Conditional Generative Adversarial Networks for Channel Estimation in RIS-Assisted ISAC Systems

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

Abstract—Integrated sensing and communication (ISAC) technology has been explored as a potential advancement for future wireless networks, striving to effectively use spectral resources for both communication and sensing. The integration of reconfigurable intelligent surfaces (RIS) with ISAC further enhances this capability by optimizing the propagation environment, thereby improving both the sensing accuracy and communication quality. Within this domain, accurate channel estimation is crucial to ensure a reliable deployment. Traditional deep learning (DL) approaches, while effective, can impose performance limitations in modeling the complex dynamics of wireless channels. This paper proposes a novel application of conditional generative adversarial networks (CGANs) to solve the channel estimation problem of an RIS-assisted ISAC system. The CGAN framework adversarially trains two DL networks, enabling the generator network to not only learn the mapping relationship from observed data to real channel conditions but also to improve its output based on the discriminator network feedback, thus effectively optimizing the training process and estimation accuracy. The numerical simulations demonstrate that the proposed CGAN-based method improves the estimation performance effectively compared to conventional DL techniques. The results highlight the CGAN's potential to revolutionize channel estimation, paving the way for more accurate and reliable ISAC deployments.

Index Terms—Integrated sensing and communication (ISAC), reconfigurable intelligent surface (RIS), channel estimation, deep learning (DL), conditional generative adversarial networks (CGAN).

I. INTRODUCTION

THE recent releases by the third generation partnership project (3GPP) represent crucial advancements in the evolution of fifth-generation (5G) networks, aiming at realizing the full potential of 5G and bridging the transition to the sixth-generation (6G) networks. One of the key focus areas of 3GPP Release-19 is the integration of sensing capabilities into communication networks, which paves the way for new applications [1], [2]. To this end, integrated sensing and communication (ISAC) is envisioned to play a key role in future generations of wireless networks by efficiently merging radar sensing with communication technologies within a single system. This convergence is expected to enhance spectrum utilization and reduce hardware costs compared to deploying

separate systems with independent hardware for each function. ISAC is well-suited for a wide range of applications, such as drone operation, industrial automation, and health monitoring, where simultaneous data transmission and environmental sensing is needed.

ISAC has been investigated in the literature from various perspectives to realize its full potential [3]–[10]. In particular, novel ISAC transceivers were investigated through the development of advanced beamforming and signal processing techniques [3]–[5]. Privacy and security concerns were addressed through the design of robust security protocols and privacy-preserving methods [6]. Some works focused on resource allocation problems in ISAC-based systems to maintain effective sensing and communication (SAC) operations [7]–[10]. Solving optimization problems in ISAC systems often involve complex trade-offs between enhancing signal quality and optimizing power consumption across both SAC tasks. ISAC faces additional challenges, such as complex interference management issues that arise from simultaneously handling SAC tasks. Furthermore, the overlapping use of the spectrum for both functions can lead to signal contamination, necessitating advanced signal processing techniques to effectively distinguish between communication signals and sensing echoes. Moreover, the dual-use can also introduce power management difficulties, as both SAC operations can be power-intensive. Addressing these challenges is crucial for realizing the potential of ISAC networks.

Besides ISAC, the reconfigurable intelligent surfaces (RIS) have emerged as a potential technology to meet the demands of next-generation wireless networks. RIS consists of a two-dimensional array of low-cost passive elements that can be individually controlled to tune the phase, amplitude, and polarization of incoming signal. This enables RIS to effectively mitigate interference, enhance signal strength at the receiver-side, and extend the wireless communication coverage [11]. Due to the unique properties of RIS and ISAC, the integration of both technologies represents a significant advancement in the development of future wireless networks. Notably, RIS can be deployed to help combat some of the ISAC system challenges. In particular, RIS can mitigate interference between SAC signals by enabling precise control over the propagation environment. Additionally, the passive nature of RIS elements can help reduce the overall energy consumption of the ISAC systems. This capability enhances the wireless communication signal quality and minimizes the system power simultaneously.

Recently, the joint integration of ISAC and RIS has been investigated in various wireless communication scenarios

[12]–[16]. For example, the authors in [12] and [13] focused on optimizing the base station (BS) transmit precoding and RIS phase shifts to improve the radar estimation performance subject to the signal-to-interference-plus-noise ratio requirements of communication users. Furthermore, the work in [14] developed a beam training scheme that enables the BS to communicate with the users while also sensing targets. This approach distinguishes between the RIS and targets based on their mixed echoes, which is crucial to ensure seamless integration of RIS-assisted ISAC systems. Since conventional methods often face limitations, either suffering from high computational complexity when using iterative approaches or poor performance when using heuristic approaches, some studies considered deep learning (DL) for resource allocation tasks [17], [18].

It is important to note that accurate channel state information (CSI) is needed for all the above designs. The aforementioned works generally assumed that CSI is available at the receiver side and only focused on the resource allocation and interference mitigation problems in RIS-assisted ISAC systems. However, in practical scenarios, obtaining accurate CSI poses a significant challenge due to the coexistence of SAC channels within the same system. This introduces the need of innovative channel estimation strategies that can effectively distinguish between the mixed signals in RIS-assisted ISAC systems, ensuring both accurate sensing data and reliable communication. Despite its importance, only a limited number of studies addressed the channel estimation challenges in RIS-assisted ISAC systems [19]–[22]. In [19], the authors focused on a single-RIS-assisted ISAC system and introduced a novel three-stage method to simplify the channel estimation process. The presented approach starts with direct channels estimation, followed by the estimation of reflected communication channel, and finally the reflected sensing channel. This framework utilizes two distinct convolutional neural network (CNN) models to accurately estimate the channels at the ISAC BS. Furthermore, the work in [20] considered a multi-user downlink RIS-assisted ISAC system and proposed two deep neural network (DNN) models to estimate the SAC channels. Specifically, the first DNN is implemented at the ISAC BS for the sensing channel estimation, while the second is deployed at each downlink user for communication channel estimation. Finally, the work in [21] considered a multi-user RIS-assisted ISAC system and proposed a two-stage framework to estimate the SAC channels, focusing sequentially on the direct and reflected links. The presented framework considered extreme learning machine to meet the system requirements with accelerated training speed.

All the above works have benchmarked their performance against the least squares (LS) estimation technique. While commonly used, the LS estimation has several limitations including its sensitivity to noise, which can significantly degrade its performance in scenarios with low signal-to-noise ratio (SNR). Additionally, the LS estimator does not account for model non-linearities in the system, which can lead to biased and inefficient estimates in RIS-assisted ISAC environments. This gap points to the necessity for innovative approaches that enhance the estimation performance efficiently. Therefore, this work applies the conditional generative adversarial networks (CGANs) for channel estimation in RIS-assisted ISAC systems. Unlike regular DL methods that typically learn to map inputs to outputs in a supervised fashion, CGANs operate by training two models: a generator that creates data following the real data distribution, and a discriminator that learns to distinguish between the real and fake (i.e., generated) data. This adversarial process allows CGANs to generate accurate data and capture complex distributions, thereby enhancing the estimation and generalization capabilities over standard DL approaches. To the best of the authors' knowledge, this is the first work to consider employing CGANs for channel estimation tasks in multi-user RIS-assisted ISAC systems. The contributions of this paper are summarized as follows:

- A CGAN-based estimation framework, incorporating two distinct DNN architectures, is devised to estimate the SAC channels for a multi-user RIS-assisted ISAC system. One DNN operates at the ISAC BS to estimate the sensing channel, while the other one is deployed at each downlink user to estimate the cascaded communication channel.
- A custom loss function is carefully designed for training the proposed CGAN-based estimator to ensure accurate and robust channel estimation performance even in environments with low SNRs.
- The proposed CGAN framework operates as a minimax two-player game, where the generative and discriminative models continuously compete and evolve during the training process. This dynamic interaction leads to a generative model capable of producing channel samples that mimic real distribution patterns. This advanced approach enables the CGAN to exhibit exceptional adaptability and scalability, which is essential for RIS-assisted ISAC systems.
- The proposed CGAN-based channel estimation approach demonstrates strong generalization capabilities, where it achieves robust performance at SNR ranges that are not considered during the training phase. This robustness indicate that the model effectively eliminates the need for the SNR estimation stage in practical deployments.
- The computational complexity of the proposed CGAN approach is evaluated based on the number of real additions and multiplications. The numerical results indicate that the proposed algorithm achieves a complexity comparable to the DL-based benchmark estimator, which satisfies the low-cost deployment demands of RIS-assisted ISAC systems.
- Extensive simulations are conducted to validate the estimation performance of the proposed approach. Numerical findings prove that the proposed method significantly outperforms the conventional DL-based approaches in the literature under a range of SNR conditions and system parameters.

The remainder of this paper[1] is organized as follows: The system model and problem formulation are introduced in Section II. The proposed channel estimation approach is presented in Section III, and the computational complexity is analyzed
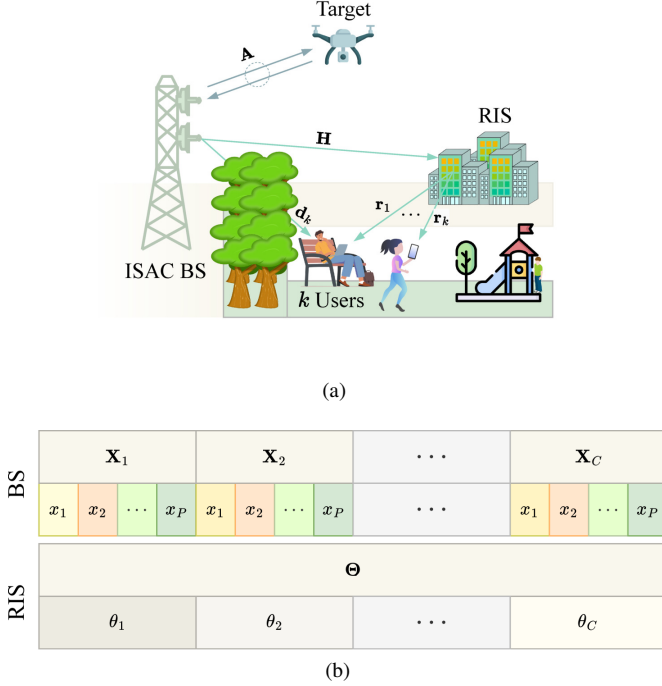
Fig. 1. RIS-assisted ISAC system. (a) System model, (b) Pilot protocol.

in Section IV. Finally, simulation results are shown in Section V and conclusions are drawn in Section VI.

II. SYSTEM MODEL

Consider an RIS-assisted multiple-input single-output (MISO) ISAC system as illustrated in Fig. 1a, where the RIS facilitates the communication between the ISAC BS and K downlink user equipments (UEs) and the BS communicates with a target for sensing purposes. The BS has M transmit antennas and one receive antenna, and each UE is equipped with one receive antenna. The RIS consists of N programmable reflecting elements. For target sensing, the BS sends radar signals towards the target and receives the echo signals through the BS-target-BS channel, denoted by $\mathbf{A} \in \mathbb{C}^{M \times M}$. The channel coefficients of the BS-RIS, RIS-UE$_k$, and BS-UE$_k$ are represented as $\mathbf{H} \in \mathbb{C}^{M \times N}$, $\mathbf{r}_k \in \mathbb{C}^{N \times 1}$, and $\mathbf{d}_k \in \mathbb{C}^{M \times 1}$, respectively. Note that all the SAC channels are assumed to be flat-fading considering the narrowband transmission [23]–[26]. Given that the ISAC BS operates in full-duplex mode, transmitting and receiving signals simultaneously, it experiences self-interference (SI); the SI channel is represented as $\mathbf{S} \in \mathbb{C}^{M \times M}$.

A pilot transmission policy is developed to estimate the SAC channels in the multi-user RIS-assisted ISAC system, as shown

---

[1]Boldface uppercase and lowercase letters denote matrices and vectors, respectively. $\mathbb{C}$ and $\mathcal{CN}$ stand for a complex-valued variable and a complex-valued normally distributed random variable, respectively. $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ denote the real and imaginary components of a variable, respectively. $\mathbb{E}\{\cdot\}$ denotes the expectation operation. The operators $(\cdot)^{\text{H}}$, $\text{vec}[\cdot]$, $(\cdot)^{-1}$, $\|\cdot\|_2$, and $\|\cdot\|_F$ represent the Hermitian, vectorization, inverse, second order norm, and Frobenius norm of their arguments, respectively. $\text{diag}\{\mathbf{x}\}$ returns a matrix whose diagonal consists of the elements of $\mathbf{x}$.

in Fig. 1b. The ISAC BS transmits pilot sequences in C sub-frames, in which each sub-frame is divided into P time slots. The pilot signal of the ISAC BS at the p-th time slot is defined as $\mathbf{x}_p \in \mathbb{C}^{M \times 1}$. To this end, the pilot matrix in sub-frame c is represented as $\mathbf{X}_c = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_P] \in \mathbb{C}^{M \times P}$. It is worth noting that the ISAC BS transmits the same pilot sequences (i.e., denoted as $\mathbf{X}$) in each sub-frame. Consequently, the RIS phase shifts remain unchanged within a single sub-frame, being denoted by $\boldsymbol{\theta}_c \in \mathbb{C}^{N \times 1}$. The corresponding phase shift matrix is represented by $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_C] \in \mathbb{C}^{N \times C}$. Let $P = M$ and $C = N$ to accommodate the necessity for low pilot overhead. This design allows to effectively capture the necessary channel information without requiring additional pilot resources, which significantly reduces redundancy and prevents excessive pilot usage. Furthermore, both $\mathbf{X}$ and $\boldsymbol{\Theta}$ are modeled as a discrete Fourier transform (DFT) matrix, expressed as

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & X^1 & \cdots & X^{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X^{M-1} & \cdots & X^{(M-1)(P-1)} \end{bmatrix}, \quad (1)$$

where $X^{(m,p)} = \frac{1}{\sqrt{M}} e^{j\frac{2\pi}{M} mp}$ is the (m,p)-th entry of $\mathbf{X}$. Modeling $\mathbf{X}$ as a DFT matrix helps combat the interference between the SAC signals and distinguish between multi-user signals. Furthermore, it has been demonstrated that designing $\boldsymbol{\Theta}$ as a DFT matrix helps to boost the power of the received signal at UEs and ensure accurate channel estimation [27].

To this end, the received signal at the k-th downlink UE for sub-frame c and time slot p is expressed as

$$y_{k,c,p}^{\text{UE}} = \left(\mathbf{r}_k^H \text{diag}\{\boldsymbol{\theta}_c^H\}\mathbf{H}^H + \mathbf{d}_k^H\right) \mathbf{x}_p + n_{k,c,p}. \quad (2)$$

Here, $n_{k,c,p} \sim \mathcal{CN}(0, \sigma^2)$ is the complex additive white Gaussian noise (AWGN) with zero-mean and variance σ^2. The RIS phase shifts are expressed as [28]

$$\boldsymbol{\theta}_c = \left[\beta_c e^{j\varphi_{c,1}}, \beta_c e^{j\varphi_{c,2}}, \cdots, \beta_c e^{j\varphi_{c,N}}\right]^T, \\ \varphi_{c,n} \in [0, 2\pi), \quad \beta_c \in [0,1], \quad (3)$$

where $\beta_c \in [0,1]$ represents the RIS element amplitude. The direct link is assumed to be blocked by obstacles and the reflected link is considered to support the communication. Here, $\text{diag}\{.\}$ transforms a vector into a diagonal matrix. This transformation implies the following

$$\text{diag}\{\boldsymbol{\theta}_c\} \mathbf{r}_k = [\theta_1 r_1, \theta_2 r_2, \cdots, \theta_N r_N]. \quad (4)$$

Similarly, constructing $\text{diag}\{\mathbf{r}_k\}$ and multiplying it by $\boldsymbol{\theta}_c$ would produce the same result, confirming the equality $\text{diag}\{\boldsymbol{\theta}_c\}\mathbf{r}_k = \text{diag}\{\mathbf{r}_k\}\boldsymbol{\theta}_c$ [29]. Using this property, the cascaded reflected channel of the BS-RIS-UE$_k$ is expressed as

$$\mathbf{G}_k = \mathbf{H}\,\text{diag}\{\mathbf{r}_k\} \in \mathbb{C}^{M \times N}. \quad (5)$$

Therefore, (2) can be formulated as

$$y_{k,c,p}^{\text{UE}} = \boldsymbol{\theta}_c^H \mathbf{G}_k^H \mathbf{x}_p + n_{k,c,p}. \tag{6}$$

On the other hand, the received signal at the ISAC BS is expressed as

$$\mathbf{y}_{c,p}^{\text{BS}} = \underbrace{\mathbf{A}^H \mathbf{x}_p}_{\text{Sensing signal}} + \underbrace{\mathbf{S}^H \mathbf{x}_p}_{\text{SI}} + \mathbf{n}_{c,p}, \tag{7}$$

where $\mathbf{n}_{c,p} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_M)$ denotes the complex AWGN with zero-mean and variance σ^2. Here, $\mathbf{I}_M$ represents an identity matrix with size M. Since the propagation environment between the ISAC BS transmitting and receiving antennas is presumed to be stable [30], the SI channel, $\mathbf{S}$, can be pre-determined at the ISAC BS. This allows for the compensation of any residual SI in (7) prior to the SAC estimation process. The estimation problems of the SAC channels in (6) and (7) are challenging. In what follows, we propose a novel approach to estimate both $\mathbf{G}_k$ and $\mathbf{A}$ based on the pilot protocol in Fig. 1b.

III. PROPOSED ESTIMATION APPROACH

In traditional DL approaches for channel estimation, performance can saturate or generalize poorly when dealing with noisy or complex data environments, due to their dependence on direct mappings from input to output data. To address these limitations, the proposed method leverages GANs, which integrate a network feedback to refine the generation process. This adversarial setup not only enhances the ability to handle diverse and noisy datasets but also improves the network capability to generate realistic channel estimates that closely mimic actual channel conditions, ensuring robust performance across various scenarios. This section provides an overview of GANs and then details the proposed framework for the channel estimation problem in RIS-assisted ISAC systems.

A. Overview of Generative Adversarial Networks

GANs are a novel class of DL frameworks introduced in [31], consisting of two DNNs: the generator and discriminator. In this framework, the generator strives to produce data that cannot be distinguished from the original dataset. The discriminator, on the other hand, tries to classify the given data as real or fake (i.e., generated samples). Both networks engage in a continuous adversarial training process, where the discriminator learns to identify generated data more accurately and the generator learns to better generate data samples given the mutual feedback. This process continues until the generator learns to produce real-like samples that the discriminator fails to identify as fake.

GANs and their variants have been utilized across a range of applications beyond the realm of wireless communications. It has been considered to generate synthetic human faces and transforming real-world scenery images into styles of famous paintings [32]. Additional examples include enhancing image resolution, and generating music, audio, and video [33]–[35]. In the field of wireless communications, GANs have emerged as powerful tool for various applications [36]. In particular,

GANs has been used to generate synthetic data that augment existing datasets in situations where collecting real wireless communication data is challenging. This is particularly useful for tasks such as signal classification, mitigating wireless jamming attacks, and spectrum sensing [37]–[39]. GANs have been further considered for optimizing wireless systems, where they learn to generate accurate resource allocation decisions by modeling the network environment and the continuous feedback from the discriminator [40]. Overall, GANs contribute significantly to advancements in network design, optimization, and data generation in wireless communications, driving forward innovations that enhance the efficiency and reliability of wireless systems.

In traditional GANs, the generator, G, takes as an input random noise vector, $\mathbf{z}$, typically drawn from a Gaussian or uniform distribution, $d(\mathbf{z})$. It outputs data that resemble the target data on which the model was trained, denoted by $G(\mathbf{z})$. The discriminator takes the target (i.e., real data) as an input, $\mathbf{x}$, with distribution $d(\mathbf{x})$, and the generated data, $G(\mathbf{z})$. The output of the discriminator is a binary classification, $D(\mathbf{x})$ and $D(G(\mathbf{z})) \in \{0, 1\}$, where the goal is to determine whether the given data are real or fake. The objective function of GANs is represented as

$$\min_G \max_D \mathbb{E}_{\mathbf{x} \sim d(\mathbf{x})} \left[\log(D(\mathbf{x}))\right] + \mathbb{E}_{\mathbf{z} \sim d(\mathbf{z})} \left[\log(1 - D(G(\mathbf{z})))\right]. \tag{8}$$

The first term in (8) represents the expected value of the discriminator output (i.e., probability that $\mathbf{x}$ is real) over all real data samples, $\mathbf{x}$. The discriminator focuses on maximize this term by assigning higher probabilities to real data, thus maximizing $\log(D(\mathbf{x}))$. Furthermore, the discriminator aims to maximize the value of the second term by minimizing $D(G(\mathbf{z}))$ (i.e., probability that input is generated). On the other hand, the generator aims to maximize $D(G(\mathbf{z}))$ by convincing the discriminator that the generated input is real.

Both the generator and discriminator are updated through an iterative process, each using a distinct loss function derived from the overall objective function in (8). This training process involves alternating updates to the discriminator and the generator with the goal of optimizing their respective loss functions. The loss functions of the discriminator and generator are respectively given as

$$L_D = -\frac{1}{2} \sum_{j=1}^{b} \left[\log D(\mathbf{x}_j) + \log(1 - D(G(\mathbf{z}_j)))\right], \tag{9}$$

and

$$L_G = -\frac{1}{b} \sum_{j=1}^{b} \log D(G(\mathbf{z}_j)), \tag{10}$$

where b is the batch size.

Given the above concept, GANs have made significant progress in the field of generative modeling, offering the ability to produce highly realistic samples across various domains. However, despite their successes, traditional GANs face a primary challenge that can affect their applicability to wireless systems. The main issue is the lack of control over the

generated data. In particular, the generator learns to create data that can fool the discriminator but does not necessarily retain the essential features or characteristics of the target. This can result in a lack of diversity in the outputs, making it difficult to capture the complexity of channel variations and the specific structure of wireless communication signals. Furthermore, traditional GANs have limited interpretability, where changes in the input noise vector do not directly correspond to meaningful variations in the generated data, which is a critical requirement in complex environments, such as wireless communication systems.

To address these limitations, CGANs have been specifically developed to incorporate additional information that guides the data generation process [41]. In CGANs, the generator and discriminator models are conditioned on auxiliary information, such as class labels or specific data attributes. This conditioning allows the generator to produce samples that align with the provided conditions, ensuring that the generated outputs maintain the necessary characteristics of the target data distribution. As a result, CGANs offer improved diversity, stability, and faster convergence in training. In the context of channel estimation, CGANs enable the generator to be conditioned on different channel states or environmental parameters, which is essential for accurately modeling the characteristics of the channel. This conditioning mechanism allows CGANs to adapt to variations in channel properties, leading to a more robust channel estimation process that can handle rapidly changing communication environments. Unlike basic GANs, which rely solely on input noise, the proposed CGAN framework in this work uses received observations as conditional inputs, allowing it to learn and reproduce complex channel structures with higher accuracy. In what follows, we detail the proposed CGAN channel estimation approach for RIS-assisted ISAC systems.

B. Proposed CGAN-Based DL Framework

This section introduces a novel channel estimation approach utilizing CGANs. Initially, the configuration of the input-output pairs for the framework is carefully designed. Building on this, a CGAN-based channel estimation approach is developed to enhance the estimation performance.

1) Input-Output Pair Design:

a) SAC Design: Consider the received signal in (6) to construct the input-output pairs for estimating the communication channel. In each sub-frame, the received signal at UE_k across P time slots is represented by

$$\mathbf{y}_{k,c}^{\text{UE}} = \boldsymbol{\theta}_c^H \mathbf{G}_k^H \mathbf{X} + \mathbf{n}_{k,c}, \qquad (11)$$

where $\mathbf{y}_{k,c}^{\text{UE}} = \left[y_{k,c,1}^{\text{UE}}, y_{k,c,2}^{\text{UE}}, \cdots, y_{k,c,P}^{\text{UE}} \right] \in \mathbb{C}^{1 \times P}$ and $\mathbf{n}_{k,c} = [n_{k,c,1}, n_{k,c,2}, \cdots, n_{k,c,P}] \in \mathbb{C}^{1 \times P}$. Based on this, the communication data design is given as

$$\mathbf{R}^{\text{UE}_k} = \left[\text{Re}\{\mathbf{y}_{k,1}^{\text{UE}}, \mathbf{y}_{k,2}^{\text{UE}}, \cdots, \mathbf{y}_{k,C}^{\text{UE}}\}, \\ \text{Im}\{\mathbf{y}_{k,1}^{\text{UE}}, \mathbf{y}_{k,2}^{\text{UE}}, \cdots, \mathbf{y}_{k,C}^{\text{UE}}\} \right]^T. \qquad (12)$$

The corresponding output (i.e., target) is the ground truth channel, $\mathbf{G}_k$, given as

$$\mathbf{O}^{\text{UE}_k} = [\text{Re}\{\text{vec}[\mathbf{G}_k]\}, \text{Im}\{\text{vec}[\mathbf{G}_k]\}]^T. \qquad (13)$$

On the other hand, to construct the input-output pairs for estimating the sensing channel, consider the received signal in (7). Similarly, the received signals at the ISAC BS are stacked across each sub-frame, C, leading to

$$\mathbf{Y}_c^{\text{BS}} = \mathbf{A}^H \mathbf{X} + \mathbf{N}_c, \qquad (14)$$

where $\mathbf{N}_c = [\mathbf{n}_{c,1}, \mathbf{n}_{c,2}, \cdots, \mathbf{n}_{c,P}] \in \mathbb{C}^{M \times P}$ and $\mathbf{Y}_c^{\text{BS}} = \left[\mathbf{y}_{c,1}^{\text{BS}}, \mathbf{y}_{c,2}^{\text{BS}}, \cdots, \mathbf{y}_{c,P}^{\text{BS}} \right] \in \mathbb{C}^{M \times P}$. To this end, the sensing data is generated as

$$\mathbf{R}^{\text{BS}} = \left[\text{Re}\{\text{vec}\left[\mathbf{Y}_1^{\text{BS}}, \mathbf{Y}_2^{\text{BS}}, \cdots, \mathbf{Y}_C^{\text{BS}}\right]\}, \\ \text{Im}\{\text{vec}\left[\mathbf{Y}_1^{\text{BS}}, \mathbf{Y}_2^{\text{BS}}, \cdots, \mathbf{Y}_C^{\text{BS}}\right]\} \right]^T. \qquad (15)$$

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

$$\mathbf{O}^{\text{BS}} = [\text{Re}\{\text{vec}[\mathbf{A}]\}, \text{Im}\{\text{vec}[\mathbf{A}]\}]^T. \qquad (16)$$

It is worth noting that ground truth channels are used solely for the training phase. During training, the availability of the true channel values enables the CGAN network to learn the underlying channel generation process effectively. However, in the testing phase, the only available information is the observation, represented by the received signal at the BS or UE. In this phase, the trained CGAN relies on this received signal to generate an estimate of the channels. This ensures that the CGAN model generalizes well to real-world scenarios, where the true channels are unknown. Note that the generated dataset will not be used for direct input-output relationship as the case with regular NNs. The following sub-section will detail the working principle of the proposed CGAN channel estimation framework.

b) Dataset Generation: The training samples are produced by utilizing Q received signals (i.e., each of the communication and sensing signals in (11) and (14), respectively), along with $V - 1$ duplicates of the q-th signal. The duplicates are created by adding synthetic AWGN to the original channel according to $\text{SNR}_{\mathbf{A}/\mathbf{G}_k} = \frac{p_{\mathbf{A}/\mathbf{G}_k}}{\sigma_{\mathbf{A}/\mathbf{G}_k}^2}$, where $p_{\mathbf{A}/\mathbf{G}_k}$ and $\sigma_{\mathbf{A}/\mathbf{G}_k}^2$ denote the channel power and synthetic noise, respectively [20]. This dataset includes both noise-free original signals and their noisy versions to enhance the robustness and performance of the CGAN estimation model. Accordingly, the dataset for communication and sensing can be respectively represented as

$$(\mathcal{R}^{\text{UE}_k}, \mathcal{O}^{\text{UE}_k}) = \left\{ \left(\mathbf{R}_{(1,1)}^{\text{UE}_k}, \mathbf{O}_{(1)}^{\text{UE}_k} \right), \left(\mathbf{R}_{(1,2)}^{\text{UE}_k}, \mathbf{O}_{(1)}^{\text{UE}_k} \right), \cdots, \\ \left(\mathbf{R}_{(1,V)}^{\text{UE}_k}, \mathbf{O}_{(1)}^{\text{UE}_k} \right), \left(\mathbf{R}_{(2,1)}^{\text{UE}_k}, \mathbf{O}_{(2)}^{\text{UE}_k} \right), \cdots, \left(\mathbf{R}_{(Q,V)}^{\text{UE}_k}, \mathbf{O}_{(Q)}^{\text{UE}_k} \right) \right\}, \qquad (17)$$

and

$$(\mathcal{R}^{\text{BS}}, \mathcal{O}^{\text{BS}}) = \left\{ \left(\mathbf{R}_{(1,1)}^{\text{BS}}, \mathbf{O}_{(1)}^{\text{BS}} \right), \left(\mathbf{R}_{(1,2)}^{\text{BS}}, \mathbf{O}_{(1)}^{\text{BS}} \right), \cdots, \\ \left(\mathbf{R}_{(1,V)}^{\text{BS}}, \mathbf{O}_{(1)}^{\text{BS}} \right), \left(\mathbf{R}_{(2,1)}^{\text{BS}}, \mathbf{O}_{(2)}^{\text{BS}} \right), \cdots, \left(\mathbf{R}_{(Q,V)}^{\text{BS}}, \mathbf{O}_{(Q)}^{\text{BS}} \right) \right\}. \qquad (18)$$
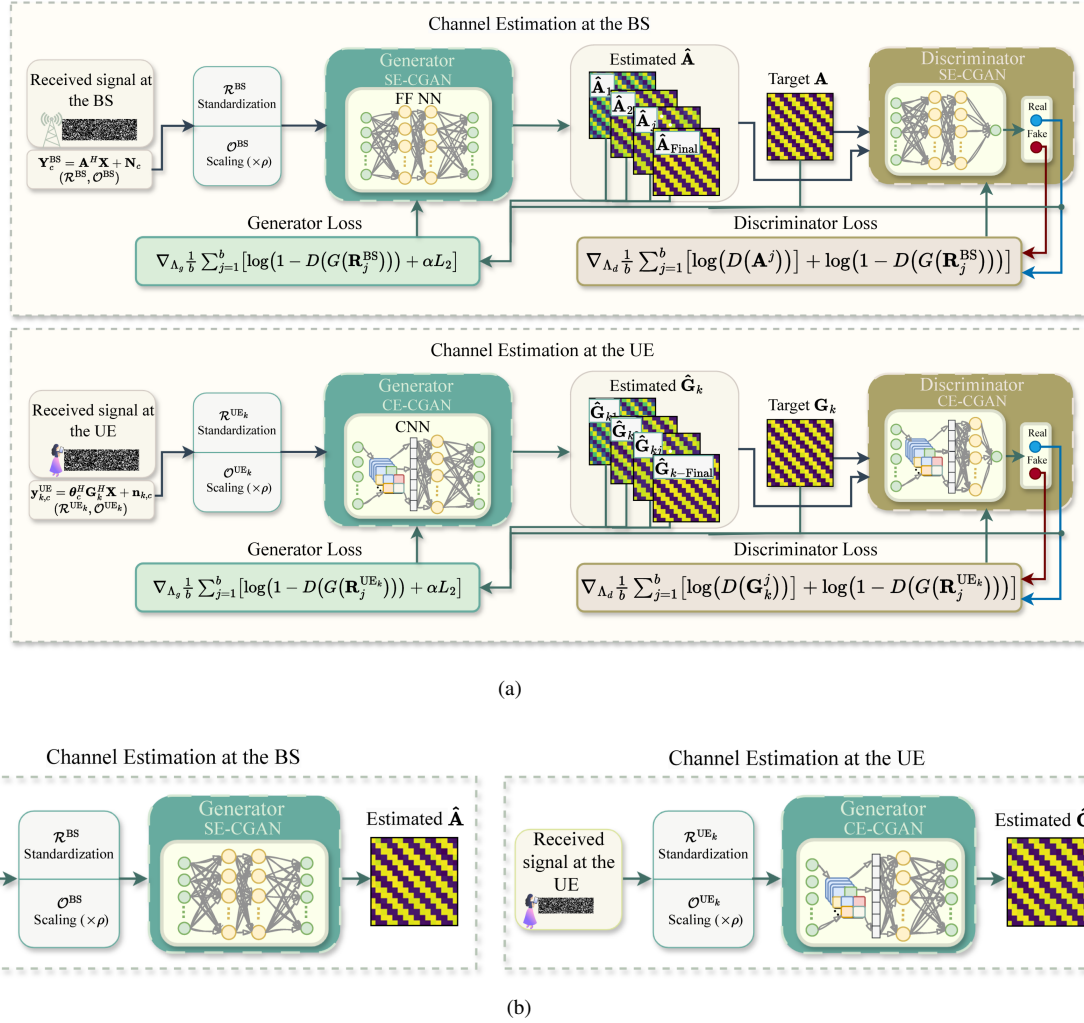
Fig. 2. Proposed CGAN-based channel estimation framework for RIS-assisted ISAC system. (a) Offline training. (b) Online testing.

2) Working Principle: The proposed CGAN-based estimation framework to estimate $\mathbf{A}$ and $\mathbf{G}_k$ is shown in Fig. 2 and detailed as follows

a) Offline Training: The generator NN aims to estimate the SAC channels based on the conditional input, which includes the received signal observations, $\mathbf{R}^{\text{BS}}$ or $\mathbf{R}^{\text{UE}_k}$. Prior to being passed into the generator, these received signals undergo preprocessing steps to ensure that the input data has a consistent range and to improve the model stability. A standardization is performed on the (q, v)-th input data, which can be mathematically expressed as

$$\mathbf{R}_{\text{std}}^{\text{BS/UE}_k}(q,v) = \frac{\mathbf{R}^{\text{BS/UE}_k}(q,v) - \mathbb{E}\{\mathbf{R}^{\text{BS/UE}_k}(q,v)\}}{\mathbb{D}\{\mathbf{R}^{\text{BS/UE}_k}(q,v)\}}, \quad (19)$$

where $\mathbb{E}\{\mathbf{R}^{\text{BS/UE}_k}(q,v)\}$ and $\mathbb{D}\{\mathbf{R}^{\text{BS/UE}_k}(q,v)\}$ denote the mean and standard deviation of the input, respectively. Additionally, the estimation target is scaled to facilitate the learning process as follows

$$\mathbf{O}_s^{\text{BS/UE}_k}(q,v) = \rho \mathbf{O}^{\text{BS/UE}_k}(q,v), \quad (20)$$

where ρ is the scaling factor.

The discriminator NN, on the other hand, aims to identify whether a given input is real (i.e., labeled 1) or fake (i.e., labeled 0), as derived from the generator output; this process is summarized as follows

$$D(\mathbf{B}) = \begin{cases} 1, & \text{if } \mathbf{B} \in \{\mathbf{A}, \mathbf{G}_k\} \\ 0, & \text{if } \mathbf{B} \in \{\hat{\mathbf{A}}, \hat{\mathbf{G}}_k\} \end{cases} \quad (21)$$

where $\mathbf{B}$ represents the input to the discriminator, either as ground truth channel matrices, $\mathbf{A}$ or $\mathbf{G}_k$, or generated estimates $\hat{\mathbf{A}}$ or $\hat{\mathbf{G}}_k$.

In this work, the overall objective function of GANs presented in (8) is further improved by adding an L_2 term. This modification guides the generator towards more accurate estimates, ensuring that the generated channel not only deceives the discriminator but also closely approximates the true channel values. Furthermore, the L_2 loss term serves as a regularizer, penalizing large deviations from the target values. This helps prevent overfitting and fosters a model that generalizes better to new, unseen data while maintaining robustness against noise in received signals. To this end, (8) becomes

$$\min_G \max_D \; \mathbb{E}\big[\log\big(D\big(\mathbf{A}, \mathbf{G}_k\big)\big)\big] +$$
$$\mathbb{E}\big[\log\big(1 - D\big(G(\mathbf{R}^{\mathrm{BS}}), G(\mathbf{R}^{\mathrm{UE}_k})\big)\big)\big] + \alpha L_2, \tag{22}$$

where

$$L_2 = \mathbb{E}\big[\big\|\mathbf{B} - G(\mathbf{R})\big\|^2\big]. \tag{23}$$

Here, $\mathbf{B}$ denotes the SAC channels (i.e., $\mathbf{B} = \{\mathbf{A}, \mathbf{G}_k\}$), and $\mathbf{R}$ is the input data (i.e., $\mathbf{R} = \{\mathbf{R}^{\mathrm{BS}}, \mathbf{R}^{\mathrm{UE}_k}\}$). Furthermore, α a weighting factor that enhances the stability of the channel estimates by penalizing large deviations from the target outputs. The generator and discriminator are trained together in an adversarial manner, where each network contributes to the loss of the other. Specifically, the generator aims to create estimates that can fool the discriminator, while the discriminator aims to correctly distinguish between real and generated data. This adversarial training is repeated until convergence, where the generator learns to produce realistic estimates. Fig. 2a illustrates the proposed CGAN DNN architectures for the SAC channels estimation and summarizes the offline training process.

In particular, two different architectures are developed within the CGAN framework: one for sensing channel estimation (SE-CGAN) and the other one for communication channel estimation (CE-CGAN). The SE-CGAN network is designed as feedforward (FF) fully connected NN, while the CE-CGAN as CNN to handle the complicated input-output relationship associated with the cascaded communication channel. In particular, the proposed SE-CGAN starts with an input layer, an FF layer (FFL) and a batch normalization layer to stabilize the NN output. The output is then fed to the second FFL followed by the second batch normalization and an output layer. On the other hand, the CE-DNN consist of one convolutional layer (CL), batch normalization layer, followed by a flatten layer, an FFL, and a second batch normalization layer layer at the end. The two hidden layers for each of the SE-CGAN and CE-CGAN networks use the *LeakyReLU* activation function. This is defined as follows

$$\mathrm{LeakyReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ \gamma x & \text{otherwise.} \end{cases} \tag{24}$$

Here, γ is a small constant value, which is set to 0.2. Both CGANs utilize a discriminator architecture that is identical to their respective generators. The only difference consists in the output layer size; the generator outputs the estimated channels, whereas the discriminator outputs a single value (i.e., 0 or 1) to evaluate the input data. The *Adam* optimizer is adopted to update the generator and discriminator parameters with a learning rate of 2×10^{-4} and 2×10^{-5}, respectively [42]. These learning rates were carefully chosen based on extensive hyperparameter tuning. During this process, various learning rates were tested to find an optimal balance between convergence speed and estimation accuracy. Lower learning rates helped achieve more stable performance. Selecting a slightly higher rate for the generator enabled faster adaptation

to the changes in the discriminator feedback. The minibatch transitions is set to size 16 and $\alpha = 100$. The detailed architectures of the SE-CGAN and CE-CGAN are summarized in Table I.

b) Online Testing: The online testing phases for SAC channel estimation are shown in Fig. 2b. During the testing phase, only the generator network is used, with no access to the discriminator network. The generator, having been trained to generate accurate channel estimates, is now applied to the testing data to predict the channel matrices. The testing dataset, $\dot{\mathcal{R}}^{\mathrm{BS}}$ and $\dot{\mathcal{R}}^{\mathrm{UE}_k}$, first undergoes preprocessing to standardize the samples according to (19), denoted by $\bar{\mathbf{R}}^{\mathrm{BS}}$ and $\bar{\mathbf{R}}^{\mathrm{UE}_k}$, and then is fed to the generator network. The estimated output is represented as

$$\hat{\mathbf{A}} = \rho^{-1} G(\bar{\mathbf{R}}^{\mathrm{BS}}; \hat{\Lambda})$$
$$\hat{\mathbf{G}}_k = \rho^{-1} G(\bar{\mathbf{R}}^{\mathrm{UE}_k}; \hat{\Lambda}), \tag{25}$$

where $\hat{\Lambda}$ represents the hyperparameters of the trained CGAN network (i.e., SE-CGAN or CE-CGAN). The output of the generator is scaled by a factor of ρ^{-1} to obtain the channel matrices, $\hat{\mathbf{A}}$ and $\hat{\mathbf{G}}_k$. The proposed CGAN-based channel estimation algorithm is detailed in Algorithm 1, where Λ_g and Λ_d denote the generator and discriminator parameters, respectively.

IV. COMPLEXITY ANALYSIS

Based on the proposed SE-CGAN and CE-CGAN channel estimation frameworks, this section derives the computational complexity required by the trained CGAN-based framework to estimate the SAC channel. The computational complexity is computed in terms of the number of real additions, $C_{\mathcal{A}}$, and multiplications, $C_{\mathcal{M}}$. To provide a comprehensive analysis, we first consider the operations involved in the generator and discriminator networks of the proposed SE-CGAN. Both networks consist of several FFL, each contributing to the overall computational load. In particular, the proposed SE-CGAN contains two hidden FFL and an input and output layers. Let η_i be the number of neurons of the i-th layer. Therefore, according to the parameters presented in Table I, the computational complexity introduced by the first FFL is $\eta_2^{\mathrm{SE}}(\eta_1^{\mathrm{SE}} + 1)$ real additions and $\eta_1^{\mathrm{SE}}\eta_2^{\mathrm{SE}}$ real multiplications. Similarly, the rest of the FFLs have the complexity corresponding to the number of neurons. To this end, the required number of additions and multiplications required by the generator and discriminator networks of SE-CGAN are respectively given as

$$C_{\mathcal{A}_G}^{\mathrm{SE}} = \sum_{i=1}^{3} \eta_i^{\mathrm{SE}} \eta_{i+1}^{\mathrm{SE}} + \sum_{i=1}^{3} \eta_{i+1}^{\mathrm{SE}}, \tag{28}$$

$$C_{\mathcal{M}_G}^{\mathrm{SE}} = \sum_{i=1}^{3} \eta_i^{\mathrm{SE}} \eta_{i+1}^{\mathrm{SE}}, \tag{29}$$

and

Algorithm 1: Training a CGAN

---

1 Initialize the generator G and discriminator D with random weights, batch size b, learning rates, and training epochs ϵ;

2 Generate $(\mathcal{R}^{\mathrm{UE}_k}, \mathcal{O}^{\mathrm{UE}_k})$ and $(\mathcal{R}^{\mathrm{BS}}, \mathcal{O}^{\mathrm{BS}})$ according to Sec. III-B1b;

3 Pre-process $(\mathcal{R}^{\mathrm{UE}_k}, \mathcal{O}^{\mathrm{UE}_k})$ and $(\mathcal{R}^{\mathrm{BS}}, \mathcal{O}^{\mathrm{BS}})$;

4 for *epoch $i = 1$ to ϵ* **do**

5   **for** *batch $j = 1$ to B* **do**

6   **Sample** batch of samples from $\mathbf{R}^{\mathrm{UE}_k}$ or $\mathbf{R}^{\mathrm{BS}}$;

7   **Sample** batch of samples from $\mathbf{O}^{\mathrm{UE}_k}$ or $\mathbf{O}^{\mathrm{BS}}$;

8   **Generate** samples from the generator $G(\mathbf{R}^{\mathrm{UE}_k})$ or $G(\mathbf{R}^{\mathrm{UE}_k})$;

9   **Evaluate** generated samples $D(G(\mathbf{R}^{\mathrm{UE}_k}))$ or $D(G(\mathbf{R}^{\mathrm{UE}_k}))$;

10   **Evaluate** target samples $D(\mathbf{O}^{\mathrm{UE}_k})$ or $D(\mathbf{O}^{\mathrm{BS}})$;

11   **Update** D by ascending its stochastic gradient:

$$\nabla_{\Lambda_d} \frac{1}{b} \sum_{j=1}^{b} \Big[\log\big(D(\mathbf{A}^j, \mathbf{G}_k^j)\big)\Big] + \\ \log\big(1 - D\big(G(\mathbf{R}_j^{\mathrm{BS}}), G(\mathbf{R}_j^{\mathrm{UE}_k})\big)\big)\Big] \qquad (26)$$

12   **Update** G by descending its stochastic gradient:

$$\nabla_{\Lambda_g} \frac{1}{b} \sum_{j=1}^{b} \Big[\log\big(1 - D\big(G(\mathbf{R}_j^{\mathrm{BS}}), G(\mathbf{R}_j^{\mathrm{UE}_k})\big)\big) + \alpha L_2\Big] \qquad (27)$$

13   **end**

14 end

15 Output the trained SE-CGAN, $G(\bar{\mathbf{R}}^{\mathrm{BS}}; \hat{\Lambda}_g)$, at the ISAC BS and the trained CE-CGAN, $G(\bar{\mathbf{R}}^{\mathrm{UE}_k}; \hat{\Lambda}_g)$, at the downlink UE$_k$.

$$C_{\mathcal{A}_D}^{\mathrm{SE}} = \sum_{i=1}^{2} \eta_i^{\mathrm{SE}} \eta_{i+1}^{\mathrm{SE}} + \sum_{i=1}^{2} \eta_{i+1}^{\mathrm{SE}} + (\eta_3^{\mathrm{SE}} + 1), \qquad (30)$$

$$C_{\mathcal{M}_D}^{\mathrm{SE}} = \sum_{i=1}^{2} \eta_i^{\mathrm{SE}} \eta_{i+1}^{\mathrm{SE}} + \eta_3^{\mathrm{SE}}. \qquad (31)$$

To this end, the overall necessary number of real additions and multiplications of the proposed SE-CGAN is represented as

$$C_{\mathcal{A}}^{\mathrm{SE}} = 2\eta_1^{\mathrm{SE}} \eta_2^{\mathrm{SE}} + \sum_{i=2}^{3} k_i \eta_i^{\mathrm{SE}} + \eta_{4_G}^{\mathrm{SE}} + 1, \qquad (32)$$

$$C_{\mathcal{M}}^{\mathrm{SE}} = \sum_{i=1}^{2} 2\eta_i^{\mathrm{SE}} \eta_{i+1}^{\mathrm{SE}} + \eta_3^{\mathrm{SE}}(\eta_{4_G}^{\mathrm{SE}} + 1). \qquad (33)$$

where $k_2 = 2(\eta_3^{\mathrm{SE}} + 1)$ and $k_3 = 3 + \eta_{4_G}^{\mathrm{SE}}$.

On the other hand, the CE-CGAN network contains a CL, FFL, input and output layers. Similarly, the computational complexity contributions arise from both generator and discriminator networks. To this end, by summing up the contri-

butions from all layers in both the generator and discriminator networks, the total computational complexity of the proposed CE-CGAN framework can be obtained. Let F_z be the filter size, F_n be the number of filters, and F_s be the stride. The output size of a CL is given by

$$\eta_F = \left\lfloor \frac{\eta_1^{\mathrm{CE}} - F_z}{F_s} + 1 \right\rfloor, \qquad (34)$$

where $\lfloor \cdot \rfloor$ is the floor operation. The computational complexity is introduced by the the second, third, and fourth layers, since the flatten layer does not contain any addition or multiplication complexity. The number of required additions is presented as $(F_z + 1)\eta_F F_n$, $\eta_3^{\mathrm{CE}}(\eta_F F_n + 1)$ and $(\eta_4^{\mathrm{CE}} + 1)\eta_3^{\mathrm{CE}}$, respectively, while the number of required multiplications is presented as $F_z \eta_F F_n$, $\eta_3^{\mathrm{CE}} \eta_F F_n$, and $\eta_3^{\mathrm{CE}} \eta_4^{\mathrm{CE}}$, respectively. To this end, the required number of additions and multiplications required by the generator and discriminator networks of CE-CGAN are respectively given as

$$C_{\mathcal{A}_G}^{\mathrm{CE}} = (F_z + \eta_3^{\mathrm{CE}} + 1)\eta_F F_n + (\eta_{4_G}^{\mathrm{CE}} + 1)\eta_3^{\mathrm{CE}} + \eta_{4_G}^{\mathrm{CE}}, \quad (35)$$

$$C_{\mathcal{M}_G}^{\mathrm{CE}} = (F_z + \eta_3^{\mathrm{CE}})\eta_F F_n + \eta_3^{\mathrm{CE}} \eta_{4_G}^{\mathrm{CE}}, \qquad (36)$$

and

$$C_{\mathcal{A}_D}^{\mathrm{CE}} = (F_z + \eta_3^{\mathrm{CE}} + 1)\eta_F F_n + 2\eta_3^{\mathrm{CE}} + 1, \qquad (37)$$

$$C_{\mathcal{M}_D}^{\mathrm{CE}} = (F_z + \eta_3^{\mathrm{CE}})\eta_F F_n + \eta_3^{\mathrm{CE}}. \qquad (38)$$

Consequently, the total computational complexity of the CE-CGAN framework is expressed as

$$C_{\mathcal{A}}^{\mathrm{CE}} = 2(F_z + \eta_3^{\mathrm{CE}} + 1)\eta_F F_n + \eta_{4_G}^{\mathrm{CE}}(\eta_3^{\mathrm{CE}} + 1) + 3\eta_3^{\mathrm{CE}} + 1, \qquad (39)$$

$$C_{\mathcal{M}}^{\mathrm{CE}} = 2(F_z + \eta_3^{\mathrm{CE}})\eta_F F_n + \eta_3^{\mathrm{CE}}(\eta_{4_G}^{\mathrm{CE}} + 1). \qquad (40)$$

To this end, the total number of additions and multiplications required by the proposed CGAN framework is $C_{\mathcal{A}}^{\mathrm{SE}} + C_{\mathcal{A}}^{\mathrm{CE}}$ and $C_{\mathcal{M}}^{\mathrm{SE}} + C_{\mathcal{M}}^{\mathrm{CE}}$, respectively. In particular, the number of additions and multiplications is respectively given as in (41) and (42).

V. SIMULATION RESULTS

This section extensively validates the performance of the proposed CGAN channel estimation framework for the RIS-assisted ISAC system. First, the simulation parameters and setup are presented. Then, the SAC channel estimation performance is evaluated under different SNR values and system conditions.

A. Simulation Parameters

Let $K = 3$, $M = 4$, and $N = 30$ for all the following simulations, unless further specified. The sensing channel is modeled according to the radar channel model as [43], [44]

$$\mathbf{A} = \mu \mathbf{a}(\theta) \mathbf{a}(\theta)^H. \qquad (43)$$

Here, μ denotes the complex-valued reflection coefficient associated with the target with phase shifts uniformly distributed from $[0, 2\pi)$, and $\mathbf{a}(\theta)$ is the steering vector, expressed as

Model	Network	Layers	Size	Filter	Activation Function
SE-CGAN	Generator	Input	$2MP$	-	-
		FFL	100	-	LeakyReLU
		FFL	200	-	LeakyReLU
		Output	$2M^2$	-	-
	Discriminator	Input	$2M^2$	-	-
		FFL	100	-	LeakyReLU
		FFL	200	-	LeakyReLU
		Output	1	-	-
CE-CGAN	Generator	Input	$2P$	-	-
		CL	132	4	LeakyReLU
		FFL	500	-	LeakyReLU
		Output	$2MN$	-	-
	Discriminator	Input	$2MN$	-	-
		CL	132	4	LeakyReLU
		FFL	500	-	LeakyReLU
		Output	1	-	-

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

where $\theta = -\frac{2\pi}{3}$, d, and λ denote the azimuth angle, BS antenna spacing, and signal wavelength, respectively. On the other hand, the communication channels (i.e., $\mathbf{H}$ and $\mathbf{r}_k$) are modeled as Rician, being expressed as

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

where $\mathbf{h}$ represents the channels (i.e., $\mathbf{h} = \{\mathbf{H}, \mathbf{r}_k\}$), PL denotes the path loss, and K_1 is the Rician factor. $\bar{\mathbf{h}}$ and $\tilde{\mathbf{h}}$ are the line-of-sight and non-line-of-sight components of the channel, respectively. Here, $\bar{\mathbf{h}} = \mathbf{a}(\dot{\theta})\mathbf{a}(\bar{\theta})^H$, where $\mathbf{a}(\dot{\theta})$ corresponds to the angle of departure from the source to destination (i.e., BS to RIS/RIS to UE$_k$) and $\mathbf{a}(\bar{\theta})$ corresponds to the angle of arrival (i.e., RIS/UE$_k$) and are set to $\frac{\pi}{3}$. $\bar{\mathbf{h}}$ is formulated similar to (44), whereas $\tilde{\mathbf{h}}$ is the random component containing independent and identical distributed $\mathcal{CN}(0,1)$ elements. The Rician factor, K_1 is set to 10 and 0 (i.e., representing a Rayleigh fading model) for $\mathbf{H}$ and $\mathbf{r}_k$, respectively [45]. Furthermore, the PL is modeled as $\text{PL} = \text{PL}_r(\frac{d_j}{d_r})^{-\zeta_j}$. d_j is the distance, PL_r is the path loss at a reference distance d_r, and ζ_j is the path loss exponent. We set $PL_r = -30\,\text{dBm}$ and $d_r = 1\,\text{m}$. The path loss exponents of the BS-target-BS, BS-RIS, and RIS-UE$_k$ links are $\zeta_1 = 3$, $\zeta_2 = 2.3$, and $\zeta_3 = 2$, respectively, while the distances are set as $d_1 = 140\,\text{m}$, $d_2 = 50\,\text{m}$, and $d_3 = 2\,\text{m}$. According to [20], [46], the transmit power of the ISAC BS and is set to $P = -20\,\text{dBm}$.

The dataset size is $Q \times V = 10^4$ for each SNR value, with $Q = 1000$ and $V = 10$. In this work, 90% of the dataset size is used for training, while the remaining 10% is used for testing. The SNR values in the training stage are SNR $= 10: 5: 20\,\text{dB}$, whereas the testing stage uses SNR $= -10: 2.5: 30\,\text{dB}$, which includes 17 values from $-10\,\text{dB}$ to $30\,\text{dB}$ with a step increment of $2.5\,\text{dB}$. The choice of the SNR region ensures that the model encounters instances from unfamiliar conditions, not only unseen samples. This way, the simulation results confirm the generalization of the model and eliminate the need of estimating the SNR prior to estimating the desired channels. Lastly, the scaling factor, ρ, is set to 10^4.

To validate the performance of the proposed CGAN channel estimation approach, the normalized mean square error (NMSE) is considered as the main performance metric being expressed as

$$\text{NMSE} = \mathbb{E}\left\{ \frac{\left\|\mathbf{h}_{\text{Estimated}} - \mathbf{h}_{\text{True}}\right\|_F^2}{\left\|\mathbf{h}_{\text{True}}\right\|_F^2} \right\}. \quad (46)$$

Furthermore, the work in [20] is considered as a benchmark to effectively evaluate the performance of the proposed algorithm in the following subsections.

B. Impact of Varying the SNR

Fig. 3 shows the estimation performance under different SNR conditions. The NMSEs of the proposed CE-CGAN and SE-CGAN demonstrate a considerable improvement compared to the benchmark estimation approach presented in [20] as well as other traditional models, including a FF network

$$C_{\mathcal{A}}^{\text{Total}} = 2\eta_1^{\text{SE}}\eta_2^{\text{SE}} + \sum_{i=2}^{3} k_i\eta_i^{\text{SE}} + \eta_{4_G}^{\text{SE}} + 1 + 2(F_z + \eta_3^{\text{CE}} + 1)\eta_F F_n + \eta_{4_G}^{\text{CE}}(\eta_3^{\text{CE}} + 1) + 3\eta_3^{\text{CE}} + 1, \quad (41)$$

$$C_{\mathcal{M}}^{\text{Total}} = \sum_{i=1}^{2} 2\eta_i^{\text{SE}}\eta_{i+1}^{\text{SE}} + \eta_3^{\text{SE}}(\eta_{4_G}^{\text{SE}} + 1) + 2(F_z + \eta_3^{\text{CE}})\eta_F F_n + \eta_3^{\text{CE}}(\eta_{4_G}^{\text{CE}} + 1. \quad (42)$$
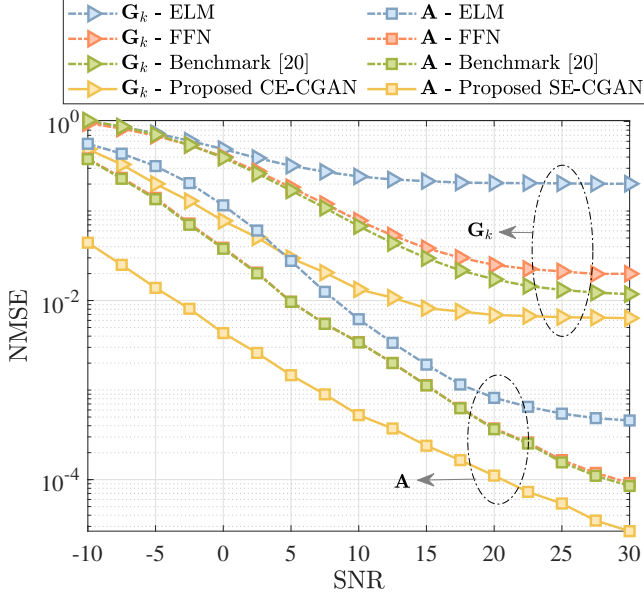
Fig. 3. NMSE performance for estimating SAC channels.

Fig. 4. NMSE performance for the communication channel estimation versus M at $N = 30$.

(FFN) and extreme learning machine (ELM). It is worth noting that the FFN model consists of two hidden layers, each containing 256 neurons, with hyperparameters aligned with the benchmark in [20]. The ELM model, on the other hand, consists of one hidden layer of 256 neurons to estimate both channels. These specifications ensure a fair comparison across all approaches. For the proposed SE-CGAN, this improvement is especially notable at lower SNR levels, where the proposed approach demonstrates good noise resilience, thus achieving lower NMSE values. This is one of the CGAN advantages, where the generator learns to create samples, not to map input data to an output. It generates data conditioned by the input sample but based on the loss function design, thereby reducing the noise impact. Furthermore, as the communication channel estimation is considered more complicated due to the cascaded relationship in (11), one can observe that the achieved NMSE is higher than that of the sensing link. However, overall, the proposed scheme outperforms the benchmark scheme and the additional traditional models substantially. In particular, the proposed approach achieves around $8\,\mathrm{dB}$ SNR improvement at NMSE $= 10^{-2}$ and NMSE $= 10^{-1}$ as compared to [20] for estimating $\mathbf{A}$ and $\mathbf{G}_k$, respectively. The figure also highlights the generalization performance of the proposed CGAN approach, where it outperforms the benchmark schemes even at SNR ranges that were excluded in the training process (i.e., SNR $= -10: 5\,\mathrm{dB}$).

C. Impact of Varying M

Fig. 4 analyzes the performance of the proposed approach in estimating the communication channel, $\mathbf{G}_k$, with respect to varying M under different SNR conditions: low (i.e., $-5\,\mathrm{dB}$) and high (i.e., $10\,\mathrm{dB}$) SNR values. As illustrated, the proposed CGAN consistently outperforms the benchmark scheme across all M values and SNR conditions, demonstrating supe-
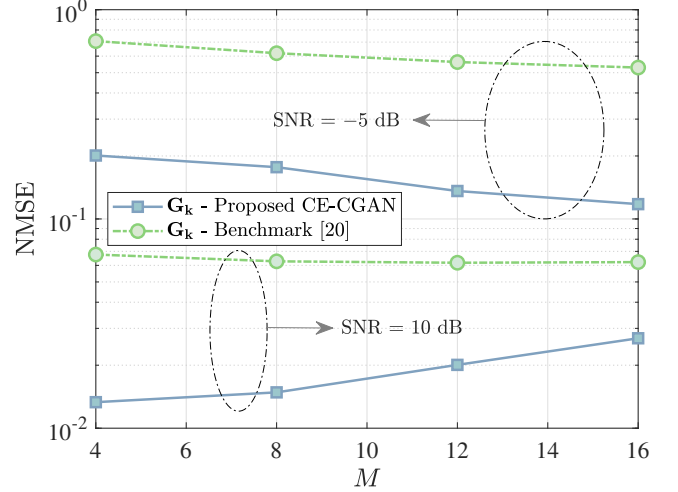
rior channel estimation accuracy. The CGAN robustness in handling higher dimensions can be attributed to its ability to better model complex channel conditions and effectively learn from the discriminator feedback, even for larger sets of training data. Furthermore, the decrease in the NMSE as M increases at SNR $= -5\,\mathrm{dB}$ indicates that the proposed framework is capable of leveraging the additional information present in the received signals to refine its estimation. This robustness stems from the ability of CGANs to capture and model the inherent spatial correlations in the communication channel. By leveraging its generation mechanism, the CGAN framework enables more effective utilization of the additional spatial diversity introduced as M increases. This results in enhanced noise suppression at low SNR values. This is a particularly useful feature for ISAC systems (i.e., where lower SNR values are adopted). At SNR $= 10\,\mathrm{dB}$, where the noise level is relatively low, the primary challenge of the proposed CGAN shifts from combating noise to accurately capturing the high-dimensional channel details. The generator network may face difficulties in fully capturing the high-dimensional dependencies, which results in a gradual increase in estimation error, especially that the overall network architecture and hyperparameters are optimized at lower channel dimensions (i.e., $M = 4$ and $N = 30$). This slight degradation at higher SNR can be attributed to increased sensitivity to minor inaccuracies in capturing subtle channel characteristics. Despite this, the NMSE increase remains minimal, demonstrating the CGAN's effectiveness in maintaining reliable performance. Moreover, the proposed approach consistently outperforms the previous work, showcasing its adaptability and robustness across diverse real-world conditions.

Fig. 5 demonstrates the performance of the proposed approach in estimating the sensing channel, $\mathbf{A}$, with respect to varying M under the same SNR conditions considered in Fig. 4. It can be seen that the proposed SE-CGAN algorithm significantly outperforms the benchmark in [20] at both SNR levels, evidencing its enhanced ability to estimate sensing channels
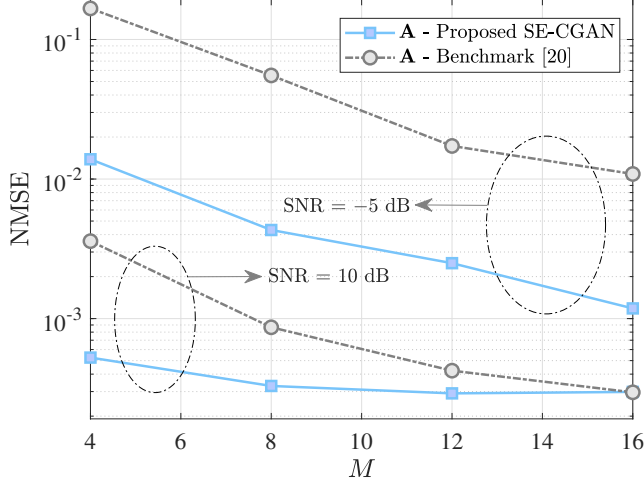
Fig. 5. NMSE performance for the sensing channel estimation versus M at $N = 30$.

Fig. 6. NMSE performance for the communication channel estimation versus N at $M = 4$.

under different channel dimension and noise conditions. At the SNR of $-5\,\mathrm{dB}$, the NMSE of both models decreases, yet the proposed SE-CGAN maintains a stable performance enhancements, showcasing its robustness against high noise levels. The consistent performance across increasing M values suggests that the proposed SE-CGAN effectively utilizes additional antennas to mitigate noise through its adaptive learning and channel modeling capabilities. On the other hand, at the SNR of $10\,\mathrm{dB}$, although the NMSE for the benchmark model decreases with higher M, the proposed algorithm still holds a significant advantage. In particular, it begins with an outperforming estimation for smaller number of antennas, demonstrating robust performance right from the onset. The performance of the proposed approach saturates as M increases, which proves its scalability and capability to handle high-dimensional channel environments without substantial loss in performance. At $M = 16$, both models converge to similar levels of NMSE, illustrating that while the benchmark in [20] improves, the proposed SE-CGAN effectively sustains its superior channel estimation capabilities across varying antenna configurations. This balance of initial superiority and scalability at high SNR settings emphasizes the proposed SE-CGAN advantage in complex RIS-assisted ISAC systems where both accuracy and adaptability are crucial.

D. Impact of Varying N

Fig. 6 demonstrates the performance of the proposed approach in estimating the communication channel, $\mathbf{G}_k$, with respect to varying N under the same SNR conditions considered throughout the simulations. Similarly, the proposed CGAN algorithm effectively outperforms the benchmark scheme in [20] across all N values and SNR conditions. At lower SNR (i.e., SNR $= -5\,\mathrm{dB}$), the proposed algorithm demonstrates a substantial improvement over the benchmark, with the NMSE slightly decreasing as N increases. This significant reduction emphasizes the CGAN robustness in noisy environments, effectively leveraging the increased number of RIS elements to enhance channel estimation. Alternatively, at higher SNR (i.e.,
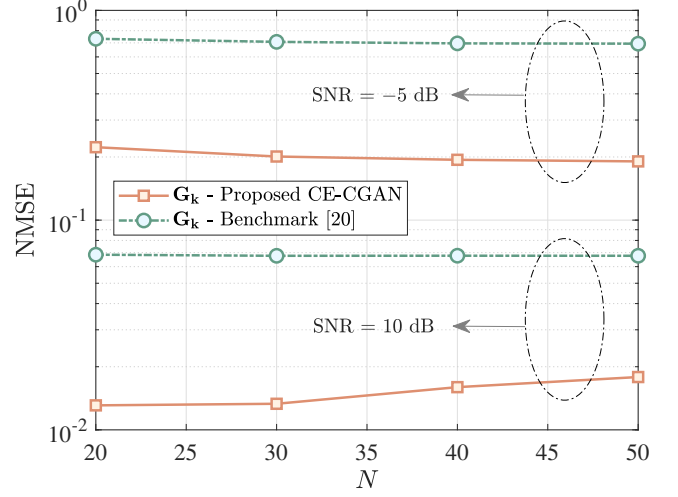
SNR $= 10\,\mathrm{dB}$), although the CGAN continues to outperform the benchmark, the performance improvement decreases, indicating less gains from additional RIS elements. However, it is worth noting that the NMSE remains relatively stable even with increasing N, reflecting the CGAN ability to maintain its estimation accuracy across larger RIS setups. The slight increase in NMSE at higher SNR is marginal and does not significantly impact the estimation performance. It is also important to note that the model was optimized for a baseline configuration (i.e., $M = 4$ and $N = 30$), which explains the observed behavior when scaling beyond these dimensions. Overall, the performance proves the CGAN scalability to larger configurations and shows the CGAN proficiency in utilizing the spatial diversity offered by larger RIS setups to optimize channel estimation, making it particularly useful for ISAC systems operating across a range of SNR scenarios.

E. Complexity Evaluation

Figs. 7a and 7b show the computational complexity of the proposed CGAN approach as compared to the benchmark scheme presented in [20]. The complexity is analyzed in terms of the required number of real additions and multiplications according to the derived formulations in Section IV. Furthermore, the figures show the complexity reduction of using the proposed CGAN algorithm over the benchmark scheme, which is expressed as

$$\text{Reduction} = \frac{C_\chi^{[20]} - C_\chi^\xi}{C_\chi^{[20]}}, \ \chi \in \{\mathcal{A}, \mathcal{M}\}, \xi \in \{\text{SE}, \text{CE}\}. \tag{47}$$

Fig. 7a illustrates the computational complexity and reduction percentage of estimating the direct channel (i.e., $\mathbf{A}$) as M increases. As can be seen, the complexities of additions and multiplications for both the proposed and benchmark schemes show a trend of increasing with M. However, the proposed algorithm complexity is notably lower compared to the benchmark method. The reduction in complexity, depicted on the secondary (i.e., right) y-axis, emphasizes the efficiency gains

Fig. 7. Computational complexity evaluation. (a) Sensing link, (b) Communication link.

from the proposed approach as M increases. The proposed SE-CGAN approach not only significantly reduces the computational complexity compared to the benchmark, enhancing system efficiency, but also delivers a superior performance across various conditions, making it an exceptionally effective solution for ISAC RIS-assisted systems.

Fig. 7b illustrates the computational complexity and reduction percentage of estimating the communication channel (i.e., $\mathbf{G}_k$) as N increases. Similarly, the number of additions and multiplications for both the proposed and benchmark schemes rapidly increase as N increases. The computational complexity of the proposed CE-CGAN approach is comparable to the benchmark scheme, even when the channel estimation dimension enlarges. This is due to the fact that estimating the communication channel is challenging and often require complex models. Despite the inherent challenges, our approach achieves a similar level of computational complexity to the benchmark, while significantly outperforming it in terms of performance. This proves the effectiveness of our model in handling the delicate features of cascaded communication channels estimation within RIS-assisted ISAC systems, ensuring practicality and timeliness in real-world deployments.

VI. Conclusion

This paper has investigated the channel estimation problem of an RIS-assisted ISAC system. A novel CGAN approach has been proposed to enhance the estimation accuracy and stability. The proposed method has leveraged the adversarial training of two deep learning networks to accurately estimate the channel coefficients, demonstrating a superior performance over conventional techniques. The numerical results have validated the efficiency of the proposed approach across different SNR conditions and system dimensions, highlighting its robustness and adaptability. Furthermore, the complexity of the proposed approach has been analyzed and compared to that of the benchmark scheme. In particular, the proposed CE-CGAN model maintains a computational complexity comparable to that of existing DL methods while achieving a better estimation accuracy, while the SE-CGAN model outperforms the existing estimation model in both performance and computational complexity. This makes the proposed CGAN a promising solution for enhancing the reliability and efficiency of RIS-assisted ISAC systems. Future works could explore extensions of the proposed CGAN-based channel estimation approach to address further practical challenges. Examples include accommodating multi-target scenarios along with mobility of users and targets, and adapting the framework for wideband system deployments.

References

[1] A. Fayad, T. Cinkler, and J. Rak, "Toward 6G optical fronthaul: A survey on enabling technologies and research perspectives," *IEEE Commun. Surveys Tuts.*, 2nd Quart. 2024, DOI: 10.1109/COMST.2024.3408090.

[2] 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. 24 634–24 658, Jul. 2024.

[3] C. B. Barneto, S. D. Liyanaarachchi, M. Heino, T. Riihonen, and M. Valkama, "Full duplex radio/radar technology: The enabler for advanced joint communication and sensing," *IEEE Wirel. Commun.*, vol. 28, no. 1, pp. 82–88, Feb. 2021.

[4] F. Bozorgi, P. Sen, A. N. Barreto, and G. Fettweis, "RF front-end challenges for joint communication and radar sensing," in *Proc. 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S)*, 2021, pp. 1–6.

[5] L. Han and K. Wu, "Multifunctional transceiver for future intelligent transportation systems," *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 7, pp. 1879–1892, May 2011.

[6] N. Su, F. Liu, and C. Masouros, "Secure radar-communication systems with malicious targets: Integrating radar, communications and jamming functionalities," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 83–95, Sep. 2021.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] H. Li, S. Shen, M. Nerini, and B. Clerckx, "Reconfigurable intelligent surfaces 2.0: Beyond diagonal phase shift matrices," *IEEE Commun. Mag.*, vol. 62, no. 3, pp. 102–108, Mar. 2024.
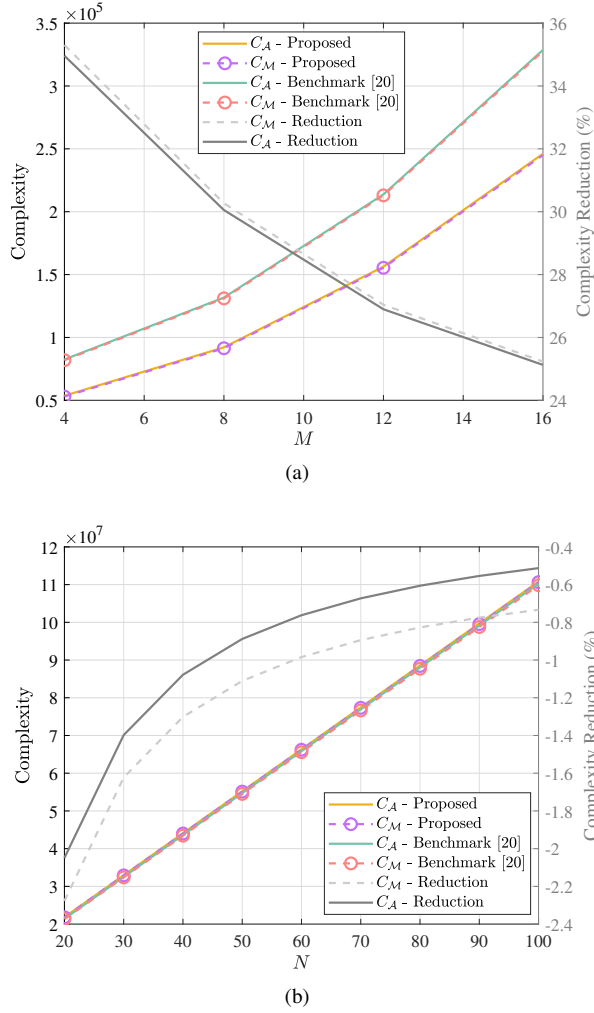
[12] J. Chu, Z. Lu, R. Liu, M. Li, and Q. Liu, "Joint beamforming and reflection design for secure RIS-ISAC systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 3, pp. 4471–4475, Oct. 2024.

[13] Q. Zhu, M. Li, R. Liu, and Q. Liu, "Cramér-rao bound optimization for active RIS-empowered ISAC systems," *IEEE Trans. Wireless Commun.*, Apr. 2024, DOI: 10.1109/TWC.2024.3384501.

[14] 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, Aug. 2024.

[15] 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.

[16] 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, Feb. 2023.

[17] 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.*, May 2024, DOI: 10.1109/LWC.2024.3402235.

[18] W. Jiang, D. Ma, Z. Wei, Z. Feng, P. Zhang, and J. Peng, "ISAC-NET: Model-driven deep learning for integrated passive sensing and communication," *IEEE Trans. Commun.*, Mar. 2024, DOI: 10.1109/TCOMM.2024.3375818.

[19] 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.

[20] ——, "Deep-learning-based channel estimation for IRS-assisted ISAC system," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2022, pp. 4220–4225.

[21] 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.

[22] Y. Liu, I. Al-Nahhal, O. A. Dobre, and F. Wang, "Chapter 4 - channel estimation in ISAC-IRS systems using machine learning approaches," in *Integrated Sensing and Communications for Future Wireless Networks*, A. Kaushik, Ed. Academic Press, 2025, pp. 97–125.

[23] M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Hybrid-layers neural network architectures for modeling the self-interference in full-duplex systems," *IEEE Trans. Vehicular Technol.*, vol. 71, no. 6, pp. 6291–6307, Jun 2022.

[24] M. Elsayed, A. A. Aziz El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Full-duplex self-interference cancellation using dual-neurons neural networks," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 557–561, Mar. 2022.

[25] M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Low complexity neural network structures for self-interference cancellation in full-duplex radio," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 181–185, 2021.

[26] X. Wang, Z. Fei, J. A. Zhang, and J. Xu, "Partially-connected hybrid beamforming design for integrated sensing and communication systems," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6648–6660, Oct. 2022.

[27] 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.

[28] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 3, pp. 990–1002, Sep. 2020.

[29] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[30] 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 International Conference on Communications (ICC)*, 2016, pp. 1–6.

[31] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 2672–2680.

[32] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2242–2251.

[33] H.-W. Dong, W.-Y. Hsiao, L.-C. Yang, and Y.-H. Yang, "MuseGAN: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment," in *Proc. AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018, pp. 34–41.

[34] J. Kong, J. Kim, and J. Bae, "Hifi-gan: Generative adversarial networks for efficient and high fidelity speech synthesis," in *Proc. Neural Inf. Process. Syst.*, vol. 33. Curran Associates, Inc., 2020, pp. 17 022–17 033.

[35] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "ESRGAN: Enhanced super-resolution generative adversarial networks," in *Proc. European Conference on Computer Vision (ECCV) Workshops*, 2018, pp. 63–79.

[36] C. Zou, F. Yang, J. Song, and Z. Han, "Generative adversarial network for wireless communication: Principle, application, and trends," *IEEE Commun. Mag.*, vol. 62, no. 5, pp. 58–64, May 2024.

[37] T. Erpek, Y. E. Sagduyu, and Y. Shi, "Deep learning for launching and mitigating wireless jamming attacks," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 1, pp. 2–14, Mar. 2019.

[38] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks," *IEEE Access*, vol. 6, pp. 15 713–15 722, Mar. 2018.

[39] Z. Liu, X. Jing, R. Zhang, and J. Mu, "Spectrum sensing based on deep convolutional generative adversarial networks," in *Proc. International Wireless Communications and Mobile Computing (IWCMC)*, 2021, pp. 796–801.

[40] Z. Li, X. Liao, J. Shi, L. Li, and P. Xiao, "MD-GAN-based UAV trajectory and power optimization for cognitive covert communications," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10 187–10 199, Jun. 2022.

[41] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014. [Online]. Available: https://arxiv.org/abs/1411.1784

[42] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. [Online]. Available: https://arxiv.org/abs/1412.6980

[43] 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.

[44] 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.

[45] 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.

[46] 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.