$ (2) Communication channel, $\mathbf{R}_k$

(b)

Fig. 10. Runtime complexity evaluation: (a) Training time, (b) Testing time, for $\mathbf{A}$ and $\mathbf{R}_k$.

In addition to the theoretical analysis of computational complexity, Fig. 10 presents the empirical runtime evaluation of the proposed framework compared to the benchmark approach in terms of training and testing times. As shown in Fig. 10(a), the proposed method achieves consistently lower training time than the benchmark across different values of M and N. For example, when $M = 16$, the training time for estimating the sensing channel $\mathbf{A}$ is approximately 60 seconds using the proposed approach, compared to over 90 seconds for the benchmark. Similarly, for the communication channel $\mathbf{R}_k$, the proposed method reduces the training time from about 12,600 seconds to 8,300 seconds when $N = 90$. These improvements are primarily attributed to the architectural efficiency of the proposed two-stage design and the lightweight networks used in both stages.

Fig. 10(b) shows the testing time per sample, which corresponds to the real-time inference phase. As expected, the testing times are significantly lower than training times. The proposed method continues to show a consistent advantage. For instance, at $M = 16$, the test time for sensing is approximately 3.6×10^{-5} seconds per sample, compared to 4.2×10^{-5} seconds in the benchmark. Similarly, for the communication channel with $N = 90$, the proposed framework reduces inference time from roughly 5.7×10^{-4} to 3.4×10^{-4} seconds per sample. These findings complement the computational complexity reductions and confirm that the proposed method not only achieves lower computational complexity but also results in faster practical execution, making it suitable for practical RIS-assisted ISAC systems.

VI. Conclusion

This paper introduced a novel two-stage approach for enhancing the channel estimation performance of RIS-assisted ISAC systems. The first stage focuses on feature refinement, where an ELM-based NN is employed to enhance the quality the received signals, effectively extracting essential channel features. In the second stage, the refined signals are used to estimate the desired channels with dedicated NNs tailored for sensing and communication tasks. The proposed method demonstrated significant improvements in channels estimation accuracy, outperforming existing benchmark methods across various SNR conditions. In terms of computational complexity, the proposed two-stage approach archived a remarkable reduction, ranging between 50% and 75% for the sensing channel, and between 10.5% and 11.8% for the communication channel. These results highlighted the efficiency and robustness of the proposed approach, making it a potential solution for improving the robustness and complexity of RIS-assisted ISAC systems. Future works could explore extensions of the proposed channel estimation approach to address more practical challenges. These include accommodating multi-target scenarios, user and target mobility, and the use of real channel measurement data to validate the model under practical propagation conditions. Additionally, transfer learning techniques can be integrated to adapt pre-trained models to new environments using limited real-world data, thereby reducing retraining costs.

References

[1] Z. Wang, Y. Zhao, Y. Zhou, Y. Shi, C. Jiang, and K. B. Letaief, "Over-the-air computation for 6G: Foundations, technologies, and applications," *IEEE Internet Things J.*, vol. 11, no. 14, pp. 24634–24658, Jul. 2024.

[2] L. Blanco, E. Zeydan, S. Barrachina-Muñoz, F. Rezazadeh, L. Vettori, and J. Mangues-Bafalluy, "A novel approach for scalable and sustainable 6G networks," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1673–1692, 2024.

[3] M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Y. Shiu, and P. Wang, "Machine learning-based self-interference cancellation for full-duplex radio: Approaches, open challenges, and future research directions," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 21–47, 2024.

[4] A. Kaushik et al., "Toward integrated sensing and communications for 6G: Key enabling technologies, standardization, and challenges," *IEEE Commun. Stand. Mag.*, vol. 8, no. 2, pp. 52–59, Jun. 2024.

[5] W. Yang et al., "Integrated sensing and communication channel modeling and measurements: Requirements and methodologies toward 6 G standardization," *IEEE Veh. Technol. Mag.*, vol. 19, no. 2, pp. 22–30, Jun. 2024.

[6] Z. He, W. Xu, H. Shen, D. W. K. Ng, Y. C. Eldar, and X. You, "Full-duplex communication for ISAC: Joint beamforming and power optimization," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2920–2936, Sep. 2023.

[7] Q. Nhat Le, V. -D. Nguyen, O. A. Dobre, and H. Shin, "RIS-assisted full-duplex integrated sensing and communication," *IEEE Wireless Commun. Lett.*, vol. 12, no. 10, pp. 1677–1681, Oct. 2023.

[8] J. Ye, L. Huang, Z. Chen, P. Zhang, and M. Rihan, "Unsupervised learning for joint beamforming design in RIS-aided ISAC systems," *IEEE Wireless Commun. Lett.*, vol. 13, no. 8, pp. 2100–2104, Aug. 2024.

[9] F. Dong, F. Liu, Y. Cui, W. Wang, K. Han, and Z. Wang, "Sensing as a service in 6G perceptive networks: A unified framework for ISAC resource allocation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3522–3536, May 2023.

[10] B. Zhao, M. Wang, Z. Xing, G. Ren, and J. Su, “Integrated sensing and communication aided dynamic resource allocation for random access in satellite terrestrial relay networks,” *IEEE Commun. Lett.*, vol. 27, no. 2, pp. 661–665, Feb. 2023.

[11] C. Qing, W. Hu, Z. Liu, G. Ling, X. Cai, and P. Du, “Sensing-aided channel estimation in OFDM systems by leveraging communication echoes,” *IEEE Internet Things J.*, vol. 11, no. 23, pp. 38023–38039, Dec. 2024.

[12] C. Qing, Z. Liu, W. Hu, Y. Zhang, X. Cai, and P. Du, “LoS sensing-based channel estimation in UAV-assisted OFDM systems,” *IEEE Wireless Commun. Lett.*, vol. 13, no. 5, pp. 1320–1324, May 2024.

[13] C. Qing, Z. Liu, G. Ling, W. Hu, and P. Du, “Channel estimation in OTFS systems by leveraging differential modulation,” *IEEE Trans. Veh. Technol.*, vol. 74, no. 5, pp. 6907–6918, May 2025.

[14] G. C. Trichopoulos et al., “Design and evaluation of reconfigurable intelligent surfaces in real-world environment,” *IEEE Open J. Commun. Soc.*, vol. 3, pp. 462–474, 2022.

[15] M. Ahmed et al., “A survey on STAR-RIS: Use cases, recent advances, and future research challenges,” *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14689–14711, Aug. 2023.

[16] H. Zhou, M. Erol-Kantarci, Y. Liu, and H. V. Poor, “A survey on model-based, heuristic, and machine learning optimization approaches in RIS-aided wireless networks,” *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 781–823, Secondquarter 2024.

[17] Z. Ding et al., “A state-of-the-art survey on reconfigurable intelligent surface-assisted non-orthogonal multiple access networks,” *Proc. IEEE*, vol. 110, no. 9, pp. 1358–1379, Sep. 2022.

[18] R. Kaur, B. Bansal, S. Majhi, S. Jain, C. Huang, and C. Yuen, “A survey on reconfigurable intelligent surface for physical layer security of next-generation wireless communications,” *IEEE Open J. Veh. Technol.*, vol. 5, pp. 172–199, 2024.

[19] C. -X. Wang et al., “On the road to 6G: Visions, requirements, key technologies, and testbeds,” *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, Secondquarter 2023.

[20] X. Zhu et al., “Enabling intelligent connectivity: A survey of secure ISAC in 6G networks,” *IEEE Commun. Surveys Tuts.*, vol. 27, no. 2, pp. 748–781, Apr. 2025.

[21] Z. Liu, Y. Liu, S. Shen, Q. Wu, and Q. Shi, “Enhancing ISAC network throughput using beyond diagonal RIS,” *IEEE Wireless Commun. Lett.*, vol. 13, no. 6, pp. 1670–1674, Jun. 2024.

[22] K. Chen, C. Qi, O. A. Dobre, and G. Y. Li, “Simultaneous beam training and target sensing in ISAC systems with RIS,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 2696–2710, Apr. 2024.

[23] X. Wang, Z. Fei, Z. Zheng, and J. Guo, “Joint waveform design and passive beamforming for RIS-assisted dual-functional radar-communication system,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 5131–5136, May 2021.

[24] X. Wang, Z. Fei, J. Huang, and H. Yu, “Joint waveform and discrete phase shift design for RIS-assisted integrated sensing and communication system under Cramer-Rao bound constraint,” *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 1004–1009, Jan. 2022.

[25] Z. Wu, X. Li, Y. Cai, and W. Yuan, “Joint trajectory and resource allocation design for RIS-assisted UAV-enabled ISAC systems,” *IEEE Wireless Commun. Lett.*, vol. 13, no. 5, pp. 1384–1388, May 2024.

[26] Y. Wang et al., “Optimizing the fairness of STAR-RIS and NOMA assisted integrated sensing and communication systems,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 6, pp. 5895–5907, Jun. 2024.

[27] P. Saikia, K. Singh, W. -J. Huang, and T. Q. Duong, “Hybrid deep reinforcement learning for enhancing localization and communication efficiency in RIS-aided cooperative ISAC systems,” *IEEE Internet Things J.*, vol. 11, no. 18, pp. 29494–29510, Sep. 2024.

[28] K. Zhong, J. Hu, C. Pan, M. Deng, and J. Fang, “Joint waveform and beamforming design for RIS-aided ISAC systems,” *IEEE Signal Process. Lett.*, vol. 30, pp. 165–169, 2023.

[29] G. Sun, Y. Zhang, W. Hao, Z. Zhu, X. Li, and Z. Chu, “Joint beamforming optimization for STAR-RIS aided NOMA ISAC systems,” *IEEE Wireless Commun. Lett.*, vol. 13, no. 4, pp. 1009–1013, Apr. 2024.

[30] Z. Yu et al., “Active RIS-aided ISAC systems: Beamforming design and performance analysis,” *IEEE Trans. Commun.*, vol. 72, no. 3, pp. 1578–1595, Mar. 2024.

[31] Q. Li, M. El-Hajjar, and L. Hanzo, “Ergodic spectral efficiency analysis of intelligent omni-surface aided systems suffering from imperfect CSI and hardware impairments,” *IEEE Trans. Commun.*, vol. 72, no. 8, pp. 5073–5086, Aug. 2024.

[32] Q. Li et al., “Achievable rate analysis of the STAR-RIS-aided NOMA uplink in the face of imperfect CSI and hardware impairments,” *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 6100–6114, Oct. 2023.

[33] Q. Li, M. El-Hajjar, I. Hemadeh, D. Jagyasi, A. Shojaeifard, and L. Hanzo, “Performance analysis of active RIS-aided systems in the face of imperfect CSI and phase shift noise,” *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8140–8145, Jun. 2023.

[34] C. Qing, L. Dong, L. Wang, J. Wang, and C. Huang, “Joint model and data-driven receiver design for data-dependent superimposed training scheme with imperfect hardware,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3779–3791, Jun. 2022.

[35] Y. Liu, I. Al-Nahhal, O. A. Dobre, and F. Wang, “Deep-learning channel estimation for IRS-assisted integrated sensing and communication system,” *IEEE Trans. Veh. Technol.*, vol. 72, no. 5, pp. 6181–6193, May 2023.

[36] Y. Liu, I. Al-Nahhal, O. A. Dobre, and F. Wang, “Deep-learning-based channel estimation for IRS-assisted ISAC system,” in *Proc. IEEE Glob. Commun. Conf.*, Dec. 2022, pp. 4220–4225.

[37] Y. Liu, I. Al-Nahhal, O. A. Dobre, F. Wang, and H. Shin, “Extreme learning machine-based channel estimation in IRS-assisted multi-user ISAC system,” *IEEE Trans. Commun.*, vol. 71, no. 12, pp. 6993–7007, Dec. 2023.

[38] A. Faisal, I. Al-Nahhal, K. Lee, O. A. Dobre, and H. Shin, “Conditional generative adversarial networks for channel estimation in RIS-assisted ISAC systems,” *IEEE Trans. Commun.*, vol. 73, no. 9, pp. 7828–7841, Sep. 2025.

[39] Å. Björck, “Least squares methods,” *Handbook Numer. Anal.*, vol. 1, pp. 465–652, 1990.

[40] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, “Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 898–912, Feb. 2022.

[41] R. Mai, D. H. N. Nguyen, and T. Le-Ngoc, “Joint MSE-based hybrid precoder and equalizer design for full-duplex massive MIMO systems,” in *Proc. IEEE Int. Conf. Commun.*, 2016, pp. 1–6.

[42] 3rd Generation Partnership Project (3GPP), “Evolved universal terrestrial radio access (e-utra); physical channels and modulation,” 3GPP, Sophia Antipolis, France, Tech. Specification 36.211, Jun. 2016.

[43] S. Biswas, C. Masouros, and T. Ratnarajah, “Performance analysis of large multiuser MIMO systems with space-constrained 2-D antenna arrays,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3492–3505, May 2016.

[44] G. -B. Huang, Q. -Y. Zhu, and C. -K. Siew, “Extreme learning machine: Theory and applications,” *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006 . [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231206000385

[45] S. Ding, X. Xu, and R. Nie, “Extreme learning machine and its applications,” *Neural Comput. Appl.*, vol. 25, no. 3/4, pp. 549–556, Sep. 2014.

[46] G. -B. Huang, D. H. Wang, and Y. Lan, “Extreme learning machines: A survey,” *Int. J. Mach. Learn. Cybern.*, vol. 2, pp. 107–122, May 2011.

[47] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, “Joint radar and communication design: Applications, state-of-the-art, and the road ahead,” *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, Jun. 2020.

[48] H. Zhang, L. Chen, K. Han, Y. Chen, and G. Wei, “Coexistence designs of radar and communication systems in a multi-path scenario,” *IEEE Trans. Veh. Technol.*, vol. 73, no. 3, pp. 3733–3749, Mar. 2024.

[49] T. Jiang, H. V. Cheng, and W. Yu, “Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1931–1945, Jul. 2021.

Alice Faisal (Graduate Student Member, IEEE) received the M.Sc. degree in electrical engineering in 2022 from Memorial University, where she is currently working toward the Ph.D. degree. Her research interests include 6G enabling technologies; reconfigurable intelligent surfaces, integrated sensing and communication, and full-duplex communications.

Ibrahim Al-Nahhal (Senior Member, IEEE) received the Ph.D. degree in electronics and communications engineering from Memorial University, Canada, in 2020. He is an Assistant Professor at Memorial University, Canada, since 2024. From 2014 to 2015, he was a physical Layer Expert at Nokia (formerly Alcatel-Lucent), Belgium. He was a Research Associate and Per-Course Instructor at Memorial University, Canada, since 2021. He holds three patents. He co-authored 50+ peer-reviewed journals and conference papers in top-ranked venues. He serves as Editor of IEEE WIRELESS COMMUNICATIONS LETTERS. His research interests are reconfigurable intelligent surfaces, full-duplex communications, integrated sensing and communication, channel estimation, and machine learning for communications.

Octavia A. Dobre (Fellow, IEEE) is a Professor and Tier-1 Canada Research Chair with Memorial University, Canada. Her research spans wireless communications and networking technologies, as well as optical and underwater communications. She has authored or co-authored over 600 publications in these areas and received ten Best Paper Awards, including the IEEE Communications Society Heinrich Hertz Award. Dr. Dobre serves as the VP Publications of the IEEE Communications Society. She was the founding Editor-in-Chief (EiC) of the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY and previously served as Editor-in-Chief of IEEE COMMUNICATIONS LETTERS. Dr. Dobre is a Clarivate Highly Cited Researcher, a Fellow of the Royal Society of Canada, the Canadian Academy of Engineering, and the Engineering Institute of Canada. She is also an elected member of the European Academy of Sciences and Arts and the Academia Europaean.

Hyundong Shin (Fellow, IEEE) received the B.S. degree in electronics engineering from Kyung Hee University (KHU), Yongin-si, South Korea, in 1999, and the M.S. and Ph.D. degrees in electrical engineering from Seoul National University, Seoul, South Korea, in 2001 and 2004, respectively. During his postdoctoral research with the Massachusetts Institute of Technology from 2004 to 2006, he was with the Laboratory for Information Decision Systems. In 2006, he joined KHU, where he is currently a Professor with the Department of Electronic Engineering. His research interests include quantum information science, wireless communication, and machine intelligence. Dr. Shin was the recipient of the IEEE Communications Society's Guglielmo Marconi Prize Paper Award and William R. Bennett Prize Paper Award. He was the Publicity Co-Chair for the IEEE PIMRC and the Technical Program Co-Chair for the IEEE WCNC and the IEEE GLOBECOM. He was the Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE COMMUNICATIONS LETTERS.