

Learning to Design Transceiver for Integrated Sensing and Communications: A Satellite Communications Perspective

William D. Lukito, *Graduate Student Member, IEEE*, Wei Xiang, *Senior Member, IEEE*, Chang Liu, *Member, IEEE*, Phu Lai, Peng Cheng, *Member, IEEE*, and Guoqiang Mao, *Fellow, IEEE*

Abstract—With its dual-functional advantages, integrated sensing and communications (ISAC) technologies can be further extended to satellite communications, enhancing global coverage services. However, achieving vast coverage would result in significant delays and considerable path losses. Motivated by this, in this paper, we focus on satellite-based ISAC (S-ISAC) systems and propose a general transceiver design framework incorporating both transmit waveform and receive filter. Unlike existing approaches, our approach uses a predictive joint transmit waveform and receive filter design that eliminates the need of channel estimation, thereby reducing time overhead. Additionally, a versatile weighting mechanism is designed to allow flexible prioritization between communications and sensing. To tackle the intractability of the ISAC transceiver design problem, we adopt a data-driven deep learning-based approach, where the model learns to design the transmit waveform and receive filter from historical channel data. Specifically, we propose a predictive optimization network (PONet), leveraging convolutional layers and a Transformer encoder to capture long-term spatial-temporal features and facilitate the learning capability. Numerical results demonstrate the effectiveness of the proposed PONet in terms of communications and sensing rates in S-ISAC networks in various system settings.

Index Terms—Integrated sensing and communications, deep learning, waveform, filter, satellite, non-terrestrial networks, Transformer.

I. INTRODUCTION

INTEGRATED sensing and communications (ISAC) has become a crucial technology for the future of wireless systems [1]. The key goal of ISAC is to integrate sensing and communications into a common resource and device, thereby enhancing spectrum efficiency and reducing hardware costs [1]–[3]. To implement ISAC effectively, it is crucial to design a waveform optimized for both communications and sensing [4], as they have differing characteristics but rely on

the same resources and devices [5], [6]. While much works have been done on the ISAC waveform design for terrestrial networks [5]–[13], satellite communications as a network platform have been largely overlooked. Satellite communications offer significant advantages, enabling ISAC systems to operate over vast areas, provide global coverage, and support applications in remote or hard-to-reach regions [1], [14]. However, due to the large distances between satellites and terrestrial systems, two unique challenges are introduced to satellite communications systems.

Firstly, the signal echoes experience significant path loss, weakening them considerably, which necessitates the use of receive filter to extract meaningful information by maximizing the useful signal [6], [15]–[17]. Therefore, in this work, we design a satellite-based ISAC (S-ISAC) transceiver framework that consists of transmit waveform and receive filter parts to represent the transmitter and receiver, respectively. Although there have been several efforts on S-ISAC [18]–[21], none have focused on the transceiver design, particularly the joint transmit waveform and receive filter design. This joint design is imperative to achieve more optimal results for practical ISAC systems, especially the sensing part, as it is affected by both transmit waveform and receive filter [6]. Secondly, designing transmit waveform and receive filter typically requires channel estimation [6], [17] that can be achieved via pilot-based [22] or two-stage [23] procedures. However, in satellite communications, this process incurs significant time overhead due to long propagation delays, which can result in outdated sensing information by the time it is extracted.

As inspired by predictive beamforming approaches for terrestrial-based ISAC systems [24], [25], we propose a predictive joint transmit waveform and receive filter scheme that eliminates the need for extensive real-time channel estimation to address the challenges caused by large path losses and delays in S-ISAC networks. By leveraging historical channel state information, our approach enables the prediction of the transmit waveform and receive filter for the next time frame, reducing overhead and ensuring sensing results as up-to-date as possible. This predictive scheme is especially beneficial for S-ISAC systems, where minimizing latency and optimizing both communications and sensing tasks are essential for efficient operation. However, implementing the predictive transmit waveform and receive filter scheme requires a design procedure that must be executed for each time frame.

Note that the joint transmit waveform and receive filter

Manuscript received Month XX, XXXX; revised Month XX, XXXX; accepted Month XX, XXXX. This work has been supported by the SmartSat CRC, whose activities are funded by the Australian Government's CRC Program. This work was supported in part by the Australian Government through the Australian Research Council's Discovery Projects Funding Scheme under Project DP220101634 (*Corresponding author: Wei Xiang*).

W. D. Lukito, W. Xiang, C. Liu, P. Lai, and P. Cheng are with the School of Computing, Engineering and Mathematical Sciences, La Trobe University, Melbourne, VIC 3086, Australia (e-mail: w.lukito@latrobe.edu.au, w.xiang@latrobe.edu.au, c.liu6@latrobe.edu.au, p.lai@latrobe.edu.au, p.cheng@latrobe.edu.au).

G. Mao is with the Research Laboratory of Smart Driving and Intelligent Transportation Systems, Southeast University, Nanjing, Jiangsu 211102, China (e-mail: g.mao@ieee.org).

design is generally a high-dimensional complex task [6], thus it is extremely challenging to acquire an optimal solution, especially we eliminate the need of channel estimation. Meanwhile, deep learning (DL)-based approaches offer a promising alternative for optimization via their powerful data-driven capability in various wireless communications problems, as it can transfer a complex computational task to the offline training and online inference phases. DL is highly effective in extracting meaningful features from data [9], and it has been proven for prediction scenarios in wireless communications systems [24], [25]. Moreover, DL has proven effective in various optimization tasks in terrestrial-based ISAC networks [8]–[11]. However, the works in [24] and [25] focus solely on terrestrial communications, without considering satellite communications and non-line-of-sight (NLOS) channels. Additionally, their reliance on long-short term memory (LSTM) models limits their ability to capture long-term dependencies effectively. As highlighted in [26], DL techniques for ISAC are also envisioned to enable effective coordination between communications and sensing tasks. In this work, we specifically leverage the advantage of convolutional neural networks (CNNs) for extracting spatial information [27] and the Transformer for capturing long-term temporal features [28].

Motivated by the aforementioned gaps and the advantages of DL, in this work, we employ a data-driven DL-based approach to design a predictive joint transmit waveform and receive filter for an S-ISAC transceiver. Our proposed predictive scheme utilizes DL to learn from historical channel information and directly predict the transmit waveform and receive filter for the next time frame, thereby significantly reducing execution time. Additionally, to ensure flexibility in terms of task preference for the ISAC deployment, we adopt a joint communications-sensing objective [6], [9], [11], [13] as the criteria for designing transmit waveform and receive filter in this work. The main contributions of this work can be summarized as follows:

- 1) We design a general transceiver framework for S-ISAC systems by formulating a general predictive joint transmit waveform and receive filter optimization problem that maximizes network utility function which aggregates both communications and sensing utilities, allowing for the flexibility of communications and sensing preference. This is achieved by utilizing a weighted normalized sum-rate to aggregate the communications and sensing utilities, subject to power constraints.
- 2) We propose a predictive design procedure to jointly design the transmit waveform and receive filter that eliminates the need for real-time channel estimation for each time frame, allowing the design to be executed using historical channel information, thereby reducing the overall time overhead. This procedure can be applied to any predictive design tasks for S-ISAC transceivers.
- 3) We propose a DL-based data-driven framework to address the predictive joint transmit waveform and receive filter design problem for S-ISAC transceivers, utilizing a penalty method and normalization to ensure power constraints are satisfied. Specifically, we propose a predictive optimization network (PONet) that effectively

captures spatial and long-term temporal features using convolutional layers and a Transformer encoder, respectively. Numerical results verify that the proposed network outperforms the baselines in terms of communications and sensing rates.

The remainder of this paper is organized as follows. Section II presents the system model of our considered ISAC framework and a general problem formulation for designing ISAC transceiver is developed. Then, in Section III, a predictive ISAC transmit waveform and receive filter design procedure is proposed to eliminate explicit channel estimation. In Section IV, we introduce a DL-based solution for predictive transmit waveform and receive filter design to solve the formulated problem. Section V verifies the proposed solution through several simulation results considering a satellite-based scenario. Finally, Section VI provides the conclusion drawn from this work.

Notation: Unless otherwise specified, bold uppercase letters, bold lowercase letters, and normal letters/symbols represent matrices, vectors, and scalars, respectively. Constants c and j denote the speed of light ($c \approx 3 \times 10^8$ meters per second) and the imaginary unit, respectively. Notations $\mathbb{R}^{a \times b}$ and $\mathbb{C}^{a \times b}$ represent real and complex number sets of size a -by- b , respectively. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the circularly symmetric complex Gaussian (CSCG) distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathcal{U}(a, b)$ denotes the uniform distribution between a and b . \mathbf{I}_N denotes N -by- N identity matrix. The function $\text{diag}(\mathbf{x})$ creates a diagonal matrix from vector \mathbf{x} . Superscripts T , $*$, and H denote transpose, conjugate, and Hermitian matrices/vectors operations, respectively. Operators \times and \otimes indicate scalar multiplication and the Kronecker product, respectively. The Euclidean distance, Frobenius norm, and absolute value are represented by $\|(\cdot)\|$, $\|(\cdot)\|_F^2$, and $|(\cdot)|$, respectively. Finally, $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ extract the real and imaginary parts of a complex value.

II. SYSTEM MODEL & PROBLEM FORMULATION

We consider an S-ISAC framework within an existing satellite-based IoT network [29], where we enable ISAC by assigning a satellite as an ISAC transceiver. This framework is operated at carrier frequency f_c . The satellite is equipped with a pair of transmit and radar receive uniform planar array (UPA) antennas. The transmit antenna array consists of $N_T = N_T^x \times N_T^y$ elements, and the receive antenna array consists of $N_R = N_R^x \times N_R^y$ elements. Single transmit and receive antennas have gains of G_T and G_R , respectively. The antennas along the x - and y -axes are separated by distances $\delta_x = 0.5\lambda_c$ and $\delta_y = 0.5\lambda_c$, respectively, where $\lambda_c = c/f_c$ is the carrier wavelength. The satellite employs an advance full-duplex technique [30], [31], and it can receive any signals during transmission without leakages from the transmitter side [17]. The position of the satellite is represented by vector $\mathbf{p}_{n,s} = [x_{n,s}, y_{n,s}, z_{n,s}]$. Fig. 1 illustrates the considered S-ISAC-assisted IoT network. It is worth noting that S-ISAC is essential for scenarios involving widely distributed sensing targets where terrestrial networks are ineffective. To illustrate the components of the satellite, Fig. 2 is presented.

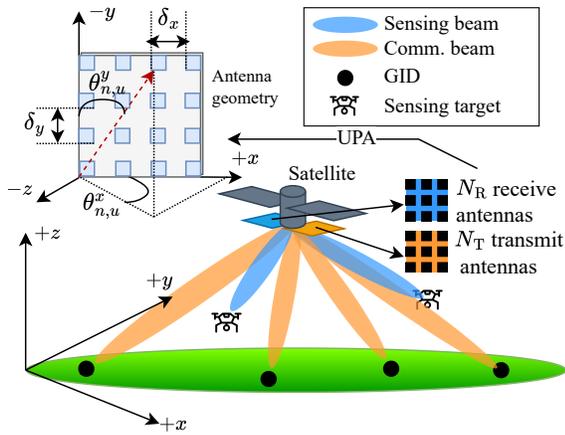


Fig. 1. ISAC-assisted satellite IoT network under consideration, featuring a satellite that communicates with multiple GIDs and senses multiple targets, where $u \in \{k, q\}$.

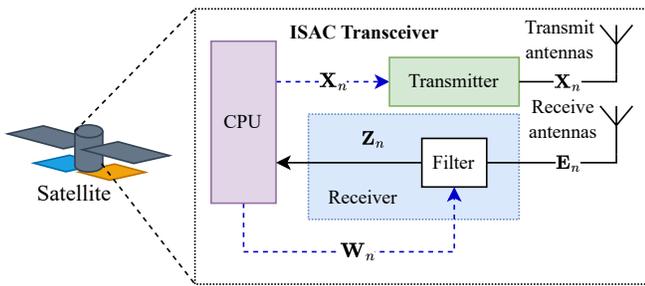


Fig. 2. Schematic of the considered S-ISAC transceiver where the design process of ISAC transmit waveform and receive filter is executed on a central processing unit (CPU).

During the n -th time frame, $n \in \{0, \dots, N-1\}$, there is a transmit waveform with length of L symbols¹. Due to the nature of inactive and active modes of ground IoT devices (GIDs) [32]–[34], satellite serves $K = \xi_K \times K_{\text{total}}$ GIDs per time frame, where $\xi_K \in (0, 1]$ and K_{total} denote the active rate and the total number of GIDs, respectively. The indices of active GIDs in the n -th time frame are represented by the set \mathcal{K}_n . At the same time, there are Q sensing targets of interest flying within the BS ground footprint area. The indices of sensing targets in the n -th time frame are represented by the set \mathcal{Q}_n . The position of active GIDs and sensing targets during the n -th time frame is represented by the following vector $\mathbf{p}_{n,u} = [x_{n,u}, y_{n,u}, z_{n,u}]$, where $u \in \{k, q\}$, $k \in \mathcal{K}_n$, and $q \in \mathcal{Q}_n$.

A. Communications Model

During the n -th time frame, the received signal matrix at active downlink GIDs is denoted as $\mathbf{Y}_n = [\mathbf{y}_{n,1}, \dots, \mathbf{y}_{n,K}]^T \in \mathbb{C}^{K \times L}$ and can be expressed as follows

$$\mathbf{Y}_n = \mathbf{H}_n \mathbf{X}_n + \mathbf{N}_n, \quad (1)$$

¹In this work, a single time frame can also be considered as a single pulse from the perspective of radar signal processing. Therefore, the duration of a time frame can be represented by a pulse repetition interval (PRI) [6], [24], [25].

where $\mathbf{H}_n = [\mathbf{h}_{n,1}, \dots, \mathbf{h}_{n,K}]^T \in \mathbb{C}^{K \times N_T}$ is the channel state information (CSI) matrix in the n -th time frame, $\mathbf{X}_n = [\mathbf{x}_{n,1}, \dots, \mathbf{x}_{n,L}] \in \mathbb{C}^{N_T \times L}$ is the transmit waveform matrix in the n -th time frame, $\mathbf{x}_{n,l}$ is the transmit waveform vector for the l -th symbol in the n -th time frame, and $\mathbf{N}_n = [\mathbf{n}_{n,1}, \dots, \mathbf{n}_{n,K}]^T \in \mathbb{C}^{K \times L}$ is the noise matrix in the n -th time frame, where $\mathbf{n}_{n,k} \sim \mathcal{CN}(0, \sigma_N^2 \mathbf{I}_L)$, and σ_N^2 is the noise variance. In our considered system model, \mathbf{X}_n serves as a common transmit waveform for both the sensing and communications tasks.

In practice, the channel between satellite and the k -th active GID is subject to propagation delay and Doppler shift. However, in this work, it is assumed that the GIDs are equipped with mechanisms to counteract these effects, enabling their mitigation at the GID end [29], [35], [36]. According to references [18], [34], [37], we model satellite channel model with Rician fading. Therefore, the channel vector between satellite and the k -th GID during the n -th time frame can be written as follows

$$\mathbf{h}_{n,k} = \mathbf{h}_{n,k}^{\text{LOS}} + \mathbf{h}_{n,k}^{\text{NLOS}}, \quad (2)$$

where $\mathbf{h}_{n,k}^{\text{LOS}}$ and $\mathbf{h}_{n,k}^{\text{NLOS}}$ are the line-of-sight (LOS) and NLOS components, respectively. $\mathbf{h}_{n,k}^{\text{LOS}}$ and $\mathbf{h}_{n,k}^{\text{NLOS}}$ can be further expanded as follows [18], [37]

$$\mathbf{h}_{n,k}^{\text{LOS}} = \sqrt{\frac{\kappa_R \varrho_{n,k}}{\kappa_R + 1}} \mathbf{a}_{n,k}^x \otimes \mathbf{a}_{n,k}^y, \quad (3)$$

$$\mathbf{h}_{n,k}^{\text{NLOS}} \sim \mathcal{CN}\left(0, \frac{\varrho_{n,k}}{\kappa_R + 1} \mathbf{I}_{N_T}\right), \quad (4)$$

where κ_R is the Rician factor, $\varrho_{n,k} = G_T N_T \left(\frac{c}{4\pi f_c d_{n,k}}\right)^2$ is the channel gain, and $d_{n,k} = \|\mathbf{p}_{n,s} - \mathbf{p}_{n,k}\|$ is the distance between the k -th GID and satellite in the n -th time frame. Due to the considerable distances involved, the channel gain is significantly weakened, impacting both (3) and (4). Meanwhile, $\mathbf{a}_{n,k}^x$ and $\mathbf{a}_{n,k}^y$ represent the array response vectors for x - and y -axis and can be written as follows

$$\mathbf{a}_{n,k}^x = \frac{1}{\sqrt{N_T^x}} [1, e^{-j\vartheta_{n,k}^x}, \dots, e^{-j\vartheta_{n,k}^x(N_T^x-1)}], \quad (5)$$

$$\mathbf{a}_{n,k}^y = \frac{1}{\sqrt{N_T^y}} [1, e^{-j\vartheta_{n,k}^y}, \dots, e^{-j\vartheta_{n,k}^y(N_T^y-1)}], \quad (6)$$

where constants $\vartheta_{n,k}^x = \frac{2\pi f_c \delta_x}{c} \sin(\theta_{n,k}^y) \cos(\theta_{n,k}^x)$ and $\vartheta_{n,k}^y = \frac{2\pi f_c \delta_y}{c} \cos(\theta_{n,k}^y)$, while $\theta_{n,k}^x$ and $\theta_{n,k}^y$ are the angle of departure pair for the k -th GID at the n -th time frame. According to [5], we assume that the downlink channel \mathbf{H}_n is considered static over a single time frame. We also consider CSI to be correlated across different time frames. Please refer to our simulation settings in Section V for further explanation.

Regarding the information symbols, let $\mathbf{S}_n = [\mathbf{s}_{n,1}, \dots, \mathbf{s}_{n,K}] \in \mathbb{C}^{K \times L}$ denote the received symbol matrix. Under an ideal condition, the received signal would match with the desired symbol matrix, such that $\mathbf{Y}_n = \mathbf{S}_n$. However, due to multi-user interference (MUI) and noise, achieving perfect conditions is impossible. By adopting

per-symbol precoding [5], [6], given \mathbf{S}_n , the received signals at the GIDs can be re-expressed as follows

$$\mathbf{Y}_n = \mathbf{S}_n + \underbrace{(\mathbf{H}_n \mathbf{X}_n - \mathbf{S}_n)}_{\text{MUI}} + \mathbf{N}_n, \quad (7)$$

where for each user, the received signal can be extracted as follows

$$\mathbf{y}_{n,k} = \mathbf{s}_{n,k} + \underbrace{(\mathbf{h}_{n,k} \mathbf{X}_n - \mathbf{s}_{n,k})}_{\text{MUI}} + \mathbf{n}_{n,k}, \quad (8)$$

where $\mathbf{s}_{n,k}$ represents the sequence of symbols for the k -th GID in the n -th time frame. Therefore, according to (8), the signal power can be expressed as $\mathbb{E}\{|\mathbf{s}_{n,k}|^2\}$ and MUI can be expressed as $\mathbb{E}\{|\mathbf{h}_{n,k} \mathbf{X}_n - \mathbf{s}_{n,k}|^2\}$, where $\mathbb{E}\{\cdot\}$ denotes the expected value operation with respect to the time frame index. As a result, the signal-to-interference-and-noise ratio (SINR) for the k -th active GID in the n -th time frame can be written as follows

$$\gamma_{n,k} = \frac{\mathbb{E}\{|\mathbf{s}_{n,k}|^2\}}{\mathbb{E}\{|\mathbf{h}_{n,k} \mathbf{X}_n - \mathbf{s}_{n,k}|^2\} + \sigma_N^2}, \quad (9)$$

and it becomes obvious that the SINR can be maximized by minimizing MUI [5], [6].

Remark 1: Although per-symbol precoding may require the channel to change slowly, which can be challenging in satellite communications due to their dynamic nature, it remains practical in scenarios with low-mobility users or stable beamforming configurations [38]–[40]. Furthermore, channel prediction techniques i.e., [41], [42], can be employed to enable per-symbol precoding under more dynamic conditions, ensuring the practicality of this approach in real-world applications.

Thus far, the necessity of CSI for maximizing the SINR has been unequivocally demonstrated, as evidenced by its role in affecting MUI, as shown in (9). By having an accurate CSI, it is possible to optimize the transmit waveform \mathbf{X}_n to substantially minimize MUI. According to Shannon's theorem, enhancing SINR directly implies into higher achievable communications rates [43]. Consequently, within an orthogonal multiple access (OMA) framework², the communications rate achievable by the active GIDs for the n -th time frame can be articulated as a function of \mathbf{X}_n , thereby establishing a weighted sum-rate (WSR) formulation as follows³

$$R_C(\mathbf{X}_n) = \frac{1}{K} \sum_{k \in \mathcal{K}_n} \beta_k \log_2(1 + \gamma_{n,k}), \quad (10)$$

where $\beta_k \in (0, 1]$ denotes the priority of the k -th active GID [44], determined according to their service level agreement (SLA).

²The application of non-orthogonal multiple access (NOMA) could be explored; however, this would require additional considerations for power allocation. To simplify our analysis, NOMA will be deferred to future work, and only OMA is addressed in this study.

³Although (10) is specific to each frame (instantaneous WSR), we consider all WSR values across multiple time frames in the expectation term of the subsequent problem formulation.

B. Sensing Model

To initialize the sensing model, we begin with considering Q sensing targets of interest. By utilizing the same transmit waveform \mathbf{X}_n , the received echoes in satellite $\mathbf{E}_n \in \mathbb{C}^{N_R \times L}$ can be written as follows

$$\mathbf{E}_n = \mathbf{G}_n \mathbf{X}_n + \mathbf{M}_n, \quad (11)$$

where $\mathbf{G}_n \in \mathbb{C}^{N_R \times N_T}$ is the target response matrix (TRM) and $\mathbf{M}_n = [\mathbf{m}_{n,1}, \dots, \mathbf{m}_{n,L}] \in \mathbb{C}^{N_R \times L}$ is the noise matrix with $\mathbf{m}_{n,l} \sim \mathcal{CN}(0, \sigma_N^2 \mathbf{I}_{N_R})$. The TRM can be further written as follows [9], [24], [37]

$$\mathbf{G}_n = \Gamma \sum_{q \in \mathcal{Q}_n} \xi_{n,q} (\mathbf{b}_{n,q}^x \otimes \mathbf{b}_{n,q}^y)^H \mathbf{a}_{n,q}^x \otimes \mathbf{a}_{n,q}^y, \quad (12)$$

where $\Gamma = \sqrt{G_T N_T G_R N_R}$ is the total antenna gain, $\xi_{n,q} = \frac{c_q}{2d_{n,q}}$ is the reflection coefficient of the q -th target, c_q denotes the radar cross section of the q -th target, and $2d_{n,q} = 2\|\mathbf{p}_{n,s} - \mathbf{p}_{n,q}\|$ denotes the round-trip distance from satellite to the q -th target [24]. In contrast to $\mathbf{a}_{n,q}^x$ and $\mathbf{a}_{n,q}^y$ those are similar to (5) and (6), $\mathbf{b}_{n,q}^x$ and $\mathbf{b}_{n,q}^y$ represent the array response vectors for x - and y - axes and can be expanded as follows

$$\mathbf{b}_{n,q}^x = \frac{1}{\sqrt{N_R^x}} [1, e^{-j\vartheta_{n,q}^x}, \dots, e^{-j\vartheta_{n,q}^x(N_R^x-1)}], \quad (13)$$

$$\mathbf{b}_{n,q}^y = \frac{1}{\sqrt{N_R^y}} [1, e^{-j\vartheta_{n,q}^y}, \dots, e^{-j\vartheta_{n,q}^y(N_R^y-1)}], \quad (14)$$

where constants $\vartheta_{n,q}^x = \frac{2\pi f_c \delta_x}{c} \sin(\theta_{n,q}^y) \cos(\theta_{n,q}^x)$ and $\vartheta_{n,q}^y = \frac{2\pi f_c \delta_y}{c} \cos(\theta_{n,q}^y)$, while $\theta_{n,q}^x$ and $\theta_{n,q}^y$ are the angle of arrival pair for the q -th target at the n -th time frame.

After satellite receives \mathbf{E}_n , it reconstructs the TRM \mathbf{G}_n to extract sensing information. However, due to substantial path loss over large distances, the received echoes are significantly attenuated [18]. Without additional processing, the system cannot fully extract the available information from these weakened echoes. Therefore, to obtain more information, a common method is to maximize the received SINR [6]. A higher SINR results in a lower mean-squared error (MSE) of the \mathbf{G}_n extraction, where it is directly related to the maximization of mutual information (MI) or sensing rate [9], [13], [45]. This can be done by designing a receive filter that optimizes the sensing performance is crucial. We model the filtered echo $\mathbf{Z}_n \in \mathbb{C}^{N_R \times L}$ as follows

$$\mathbf{Z}_n = \mathbf{W}_n \mathbf{E}_n, \quad (15)$$

where $\mathbf{W}_n \in \mathbb{C}^{N_R \times N_R}$ is the receive filter matrix.

Remark 2: To sense an entire satellite footprint, we can set Q to approximately infinity ($Q \approx \infty$). Additionally, the sensing targets of interest can be either static or dynamic. However, for simplicity, in this work, we deliberately choose a finite–small number for Q and consider the sensing targets are moving along their trajectories. We will elaborate this further in Section V.

To obtain the sensing rate, we first need to calculate MI between \mathbf{Z}_n and \mathbf{G}_n given \mathbf{X}_n , then divide it by the time

frame duration. According to the information theory, MI can be written as follows

$$I(\mathbf{Z}_n; \mathbf{G}_n | \mathbf{X}_n) = h(\mathbf{Z}_n | \mathbf{X}_n) - h(\mathbf{Z}_n | \mathbf{G}_n, \mathbf{X}_n), \quad (16)$$

where $h(\cdot)$ denotes the differential entropy. By having $h(\mathbf{Z}_n | \mathbf{X}_n)$ and $h(\mathbf{Z}_n | \mathbf{G}_n, \mathbf{X}_n)$, MI can be further expressed as follows

$$I(\mathbf{Z}_n; \mathbf{G}_n | \mathbf{X}_n) = \log_2 \left\{ \frac{\det(\mathbf{W}_n^H \mathbf{R}_{EE} \mathbf{W}_n)}{\det(\sigma_N^2 \mathbf{W}_n^H \mathbf{W}_n)} \right\}, \quad (17)$$

where $\mathbf{R}_{EE} = \mathbf{G}_n \mathbf{X}_n \mathbf{X}_n^H \mathbf{G}_n^H + \sigma_N^2 \mathbf{I}_{N_R}$. For the derivation of the necessary components in (16) leading to (17), please refer to Appendix A. To this point, we show that MI depends on both \mathbf{X}_n and \mathbf{W}_n . Consequently, the sensing rate can be formulated as follows [9], [46]

$$R_S(\mathbf{X}_n, \mathbf{W}_n) = \frac{1}{L} I(\mathbf{Z}_n; \mathbf{G}_n | \mathbf{X}_n) = \frac{1}{L} \log_2 \left\{ \frac{\det(\mathbf{W}_n^H \mathbf{R}_{EE} \mathbf{W}_n)}{\det(\sigma_N^2 \mathbf{W}_n^H \mathbf{W}_n)} \right\}. \quad (18)$$

Remark 3: We consider using the sensing rate as a general metric for the sensing performance. The sensing rate represents the limit of information acquisition, therefore it can represent any physical sensing measurements [13].

Since both the communications and sensing rates depend on \mathbf{X}_n as shown in (10) and (18), it is logical to design an ISAC transmit waveform that enhances both performance metrics. Furthermore, to effectively extract sensing information from the echoes, an optimal receive filter \mathbf{W}_n is necessary to maximize the sensing performance, as shown in (18). Therefore, this work focuses on optimizing both \mathbf{X}_n and \mathbf{W}_n to improve the communications and sensing performances simultaneously.

C. Problem Formulation

The main objective of this paper is to design an ISAC transceiver by maximizing the network utility function via jointly optimizing transmit waveform \mathbf{X}_n and sensing receive filter \mathbf{W}_n subject to the transmit waveform and filter power constraints, given the n -th symbol matrix, CSI, and TRM⁴. Generally, in the n -th frame, the problem can be formulated as follows

$$\max_{\mathbf{X}_n, \mathbf{W}_n} \mathbb{E}_{\mathbf{S}_n, \mathbf{H}_n, \mathbf{G}_n} U \{ U_C(\mathbf{X}_n, \mathbf{S}_n, \mathbf{H}_n), U_S(\mathbf{X}_n, \mathbf{W}_n, \mathbf{G}_n) \}, \quad (19)$$

$$\text{s.t.} \quad \|\mathbf{X}_n\|_F^2 \leq P_{\max}, \quad \|\mathbf{W}_n\|_F^2 \leq 1, \quad (19a)$$

where $U(\cdot)$, $U_C(\cdot)$, and $U_S(\cdot)$ represent the utility functions for the network, communications, and sensing, respectively. The usage of \mathbf{S}_n , \mathbf{H}_n , and \mathbf{G}_n in the objective function indicates which utility depends on it. Meanwhile, the term $\mathbb{E}_{\mathbf{S}_n, \mathbf{H}_n, \mathbf{G}_n} |_{\Omega_{n-1}^\tau, \Phi_{n-1}^\tau}$ in the objective function denotes the ergodic average with respect to \mathbf{S}_n , \mathbf{H}_n and

⁴If we have the information symbols for future time frames, this problem formulation can be extended to consider multiple future time frames, e.g. $\{\mathbf{x}_n, \dots, \mathbf{x}_{n+N}\}$. However, to solve this, a minor modification in the proposed solution is required, specifically an adjustment to the output dimensions.

\mathbf{G}_n , given a sequence of historical estimated CSI, $\Omega_{n-1}^\tau \triangleq [\tilde{\mathbf{H}}_{n-1}, \dots, \tilde{\mathbf{H}}_{n-\tau}]$, and TRM, $\Phi_{n-1}^\tau \triangleq [\tilde{\mathbf{G}}_{n-1}, \dots, \tilde{\mathbf{G}}_{n-\tau}]$, of length τ frames⁵. The use of the ergodic average characterizes the performance of communications and sensing in the n -th frame, since only historical CSI and TRM data (i.e., from frame $n-1$ to $n-\tau$) can be utilized when designing the transmit waveform and receive filter for the n -th frame [43]. As for the constraint, P_{\max} denotes the maximum transmitted power and a unit power is considered for receive filter at satellite for each frame n .

Remark 4: It is important to note that the expectation term in the objective function (19) is formulated to ensure the generalizability of the problem. In the general case, both \mathbf{H}_n and \mathbf{G}_n may depend on both Ω_{n-1}^τ and Φ_{n-1}^τ . For instance, in a scenario where a communications user is also observed as a sensing target. However, in this work, we focus on a specific scenario where \mathbf{H}_n depends solely on Ω_{n-1}^τ and \mathbf{G}_n depends solely on Φ_{n-1}^τ .

Remark 5: Given that $U(\cdot)$ is fundamentally a function of $U_C(\cdot)$ and $U_S(\cdot)$, it follows intuitively that maximizing $U(\cdot)$ involves optimizing both $U_C(\cdot)$ and $U_S(\cdot)$. This approach is applicable to any network utility functions [47], as well as communications and sensing utility functions, provided they share the same directional metric (i.e., higher values indicate better performance). This problem formulation is applicable to any utility functions.

As a demonstration for this paper, we use a weighted sum as the network utility to aggregate the communications and sensing utilities based on their respective normalized rates. Therefore, the objective function can be rewritten as follows⁶

$$U(\mathbf{X}_n, \mathbf{W}_n) \triangleq \frac{\rho_C R_C(\mathbf{X}_n)}{\mu_{C,n}} + \frac{(1 - \rho_C) R_S(\mathbf{X}_n, \mathbf{W}_n)}{\mu_{S,n}}, \quad (20)$$

where ρ_C and $(1 - \rho_C)$ correspond to the weights for communications and sensing utilities, respectively. As a default value, we set $\rho_C = 0.5$. Meanwhile, $\mu_{C,n}$ and $\mu_{S,n}$ denote the maximum rates for communications and sensing as derived in Appendix B.

III. PREDICTIVE ISAC TRANSMIT WAVEFORM AND RECEIVE FILTER DESIGN PROCEDURE FOR SATELLITE-BASED TRANSCEIVER

Prior to signal transmission and echo processing, the ISAC transceiver requires the current CSI \mathbf{H}_n and TRM \mathbf{G}_n to design the transmit waveform \mathbf{X}_n and receive filter \mathbf{W}_n [6], [23]. Traditionally, these information is acquired through a channel estimator, which is typically incorporated into a procedure that must be executed for each time frame n . The existing approaches [22], [23] to obtaining \mathbf{H}_n and \mathbf{G}_n require substantial estimation overhead, therefore, it may lead to outdated

⁵ $\tilde{\mathbf{H}}_n$ is essentially similar to \mathbf{H}_n ; however, for each component $\mathbf{h}_{n,k}$ as described in (2), it represents the estimated GID's distance and direction, $\tilde{d}_{n,k}$ and $\tilde{\theta}_{n,k}$, instead of the exact values. Similarly, $\tilde{\mathbf{G}}_n$ mirrors \mathbf{G}_n as described in (12), but represents the estimated target's distance and direction, $\tilde{d}_{n,q}$ and $\tilde{\theta}_{n,q}$, respectively.

⁶As we aim to eliminate the use of CSI \mathbf{H}_n and TRM \mathbf{G}_n , and given that the symbol matrix \mathbf{S}_n is provided, we omit these elements from (20).

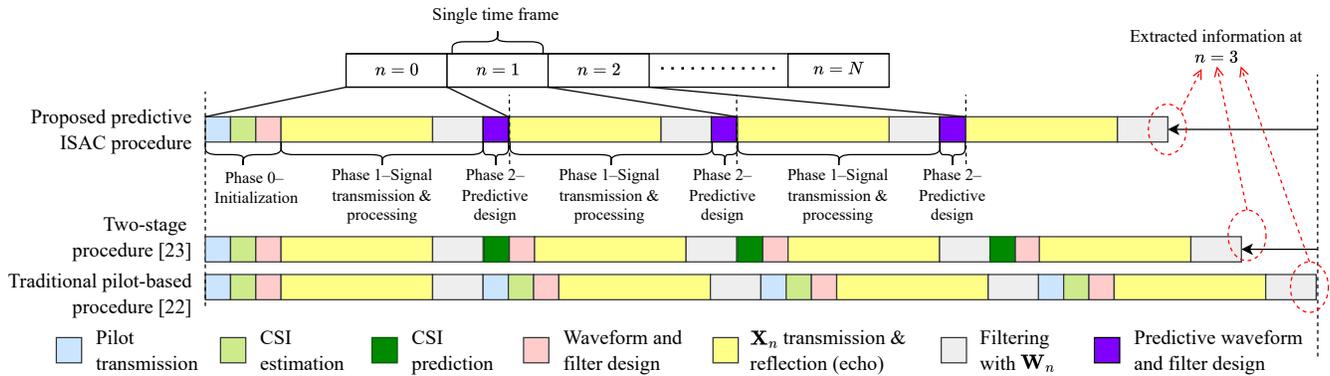


Fig. 3. Proposed predictive ISAC transmit waveform and receive filter design procedure compared with two-stage procedure [23] and traditional procedure [22], where the left arrows indicate the time overhead is being saved. The phase 2–predictive design of our proposed procedure will be realized with a DNN later in Section IV.

estimates due to the delays caused by the large distances involved in satellite communications [29], [33], [48]. Therefore, to address this issue, inspired by [24], we propose a predictive transmit waveform and receive filter design procedure that eliminates the \mathbf{H}_n and \mathbf{G}_n estimation process. Therefore, the prediction of transmit waveform \mathbf{X}_n and receive filter \mathbf{W}_n is realized by utilizing historical estimated CSI Ω_{n-1}^τ and TRM Φ_{n-1}^τ , each of length τ .

In general, our proposed predictive transmit waveform and receive filter design procedure consists of initialization, signal transmission and processing, as well as predictive design phases. However, the initialization phase is only required during $n = 0$, where the other phases are required for each time frame⁷. To illustrate the proposed predictive transmit waveform and receive filter design procedure, we present Fig. 3. The phases of our proposed predictive ISAC procedure are described as follows:

1) *Phase 0–Initialization*: This phase only happens when $n = 0$. During this phase, a conventional channel estimator is used to obtain the initial CSI and TRM. Although nonlinearity in satellite channels can significantly affect channel estimation and detection, we assume that the adopted channel estimator effectively address these effects [50]. Furthermore, to gather more initial data points, the conventional channel estimation can be repeated consecutively during this phase. For the sake of simplicity, it does make sense to assume that the prior information (i.e., CSI and TRM) have been obtained in the detection phase [7], [24], [25].

2) *Phase 1–Signal transmission & processing*: After the initial transmission at $n = 0$, the transceiver transmits the waveform predicted during the phase 2 of the previous time frame ($n - 1$). When the echoes return to the transceiver, the receive filter, which also predicted during the phase 2 of the ($n - 1$)-th time frame, is used to process the incoming echoes. However, due to the large distances, propagation delays are unavoidable in both signal transmission and echo reflection.

⁷Due to the prediction drift [49] and the unavailability of initial historical data, it is necessary to periodically re-execute the initialization phase at defined time frame intervals. This re-execution is considered as one cycle of design procedure. However, to simplify implementation, this study considers only a single cycle of the design procedure.

3) *Phase 2–Predictive design*: This phase solves the design problem as described in (19). Given the historical estimated CSI and TRM, Ω_{n-1}^τ and Φ_{n-1}^τ , we introduce an intermediate function $f_P(\cdot)$ to implicitly replace the needs of current CSI \mathbf{H}_n and TRM \mathbf{G}_n . Thus, the solution to problem (19) can be obtained by

$$(\mathbf{X}_n, \mathbf{W}_n) = f_O(\mathbf{S}_n, f_P(\Omega_{n-1}^\tau, \Phi_{n-1}^\tau)). \quad (21)$$

Therefore, as an intermediate approach, optimizing \mathbf{X}_n and \mathbf{W}_n is equivalent to optimizing the chain of functions in (21). However, solving problem (19) is inherently challenging due to the intractability of deriving a closed-form solution, even for the intermediate approach. Furthermore, the objective function and constraints are non-convex with respect to \mathbf{X}_n and \mathbf{W}_n . As a realization, we will adopt a data-driven approach to represent the function $f_O(\mathbf{S}_n, f_P(\Omega_{n-1}^\tau, \Phi_{n-1}^\tau))$ using a DNN and learning the parameters from data, inspired by the learning-to-optimize approach [51].

Remark 6: Within each time frame, the transmit waveform \mathbf{X}_n and receive filter \mathbf{W}_n are jointly designed based on the given information symbol matrix \mathbf{S}_n . This approach optimizes communications by tailoring \mathbf{X}_n to the structure of the transmitted information symbols, which consequently influences the design of \mathbf{W}_n as well. For demonstration purposes in this work, we assume \mathbf{S}_n is a modulated symbol matrix derived from binary information, where the original bits (ones and zeros) follow a uniform distribution.

Additionally, Fig. 3 also compares our approach with traditional pilot-based [22] and two-stage procedures [23]. It is clearly demonstrated that by employing the predictive design phase, the entire process is expedited (shown by left arrows at the end of last time frame), ensuring the use of the most current information. As a result, the transmit waveform and filter design are as timely as possible.

Remark 7: The proposed predictive transmit waveform and receive filter design procedure can be applied to any design scenarios for S-ISAC framework that typically require a channel estimation process, as it effectively eliminates the need for channel estimation. In this work, however, we specifically

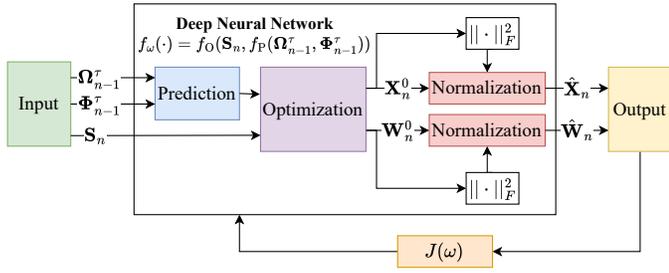


Fig. 4. The proposed DL-based framework to design transmit waveform and receive filter given historical CSI, TRM, and communications symbols.

leverage its advantage in the context of ISAC transmit waveform and receive filter design in satellite-based networks.

IV. DEEP LEARNING-BASED PREDICTIVE DESIGN FOR ISAC: TRANSMIT WAVEFORM AND RECEIVE FILTER

To implement the predictive joint transmit waveform and receive filter design phase, we aim to solve problem (19) with a function, as described in (21), and parameterizing this function using a DNN via a data-driven approach. First, let $f_\omega(\cdot)$ denote the network that parameterizes the chain of functions (21), where ω is the network parameters. Then, using a training dataset, we train the DNN to optimize its parameters by minimizing the loss function $J(\omega)$ in an unsupervised fashion. The proposed DL-based scheme to design transmit waveform and receive filter is illustrated in Fig. 4. In this section, we will present our proposed solution to realize the function $f_\omega(\cdot)$.

A. Proposed DL-Based Predictive Design Framework

Our proposed predictive DL-based design framework generally involves several key steps. First, we handle constraints by transforming the constrained problem into an unconstrained one. Next, we approximate the objective function using the Monte Carlo method. This approximation allows us to define the DNN training loss function. Finally, once the DNN is well-trained, we obtain a model that results in the optimized transmit waveform and receive filter. The details of each key steps are described as follows:

1) *Constraints handling*: It is important noting that DL-based approaches are typically designed for unconstrained optimization problems [27]. To adapt these methods for our constrained optimization problem, we employ the penalty method, which effectively transforms the constrained problem into an equivalent unconstrained one [52]. The transformed problem (19) can be written as follows

$$\begin{aligned} \max_{\mathbf{X}_n, \mathbf{W}_n} \mathbb{E}_{\mathbf{S}_n, \mathbf{H}_n, \mathbf{G}_n} U \{U_C(\mathbf{X}_n, \mathbf{S}_n, \mathbf{H}_n), U_S(\mathbf{X}_n, \mathbf{W}_n, \mathbf{G}_n)\} \\ \text{subject to } \Omega_{n-1}^\tau, \Phi_{n-1}^\tau \\ - \psi_1 \left[\max \left(0, \|\mathbf{X}_n\|_F^2 - P_{\max} \right) \right]^2 \\ - \psi_2 \left[\max \left(0, \|\mathbf{W}_n\|_F^2 - 1 \right) \right]^2, \end{aligned} \quad (22)$$

where $\psi_i \gg 0$, $i \in \{1, 2\}$ denotes the constant imposed for violation of the power constraints⁸. However, it is still

⁸Determining the value of ψ_i requires fine-tuning, which can be either fixed or dynamic [52]. In this work, we select a fixed ψ_i for ease of implementation.

intractable to derive a closed-form solution for the transformed problem.

Additionally, to ensure the satisfaction of the constraints in our particular problem, we deliberately normalize the output by the Frobenius norm of the original output, as inspired by [9]. Therefore, let say the original outputs of the DNN are \mathbf{X}_n^0 and \mathbf{W}_n^0 ; then the actual results will be

$$\hat{\mathbf{X}}_n = \frac{\mathbf{X}_n^0}{\|\mathbf{X}_n^0\|_F^2} P_{\max}, \quad \hat{\mathbf{W}}_n = \frac{\mathbf{W}_n^0}{\|\mathbf{W}_n^0\|_F^2}, \quad (23)$$

thus ensuring that the constraints are always satisfied. This process compensates for the large path losses by utilizing the maximum available power. However, this normalization becomes ineffective for non-power constraints.

Remark 8: After normalization, the penalty terms in (22) become ineffective, as they will result in zeros. However, to ensure the generalizability of the framework, we retain the penalty method as the first step to address constraints, particularly in cases where the constraints cannot be resolved through normalization alone.

2) *Data-driven approximation*: To address the intractability of the closed-form solution to problem (22), we adopt a data-driven approach to asymptotically approximate the statistical expectation involved in the objective function. By leveraging the powerful feature extraction capabilities of DNN, we can obtain the solution to the transformed problem. Firstly, let $\mathbb{E}\{g(\mathbf{X}_n, \mathbf{W}_n)\}$ represent the transformed objective function (22). Then, we utilize the Monte Carlo method to approximate it as follows

$$\begin{aligned} \mathbb{E}\{g(\mathbf{X}_n, \mathbf{W}_n)\} &\approx \frac{1}{N_s} \sum_{i=1}^{N_s} g(\mathbf{x}_n^{(i)}, \mathbf{w}_n^{(i)}), \quad (24) \\ &= \frac{1}{N_s} \sum_{i=1}^{N_s} g\left(f_\omega(\mathbf{S}_n^{(i)}, \Omega_{n-1}^{\tau(i)}, \Phi_{n-1}^{\tau(i)})\right), \quad (25) \end{aligned}$$

where N_s is the number of training samples [53]. Nonetheless, due to the universal approximation theory [51], the approximation becomes valid when N_s is sufficiently large [27].

3) *Training loss function*: Given the expected value approximation, the loss function for the DNN training can be written as follows

$$J(\omega) = -\frac{1}{N_s} \sum_{i=1}^{N_s} g\left(f_\omega(\mathbf{S}_n^{(i)}, \Omega_{n-1}^{\tau(i)}, \Phi_{n-1}^{\tau(i)})\right), \quad (26)$$

where the negative objective function is used to represent the original maximization problem as an equivalent minimization problem.

4) *Training and inference*: The optimized transmit waveform and receive filter can be acquired from the DNN training by optimizing ω that minimizes the loss function in an unsupervised manner. Therefore, the final solution can be represented as follows

$$(\hat{\mathbf{X}}_n, \hat{\mathbf{W}}_n) = f_{\hat{\omega}}(\mathbf{S}_n, \Omega_{n-1}^\tau, \Phi_{n-1}^\tau), \quad (27)$$

where $\hat{\mathbf{X}}_n$ and $\hat{\mathbf{W}}_n$ represent the optimized transmit waveform and receive filter, respectively, and $\hat{\omega} = \arg \min_{\omega} J(\omega)$ is

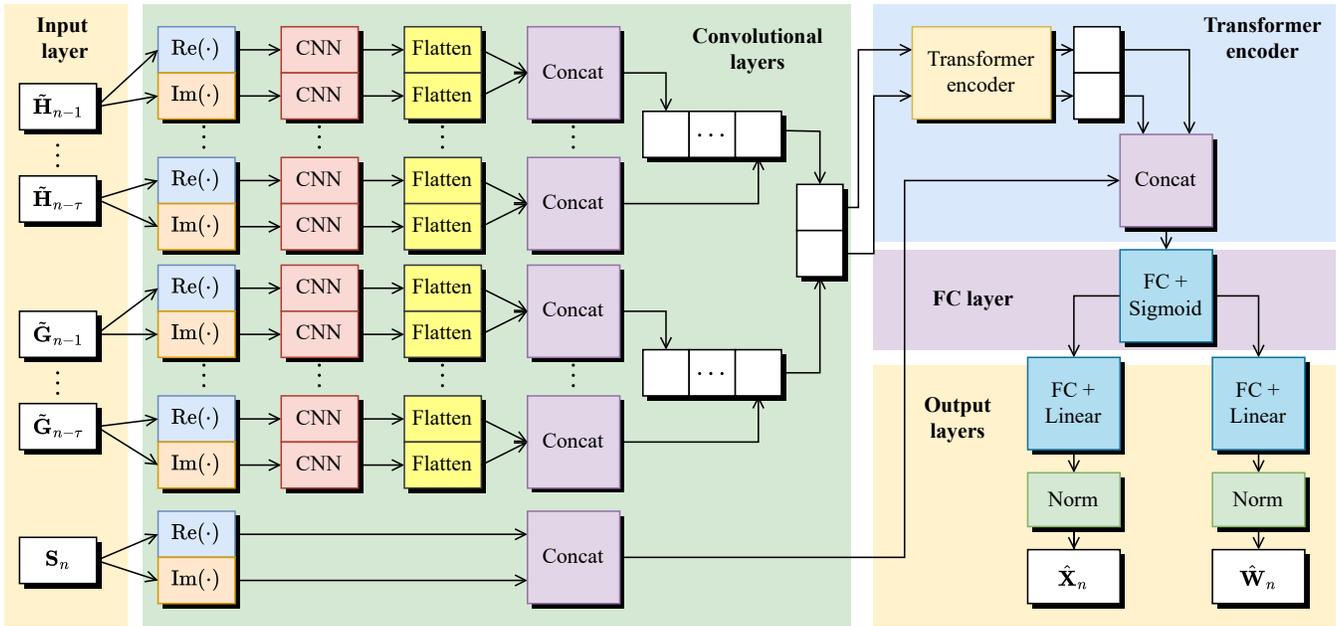


Fig. 5. Proposed PONet architecture to design predictive transmit waveform and receive filter in ISAC system. This network is a realization of the phase 2-predictive design within our proposed predictive transmit waveform and filter procedure (please refer to Fig. 3).

the well-trained network parameters. Once the model is well-trained, the optimized transmit waveform and receive filter can be acquired by feeding \mathbf{S}_n , $\mathbf{\Omega}_{n-1}^\tau$, and $\mathbf{\Phi}_{n-1}^\tau$ to the model without additional training required.

B. Predictive Optimization Network (PONet)

To realize the DNN, we propose PONet, specifically designed for predictive joint ISAC transmit waveform and receive filter design that jointly optimizes communications and sensing tasks in ISAC systems as shown in Fig. 5. This network combines CNNs with Transformer model to process complex-valued data, enabling the joint design of predictive transmit waveform and receive filter. The layers of PONet are summarized as follows:

1) *Input layer*: The input layer of the PONet accepts three complex-valued matrices: $\mathbf{\Omega}_{n-1}^\tau$, $\mathbf{\Phi}_{n-1}^\tau$, and \mathbf{S}_n . Each of these matrices contains both real and imaginary components, and they are processed separately in the subsequent layers. The temporal dimension τ is maintained throughout the network to enable temporal modeling.

2) *Convolutional layers*: The convolutional layers are designed to extract spatial features for each temporal slice independently. For each frame τ , a pair of 2D CNN are applied to process the real and imaginary components of $\mathbf{\Omega}_{n-1}^\tau$ and $\mathbf{\Phi}_{n-1}^\tau$. Therefore, the architecture includes $\tau \times 2 \times 2$ CNNs. Then, the outputs from the CNNs are flattened and transformed to match the input dimensions required by the Transformer encoder. This transformation is achieved by concatenating the outputs, with the first half representing the real component and the second half the imaginary component. The concatenated outputs are then organized in temporal sequence before being fed into the Transformer encoder.

3) *Transformer encoder*: The Transformer encoder consists of multiple layers of self-attention and feedforward networks, allowing it to capture long-term dependencies and interactions across the entire sequence of time frame indices [28]. This design enables the network to effectively model the temporal evolution of CSI and TRM, maintaining adaptability even over extended sequences. In our network, the Transformer encoder processes the sequence of features extracted by the convolutional layers. Before being passed to the next layer, the outputs of the Transformer encoder are concatenated with the processed information symbols.

4) *Fully-connected (FC) layer*: After processing the temporal sequence with Transformer, the network concatenates the Transformer's output with $[\text{Re}(\mathbf{S}_n), \text{Im}(\mathbf{S}_n)]$. This combined output is then fed to an FC layer with a Sigmoid activation function.

5) *Output layers*: The output layers comprise two branches that produce the intermediate complex-valued matrices $\hat{\mathbf{X}}_n^0$ and $\hat{\mathbf{W}}_n^0$. In each branch, the real and imaginary components are represented by even and odd indices, and the final outputs are reconstructed as complex-valued matrices. Subsequently, after applying normalization as described in (23), the network returns $\hat{\mathbf{X}}_n$ and $\hat{\mathbf{W}}_n$.

C. Overall Algorithm

The realization of our proposed framework involves two stages. The first stage is offline training, where the model learns from a set of training samples to optimize ω , which, in our case, correspond to PONet's parameters. The goal of this training is to minimize the loss function as described in (26). The second stage is online inference, where the transmit waveform and receive filter are designed by feeding the information symbol matrix \mathbf{S}_n , historical CSI $\mathbf{\Omega}_n^\tau$, and

historical TRM Φ_n^τ into $f_{\hat{\omega}}$. It is important to emphasize that the predictive design phase in our design procedure, as illustrated in Fig. 3, corresponds to the online inference stage. The complete algorithm is detailed in Algorithm 1.

Algorithm 1 Predictive Transmit Waveform and Receive Filter Design for S-ISAC

Stage 1: Offline training

Input: $\{(\mathbf{S}_n^{(1)}, \Omega_{n-1}^{\tau(1)}, \Phi_{n-1}^{\tau(1)}), \dots, (\mathbf{S}_n^{(N_s)}, \Omega_{n-1}^{\tau(N_s)}, \Phi_{n-1}^{\tau(N_s)})\}$
Output: $f_{\hat{\omega}}(\cdot)$

- 1: Initiate ω ;
- 2: **repeat**
- 3: Update ω using the backpropagation algorithm to minimize the loss function in (26);
- 4: **until** the maximum number of training epochs is reached or the loss converges
- 5: $\hat{\omega} \leftarrow \omega$; $f_{\hat{\omega}} \leftarrow f_{\omega}$;

Stage 2: Online inference

Input: $\mathbf{S}_n, \Omega_{n-1}^\tau, \Phi_{n-1}^\tau$
Output: $\hat{\mathbf{X}}_n, \hat{\mathbf{W}}_n$

- 6: Compute $(\hat{\mathbf{X}}_n, \hat{\mathbf{W}}_n) = f_{\hat{\omega}}(\mathbf{S}_n, \Omega_{n-1}^\tau, \Phi_{n-1}^\tau)$;

D. Complexity Analysis

The complexity of the proposed solution encompasses both the offline training and online inference phases, with each phase involving the computational requirements of the proposed neural network. First, the complexity of the CNNs is given by $\mathcal{C}_{\text{CNN}} = \mathcal{O}(\tau \mathcal{L}_h \kappa_H \kappa_W (KN_T + N_R N_T))$, where \mathcal{L}_h represents the hidden size, and κ_H and κ_W denote the height and width of the kernel size, respectively. Next, the complexity of dimension matching after the CNN layer can be expressed as $\mathcal{C}_{\text{DM}} = \mathcal{O}(\mathcal{L}_h (KN_T + N_R N_T) 2\eta)$, where η represents the dimensionality of the input and output vectors in the Transformer model.

For the Transformer encoder, the complexity is written as $\mathcal{C}_{\text{Trans}} = \mathcal{O}(\mathcal{L}_E \tau \eta^2)$, where \mathcal{L}_E represents the number of encoder layers. Leveraging the self-attention mechanism, the Transformer achieves reduced per-layer computational complexity and enables more extensive parallel processing compared to traditional recurrent layers [28]. The final design process includes FC and output layers, where the complexity of the concatenation process and the fully connected (FC) layer can be expressed as $\mathcal{C}_{\text{FC}} = \mathcal{O}(\eta(\eta + 2KL))$. For the output layers, there are two branches of FC layers, each representing a specific output, with complexities given by $\mathcal{C}_X = \mathcal{O}(2\eta N_T L)$ and $\mathcal{C}_W = \mathcal{O}(2\eta N_R^2)$.

However, from the above components, the complexity of the Transformer encoder dominates. Therefore, the complexity for the offline training and online inference phases can be approximated as $\mathcal{C}_{\text{offline}} \approx \mathcal{O}(I_{\text{max}} N_s \mathcal{L}_E \tau \eta^2)$ and $\mathcal{C}_{\text{online}} \approx \mathcal{O}(\mathcal{L}_E \tau \eta^2)$, respectively, where I_{max} denotes the maximum iteration number.

V. NUMERICAL RESULTS AND DISCUSSION

In this section, we conduct several simulations to demonstrate and evaluate our proposed framework. Unless otherwise

TABLE I
DEFAULT SIMULATION PARAMETERS

Parameters	Default Values
Carrier frequency	$f_c = 11$ GHz
Number of active GIDs	$K = 3$ GIDs
Number of sensing targets	$Q = 3$ targets
Signal length	$L = 10$ symbols
Satellite's position	$(x_{n,s}, y_{n,s}, z_{n,s}) = (0, 0, 400)$ km
Satellite's transmit antenna size	$N_T^x = N_T^y = 3$ antennas
Satellite's receive antenna size	$N_R^x = N_R^y = 3$ antennas
Satellite's antenna separation	$\delta_x = \delta_y = 0.5\lambda_c$
Satellite's antennas gain	$G_T = G_R = 6$ dBi
Satellite's max. transmission power	$P_{\text{max}} = 40$ dBm
Noise power	$\sigma_N^2 = -100$ dBm
Number of historical frame	$\tau = 3$
Duration of signal length	$\Delta_L = 3$ ms

specified, the default parameters for the simulation is given in Table I. In our system model, we consider the presence of communications users and sensing targets.

- *Communications users:* For ease of implementation, we deliberately set $K_{\text{total}} = 3$ GIDs and consider $\xi_K = 1$, which means all GIDs are always active. The GIDs are distributed across an area represented by $x_{n,k} = x_k \sim \mathcal{U}(-10 \text{ km}, 10 \text{ km})$, $y_{n,k} = y_k \sim \mathcal{U}(-10 \text{ km}, 10 \text{ km})$, and $z_{n,k} = 0$ m. Since the satellite and GIDs remain at static positions⁹, the LOS component is effectively invariant from one time frame to the next, resulting in a correlated CSI across adjacent time frames. Additionally, all users are being treated equally, therefore, $\beta_k = 1$. For symbol mapping, we use quadrature phase-shift keying (QPSK) modulation for demonstration.
- *Sensing targets:* As we have mentioned in Remark 2, we consider the presence of Q sensing targets and we also consider their mobility¹⁰. In this demonstration, we consider a static altitude for each target. The trajectory of q -th target can be modeled as follows

$$x_{n+1,q} = x_{n,q} + \Delta_L v_{n,q} \cos \Theta_{n,q}, \quad (28a)$$

$$y_{n+1,q} = y_{n,q} + \Delta_L v_{n,q} \sin \Theta_{n,q}, \quad (28b)$$

where $x_{n,q}$ and $y_{n,q}$ are the x - and y -coordinates, and $v_{n,q}$ is the resultant velocity of the q -th target at the n -th time frame. The orientation of the q -th target at the n -th frame is generated randomly each sample realization following $\Theta_{n,q} \sim \mathcal{U}(0, 2\pi)$. For the ease of implementation, we deliberately set the altitude and velocity of sensing targets as $z_{n,q} = 50$ m and $v_{n,q} = 1$ m/s, respectively. The initial position of the targets are $x_{0,q} \sim \mathcal{U}(-10 \text{ km}, 10 \text{ km})$ and $y_{0,q} \sim \mathcal{U}(-10 \text{ km}, 10 \text{ km})$. Regarding the reflection coefficient of the target, we intentionally assume a unit radar cross section $\varsigma_q = 1$ without loss of generality, meaning the reflection coefficient is influenced solely by the round-trip distance [54].

⁹For short time durations, we assume a static satellite position to simplify the analysis [18], [37], as small satellite movements during such intervals have a negligible effect on the network geometry.

¹⁰In this demonstration, we consider several UAVs as sensing targets.

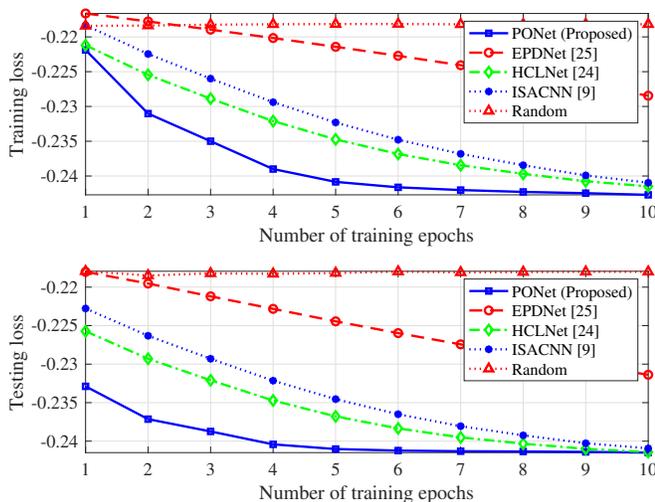


Fig. 6. Training and testing losses of various comparison schemes.

As for the DL training, we use 2000 data samples¹¹ with a split factor of 0.2. To optimize the network parameters, we use the Adam optimizer [55] with an initial learning rate of 10^{-3} . The learning rate is adjusted adaptively using the ReduceLRonPlateau method, with a reduction factor of 0.5 and a patience value of 10. The offline training and online inference are conducted on an AMD EPYC 7313 CPU and an NVIDIA GeForce RTX 3090 GPU.

To verify the performance of the proposed PONet, we compare it against the following baselines as their framework can be easily adopted to our problem¹²:

- EPDNet [25]: This approach employs a CNN to capture spatial features, followed by an LSTM module to extract temporal features. The outputs are then passed through fusion and decoder modules to generate the final results.
- HCLNet [24]: This method uses a CNN to extract spatial features, which are then processed by an LSTM module for each historical time frame. The output is subsequently combined using FC and linear layers.
- ISACNN [9]: This approach applies a series of CNNs and a multi-layer perceptron (MLP) to map the input to the output, using a Sigmoid activation function to produce the final output.
- Random: This method generates random transmit waveforms and receive filters, normalizing them to adhere to power budget constraints.

To begin with, we compare the performance of the proposed PONet with the baseline models in terms of convergence behavior during the offline training and online testing stages, as illustrated in Fig. 6. Among all the comparison schemes, PONet demonstrates the fastest convergence, reaching it by

¹¹The data is generated based on the system model and is subsequently divided into training and testing datasets [9], [10], [24], [25], [41].

¹²As the original objectives are different, we made slight modifications and retrained all the comparison schemes to ensure they align with our input data, i.e., historical CSI and TRM, and the desired output, i.e., the ISAC transmit waveform and receive filter.

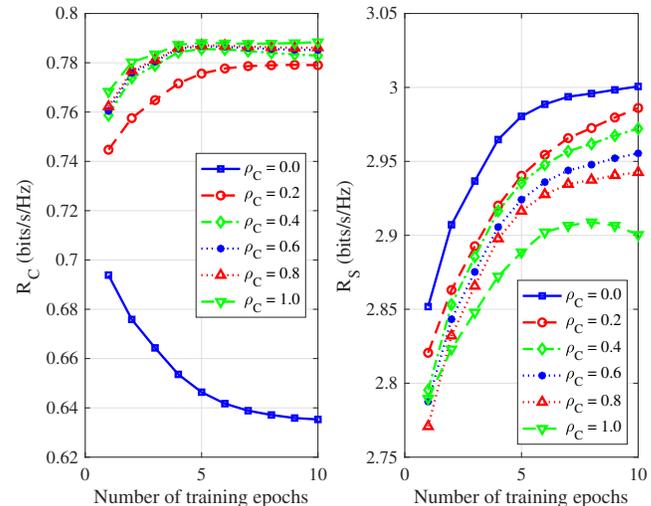


Fig. 7. Testing performance of the proposed scheme in terms of the communications and sensing rates under different communications weights.

the fifth training epoch, although HCLNet and ISACNN eventually converge to a similar minimum point as PONet later. This result emphasizes that our proposed PONet requires only a few training epochs to achieve convergence. Additionally, it can be observed that the offline training and online testing losses exhibit approximately similar values and convergence behavior, indicating that our model maintains consistency between the offline training and online testing stages. While all comparison schemes decrease to a certain minimum point, the random scheme remains constant. This is because no optimization takes place in the random scheme, as it generates a random transmit waveform and receive filter in each iteration. Particularly, for PONet, we present Fig. 7 that illustrates its effectiveness in achieving testing convergence across various communications weights. The results show that the scheme quickly converges for all communications weights, with both communications rate and sensing rate convergence within just a few training epochs. It is important to note that in the sensing-focused case, when $\rho_C = 0.0$, the communications rate decreases because the communications utility is excluded from the objective function. Conversely, in the communications-focused case, when $\rho_C = 1.0$, the sensing rate initially increases but then declines, as the sensing utility is no longer included in the objective function. The proposed PONet demonstrates robust performance across all weights, proving its adaptability in different S-ISAC scenarios. This highlights the efficiency of PONet in optimizing both communications and sensing tasks, making it well-suited for practical applications. In practice, the training is terminated once the model converges. This early convergence not only demonstrates the efficiency of our model but also indicates its ability to quickly optimize the objective function, reducing overall training time compared to baselines.

After PONet is well-trained, we present more results to assess the testing performance. In this testing, we intentionally exclude the performance of the network utility function as

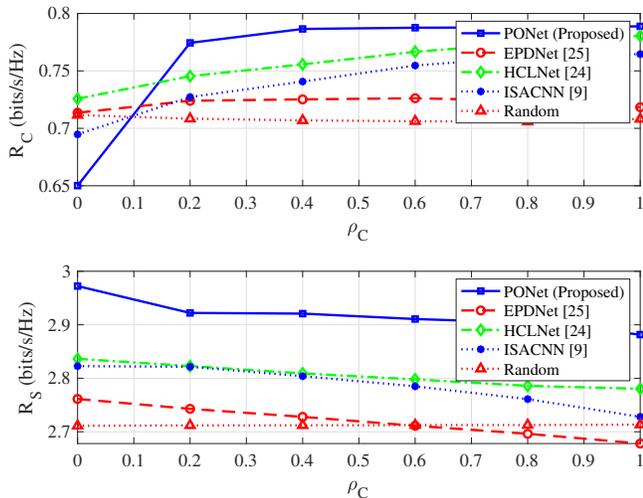


Fig. 8. Communications and sensing rates with different communications weights of various comparison schemes.

described in (20) since it lacks direct physical interpretation. Instead, we present the actual performance in terms of communications and sensing rates.

First, we present Fig. 8 to further evaluate the performance of communications and sensing rates under different communications weights. Across all comparison schemes, it is evident that as the communications weight increases, the communications rate also rises while the sensing rate decreases, except for the random scheme as it is constant for all communications weights. Therefore, it demonstrates that our proposed PONet exhibits the same trade-off behavior as the existing baselines in terms of the performance between communications and sensing. Moreover, Fig. 8 also demonstrates that our proposed PONet achieves the best sensing rate across all communications weights. It is worth noting that the communications rate is more sensitive to changes in the communications weight than the sensing rate, making the selection of an appropriate communications weight crucial for optimizing overall S-ISAC performance.

To evaluate the impact of the number of historical time frames τ on both communications and sensing rates, we present Fig. 9. Specifically for PONet, we observe that using more historical time frames leads to a slight increase in communications rate, while the sensing rate remains constant¹³. Additionally, PONet consistently achieves the highest communications and sensing rates across all comparison schemes. Although other benchmarks do not exhibit a clear trend, PONet proves to be the most effective in capturing temporal dependencies within the satellite communications network. While HCLNet [24] and EPDNet [25] have demonstrated success in predictive tasks for ISAC-assisted ground-based vehicular networks, they fall short in transmit waveform and receive

¹³Although increasing τ improves communications performance, using a large number of τ is not advisable, as it needs more input data and imposes additional burden on the CPU when designing the transmit waveform and receive filter.

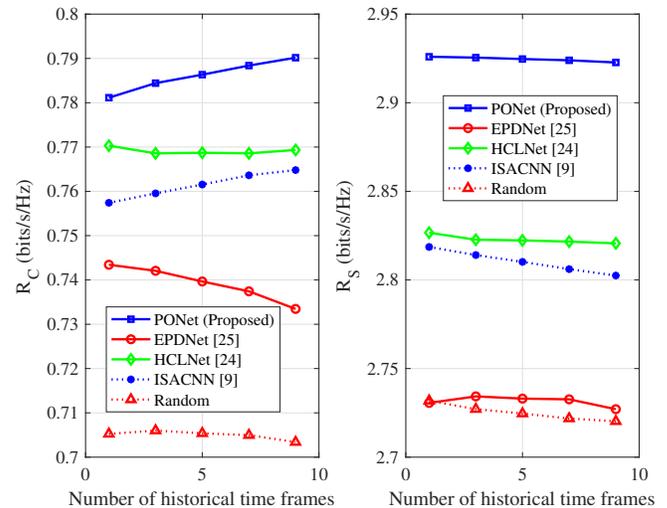


Fig. 9. Effect of historical time frames number of various comparison schemes.

filter design for S-ISAC systems. These findings confirm that our proposed PONet performs well in the predictive design of transceivers for the considered satellite communications scenario.

Next, we present Fig. 10 to illustrate the impact of the number of data samples on performance. We deliberately compare the performance with total data samples of sizes 10, 100, 1000, 2000, and 5000¹⁴. For PONet, it is clearly shown that as the number of data samples increases, the performance improves accordingly. Notably, we observe that 2000 data samples are sufficient to achieve a well-optimized design. In contrast, for smaller sample sizes, some benchmarks exhibit inconsistent slopes for communications and sensing rates, along with instability due to overfitting. However, for our proposed PONet, the relationship between data samples and performance is linear and smooth. This experiment demonstrates that PONet performs well with just 2000 data samples, maintaining a consistent slope for both communications and sensing performances, thereby underscoring the superiority of PONet compared to other baseline schemes in terms of the S-ISAC systems.

Regarding the design parameter, we examine the impact of signal length on the performance of the S-ISAC system, as it is a crucial parameter for transmit waveform design in both communications and sensing tasks. Fig. 11 illustrates that across all comparison schemes, increasing the signal length leads to an increase in the communications rate. However, the sensing rate decreases as the signal length grows. Our proposed PONet consistently delivers the best performance for signal lengths of six or more ($L \geq 6$). In practical terms, a signal length of less than six ($L < 6$) is inadequate for accommodating more data in communications. Additionally, we observe that the performance gap between PONet and the second-best scheme widens as the signal length increases,

¹⁴With 80% of the data utilized for training, the size of training samples are as follows: 8, 80, 800, 1600, and 4000.

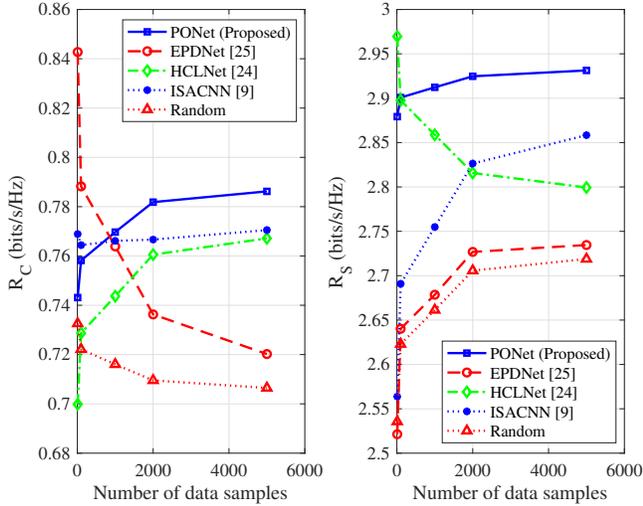


Fig. 10. Number of data samples vs. the communications and sensing rates of various comparison schemes.

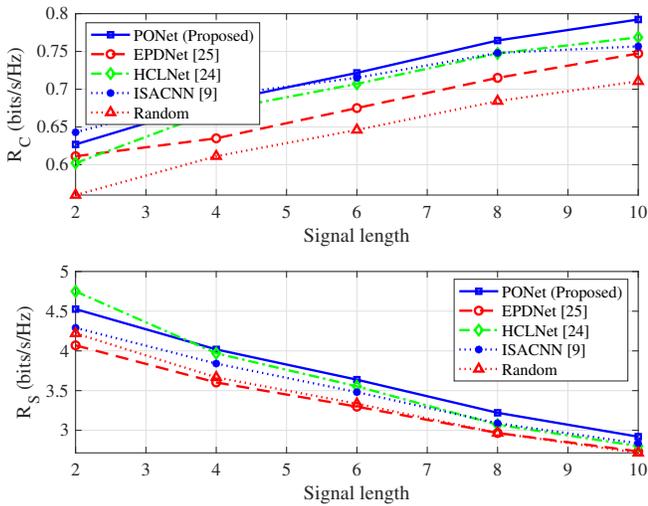


Fig. 11. Signal length vs. the communications and sensing rates of various comparison schemes.

further highlighting the superiority of the proposed PONet in communications performance. It is also worth noting that the relationship with sensing rate is inverse; longer signal lengths result in lower sensing rates, which aligns with the expected behavior indicated by (18). Therefore, it is crucial to choose a signal length L that optimizes performance for one task without significantly compromising the other.

Finally, we present Fig. 12 to illustrate that our proposed scheme consistently adheres to the power budget constraints in both transmit waveform and receive filter design. The figure clearly shows that across various simulation settings, represented by different communications weights, the power of the transmit waveform remains bounded by 40 dBm, and the power of the receive filter stays within 30 dBm (unit power). This demonstrates that our proposed scheme is ap-

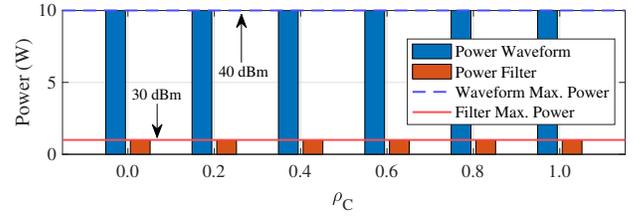


Fig. 12. Required power for transmit waveform and receive filter across different weighting scenarios.

pliable across various scenarios. Furthermore, this highlights the efficacy of the normalization approach outlined in (23) for rigorously enforcing the power budget constraints.

VI. CONCLUSION

In this paper, we investigated ISAC in satellite communications networks and develop a predictive S-ISAC transceiver design framework. The design objective is to maximize network utility, by jointly optimizing ISAC transmit waveform and receive filter design, with the objective realization being a weighted sum of communications and sensing normalized rates maximization. Firstly, we proposed a predictive ISAC procedure that eliminates the channel estimation overhead, ensuring that the information remains as current as possible. Then, given the formulated problem is intractable for conventional optimization solvers, we adopted a DL-based data-driven framework as a solution, specifically we propose PONet to solve the design problem. The effectiveness of the proposed solution in terms of communications and sensing rates were validated through numerical simulations, which compared its convergence behavior, trade-off between communications and sensing, and the impact of historical time frames, data samples, and signal length. Last but not least, we demonstrated that the power budget constraints are satisfied.

APPENDIX A DERIVATION OF MI (17)

To determine MI $I(\mathbf{Z}_n; \mathbf{G}_n | \mathbf{X}_n)$ as outlined in (16), we need to calculate two entropies: the entropy of \mathbf{Z}_n given \mathbf{X}_n , and the entropy of \mathbf{Z}_n given both \mathbf{G}_n and \mathbf{X}_n . Firstly, when \mathbf{X}_n is known, \mathbf{Z}_n is expressed as follows

$$\mathbf{Z}_n | \mathbf{X}_n = \mathbf{W}_n^H \mathbf{G}_n \mathbf{X}_n + \mathbf{W}_n^H \mathbf{M}_n. \quad (29)$$

The covariance matrix of \mathbf{Y}_n given \mathbf{X}_n as follows

$$\Sigma_{\mathbf{Z}_n | \mathbf{X}_n} = \mathbf{W}_n^H (\mathbf{G}_n \mathbf{X}_n \mathbf{X}_n^H \mathbf{G}_n^H + \sigma_N^2 \mathbf{I}) \mathbf{W}_n. \quad (30)$$

The differential entropy can thus be written as follows

$$h(\mathbf{Z}_n | \mathbf{X}_n) = \log_2 \det (\pi e \Sigma_{\mathbf{Z}_n | \mathbf{X}_n}). \quad (31)$$

Secondly, when both \mathbf{G}_n and \mathbf{X}_n are known, \mathbf{Z}_n is given by

$$\mathbf{Z}_n | \mathbf{G}_n, \mathbf{X}_n = \mathbf{W}_n^H \mathbf{M}_n. \quad (32)$$

The covariance matrix of \mathbf{Z}_n given \mathbf{G}_n and \mathbf{X}_n can be written as follows

$$\Sigma_{\mathbf{Z}_n | \mathbf{G}_n, \mathbf{X}_n} = \sigma_N^2 \mathbf{W}_n^H \mathbf{W}_n. \quad (33)$$

Thus, the differential entropy can be expressed as follows

$$h(\mathbf{Z}_n | \mathbf{G}_n, \mathbf{X}_n) = \log_2 \det(\pi e \sigma_N^2 \mathbf{W}_n^H \mathbf{W}_n). \quad (34)$$

Finally, by substituting (31) and (34) into (16), we can write MI as follows

$$I(\mathbf{Z}_n; \mathbf{G}_n | \mathbf{X}_n) = \log_2 \det(\Sigma_{\mathbf{Z}_n | \mathbf{X}_n}) - \log_2 \det(\sigma_N^2 \mathbf{W}_n^H \mathbf{W}_n), \quad (35)$$

where the πe term is canceled out due to the properties of the logarithm. Therefore, using another logarithmic property $\log a - \log b = \log \frac{a}{b}$, (35) can be re-expressed as (17). This concludes the derivation.

APPENDIX B

THEORETICAL MAXIMUM COMMUNICATIONS AND SENSING RATES FOR (20)

The maximum communications rate $\mu_{C,n}$ can be intuitively obtained under the assumption of ideal conditions, where MUI is zero. Therefore, (9) can be simplified to $\gamma_{n,k} = \mathbb{E}\{|s_{n,k}|^2\} / \sigma_N^2$. Since we assume the same constellation for every GIDs, we can write $\mathbb{E}\{|s_{n,k}|^2\} = P_{n,k}$. Therefore, we can substitute into (10) as follows

$$\mu_{C,n} = \frac{1}{K} \sum_{k \in \mathcal{K}_n} \beta_k \log_2 \left(1 + \frac{P_{n,k}}{\sigma_N^2} \right), \quad (36)$$

where $P_{n,k}$ is the k -th GID's received power in the n -th time frame.

To determine the theoretical upper bound of the sensing rate $\mu_{S,n}$, we begin by exploiting eigenvalue decomposition on $\mathbf{G}_n \mathbf{X}_n \mathbf{X}_n^H \mathbf{G}_n^H$. Let λ_i be the i -th eigenvalue, we can rewrite \mathbf{R}_{EE} as utilized in (17) as follows

$$\mathbf{R}_{EE} = \mathbf{U} (\mathbf{\Lambda} + \sigma_N^2 \mathbf{I}) \mathbf{U}^H, \quad (37)$$

where $\mathbf{\Lambda}$ is a diagonal matrix with $\lambda_i, i \in \{1, \dots, N_R\}$ become the diagonal elements, and \mathbf{U} is a unitary matrix. Since the determinant of \mathbf{R}_{EE} can be expressed as $\det(\mathbf{R}_{EE}) = \prod_{i=1}^{N_R} (\lambda_i + \sigma_N^2)$, thus, the numerator inside the logarithmic term of (17) can be written as follows

$$\det(\mathbf{W}_n^H \mathbf{R}_{EE} \mathbf{W}_n) = \det(\mathbf{W}_n^H \mathbf{U} (\mathbf{\Lambda} + \sigma_N^2 \mathbf{I}) \mathbf{U}^H \mathbf{W}_n). \quad (38)$$

By utilizing a property of determinant that $\det(\mathbf{ABC}) = \det(\mathbf{A}) \det(\mathbf{B}) \det(\mathbf{C})$ and the fact that \mathbf{U} is unitary ($\mathbf{U}^H \mathbf{U} = \mathbf{I}$), we can further expand (38) as $\det(\mathbf{W}_n^H \mathbf{R}_{EE} \mathbf{W}_n) = \det(\mathbf{W}_n^H \mathbf{U}) \det(\mathbf{\Lambda} + \sigma_N^2 \mathbf{I}) \det(\mathbf{U}^H \mathbf{W}_n)$.

For the denominator inside the logarithmic term of (17), the determinant of $\mathbf{W}_n^H \mathbf{W}_n$ can be expressed as $\det(\mathbf{W}_n^H \mathbf{W}_n) = \prod_{i=1}^{N_R} \mu_i$, where μ_i is the i -th eigenvalue of $\mathbf{W}_n^H \mathbf{W}_n$. Therefore, it is obvious that to maximize MI, we need to maximize the determinant ratio.

Given that $\det(\mathbf{W}_n^H \mathbf{U}) = \det(\mathbf{U}^H \mathbf{W}_n)$ and using the properties of determinants, the expression can be simplified as follows

$$I(\mathbf{Z}_n; \mathbf{G}_n | \mathbf{X}_n) = \log_2 \left(\prod_{i=1}^{N_R} \left(\frac{\lambda_i + \sigma_N^2}{\sigma_N^2 \mu_i} \right) \right). \quad (39)$$

The eigenvalues λ_i are maximized when the power constraint is fully utilized. Since $\|\mathbf{G}_n \mathbf{X}_n \mathbf{X}_n^H \mathbf{G}_n^H\|_F^2 = \sum_{i=1}^{N_R} \lambda_i \leq$

$\lambda_{\max} N_R$, the upper bound of MI occurs when $\lambda_i = \lambda_{\max}, \forall i \in \{1, \dots, N_R\}$, where λ_{\max} is the maximum eigenvalue of $\mathbf{G}_n \mathbf{G}_n^H$. For the normalization of \mathbf{W}_n , the eigenvalues μ_i of $\mathbf{W}_n^H \mathbf{W}_n$ must satisfy $\|\mathbf{W}_n\|_F^2 \leq 1$, ensuring that $\sum_{i=1}^{N_R} \mu_i \leq N_R$. The term $\frac{1}{\mu_i}$ is maximized when $\mu_i = 1, \forall i$, making $\prod_{i=1}^{N_R} \mu_i = 1$. Therefore, the upper bound for MI can be simplified as $I(\mathbf{Z}_n; \mathbf{G}_n | \mathbf{X}_n) = \log_2 \left(\prod_{i=1}^{N_R} \left(1 + \frac{\lambda_i}{\sigma_N^2} \right) \right)$, with $\lambda_i = \lambda_{\max}, \forall i$. Finally, the theoretical maximum sensing rate can be written as follows

$$\mu_{S,n} = \frac{N_R}{L} \log_2 \left(1 + \frac{\lambda_{\max}}{\sigma_N^2} \right). \quad (40)$$

This completes the derivation.

REFERENCES

- [1] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.
- [2] 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.
- [3] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, and K. Chetty, "A survey on fundamental limits of integrated sensing and communication," *IEEE Commun. Surv. Tutor.*, vol. 24, no. 2, pp. 994–1034, 2022.
- [4] W. Zhou, R. Zhang, G. Chen, and W. Wu, "Integrated sensing and communication waveform design: A survey," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1930–1949, 2022.
- [5] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4264–4279, Aug. 2018.
- [6] C. G. Tsinos, A. Arora, S. Chatzinotas, and B. Ottersten, "Joint transmit waveform and receive filter design for dual-function radar-communication systems," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1378–1392, Nov. 2021.
- [7] M. Shi, X. Li, J. Liu, and S. Lv, "Constant modulus waveform design for RIS-aided ISAC system," *IEEE Trans. Veh. Technol.*, vol. 73, no. 6, pp. 8648–8659, Jun. 2024.
- [8] P. Pulkkinen and V. Koivunen, "Model-based online learning for active ISAC waveform optimization," *IEEE J. Sel. Topics Signal Process. (Early Access)*, pp. 1–15, Apr. 2024.
- [9] Q. Qi, X. Chen, C. Zhong, C. Yuen, and Z. Zhang, "Deep learning-based design of uplink integrated sensing and communication," *IEEE Trans. Wirel. Commun. (Early Access)*, pp. 1–14, Mar. 2024.
- [10] 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. (Early Access)*, pp. 1–16, Mar. 2024.
- [11] Y. Wu, F. Lemic, C. Han, and Z. Chen, "Sensing integrated DFT-spread OFDM waveform and deep learning-powered receiver design for terahertz integrated sensing and communication systems," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 595–610, Jan. 2023.
- [12] C. Wen, Y. Huang, and T. N. Davidson, "Efficient transceiver design for MIMO dual-function radar-communication systems," *IEEE Trans. Signal Process.*, vol. 71, pp. 1786–1801, 2023.
- [13] Z. Wei, J. Piao, X. Yuan, H. Wu, J. A. Zhang, Z. Feng, L. Wang, and P. Zhang, "Waveform design for MIMO-OFDM integrated sensing and communication system: An information theoretical approach," *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 496–509, Jan. 2024.
- [14] L. Yin, Z. Liu, M. R. Bhavani Shankar, M. Alae-Kerahroodi, and B. Clerckx, "Integrated sensing and communications enabled low earth orbit satellite systems," *IEEE Netw. (Early Access)*, pp. 1–7, 2024.
- [15] Y. Zhang, W. Ni, J. Wang, W. Tang, M. Jia, Y. C. Eldar, and D. Niyato, "Robust transceiver design for covert integrated sensing and communications with imperfect CSI," *IEEE Trans. Commun. (Early Access)*, pp. 1–15, Apr. 2024.
- [16] B. He, F. Wang, and J. Cheng, "Joint secure transceiver design for integrated sensing and communication," *IEEE Trans. Wirel. Commun. (Early Access)*, pp. 1–17, May 2024.

- [17] Q. Zhu, M. Li, R. Liu, and Q. Liu, "Joint transceiver beamforming and reflecting design for active RIS-aided ISAC systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 7, pp. 9636–9640, Jul. 2023.
- [18] L. You, X. Qiang, C. G. Tsinos, F. Liu, W. Wang, X. Gao, and B. Ottersten, "Beam squint-aware integrated sensing and communications for hybrid massive MIMO LEO satellite systems," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2994–3009, Oct. 2022.
- [19] 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.
- [20] M. Huang, F. Gong, G. Li, N. Zhang, and Q.-V. Pham, "Secure precoding for satellite NOMA-aided integrated sensing and communication," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 29533–29545, Sep. 2024.
- [21] Z. Liu, L. Yin, W. Shin, and B. Clerckx, "Rate-splitting multiple access for quantized ISAC LEO satellite systems: A max-min fair energy-efficient beam design," *IEEE Trans. Wirel. Commun. (Early Access)*, pp. 1–15, 2024.
- [22] D. Zhang, A. Li, M. Shirvanimoghaddam, P. Cheng, Y. Li, and B. Vucetic, "Codebook-based training beam sequence design for millimeter-wave tracking systems," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 11, pp. 5333–5349, Nov. 2019.
- [23] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 11, pp. 7704–7719, Nov. 2020.
- [24] C. Liu, W. Yuan, S. Li, X. Liu, H. Li, D. W. K. Ng, and Y. Li, "Learning-based predictive beamforming for integrated sensing and communication in vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2317–2334, Aug. 2022.
- [25] X. Zhang, W. Yuan, C. Liu, J. Wu, and D. W. K. Ng, "Predictive beamforming for vehicles with complex behaviors in ISAC systems: A deep learning approach," *IEEE J. Sel. Topics Signal Process. (Early Access)*, pp. 1–14, May 2024.
- [26] N. Wu, R. Jiang, X. Wang, L. Yang, K. Zhang, W. Yi, and A. Nallanathan, "AI-enhanced integrated sensing and communications: Advancements, challenges, and prospects," *IEEE Commun. Mag.*, vol. 62, no. 9, pp. 144–150, 2024.
- [27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *NeurIPS 2017*, Long Beach, CA, USA, Jun. 2017.
- [29] X. Ding, Y. Ren, X. Xie, Y. Zou, M. Jia, and G. Zhang, "Improving user capacity of satellite internet of things via joint user grouping and multi-beam processing," *IEEE Trans. Commun. (Early Access)*, pp. 1–15, Feb. 2024.
- [30] E. Grayver, R. Keating, and A. Parower, "Feasibility of full duplex communications for LEO satellite," in *2015 IEEE Aerosp. Conf. Proc.*, Big Sky, MT, USA, Mar. 2015, pp. 1–8.
- [31] Q. T. Ngo, K. T. Phan, W. Xiang, A. Mahmood, and J. Slay, "Two-tier cache-aided full-duplex hybrid satellite-terrestrial communication networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 3, pp. 1753–1765, Jun. 2022.
- [32] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, wide-area, and non-terrestrial IoT: a survey on 5G advances and the road toward 6G," *IEEE Commun. Surv. Tutor.*, vol. 24, no. 2, pp. 1117–1174, 2022.
- [33] M. Centenaro, C. E. Costa, F. Granelli, C. Sacchi, and L. Vangelista, "A survey on technologies, standards and open challenges in satellite IoT," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 3, pp. 1693–1720, Jul. 2021.
- [34] W. D. Lukito, W. Xiang, P. Lai, P. Cheng, C. Liu, K. Yu, and X. Zhu, "Integrated STAR-RIS and UAV for satellite IoT communications: an energy-efficient approach," *IEEE Internet Things J. (Early Access)*, pp. 1–15, Oct. 2024.
- [35] Y. Zuo, M. Yue, M. Zhang, S. Li, S. Ni, and X. Yuan, "OFDM-based massive connectivity for LEO satellite internet of things," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 11, pp. 8244–8258, Nov. 2023.
- [36] K. Ying, Z. Gao, S. Chen, M. Zhou, D. Zheng, S. Chatzinotas, B. Ottersten, and H. V. Poor, "Quasi-synchronous random access for massive MIMO-based LEO satellite constellations," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1702–1722, Jun. 2023.
- [37] L. You, X. Qiang, Y. Zhu, F. Jiang, C. G. Tsinos, W. Wang, H. Wymeersch, X. Gao, and B. Ottersten, "Integrated communications and localization for massive MIMO LEO satellite systems," *IEEE Trans. Wirel. Commun.*, pp. 1–15, 2024.
- [38] D. Spano, S. Chatzinotas, J. Krause, and B. Ottersten, "Symbol-level precoding with per-antenna power constraints for the multi-beam satellite downlink," in *2016 ASMS/SPSC*, Palma, Spain, Sep. 2016, pp. 1–8.
- [39] X. Zhang, J. Wang, C. Jiang, C. Yan, Y. Ren, and L. Hanzo, "Robust beamforming for multibeam satellite communication in the face of phase perturbations," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3043–3047, Mar. 2019.
- [40] B. S. Mysore, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Precoding for satellite communications: Why, how and what next?" *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2453–2457, Aug. 2021.
- [41] M. Ying, X. Chen, Q. Qi, and W. Gerstacker, "Deep learning-based joint channel prediction and multibeam precoding for LEO satellite internet of things," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 10, pp. 13946–13960, 2024.
- [42] Y. Zhang, A. Liu, P. Li, and S. Jiang, "Deep learning (DL)-based channel prediction and hybrid beamforming for LEO satellite massive MIMO system," *IEEE Internet Things J.*, pp. 23705–23715, Dec. 2022.
- [43] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge University Press, 2005.
- [44] Q. Qi, X. Chen, C. Zhong, and Z. Zhang, "Integrated sensing, computation and communication in B5G cellular internet of things," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 1, pp. 332–344, Jan. 2021.
- [45] B. Tang and J. Li, "Spectrally constrained MIMO radar waveform design based on mutual information," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 821–834, Feb. 2019.
- [46] C. Ouyang, Y. Liu, and H. Yang, "On the performance of uplink ISAC systems," *IEEE Commun. Lett.*, vol. 26, no. 8, pp. 1769–1773, Aug. 2022.
- [47] 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.
- [48] J. A. Fraire, O. Iova, and F. Valois, "Space-terrestrial integrated internet of things: challenges and opportunities," *IEEE Commun. Mag.*, vol. 60, no. 12, pp. 64–70, Dec. 2022.
- [49] J. Gama, P. Medas, G. Castillo, and P. Rodrigues, "Learning with drift detection," in *SBLA 2004*, Sao Luis, Maranhao, Brazil, 2004, pp. 286–295.
- [50] Y. Zhang, B. Li, N. Wu, Y. Ma, W. Yuan, and L. Hanzo, "Message passing-aided joint data detection and estimation of nonlinear satellite channels," *IEEE Trans. Veh. Technol.*, vol. 72, no. 2, pp. 1763–1774, 2023.
- [51] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5438–5453, Oct. 2018.
- [52] A. E. Smith and D. W. Coit, "Constraint-Handling Techniques - Penalty Functions," in *Handbook of Evolutionary Computation*. Bristol, U.K.: Institute of Physics Publishing and Oxford University Press, 1997.
- [53] G. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*. New York, NY, USA: Springer, 2013.
- [54] Z. Xiao and Y. Zeng, "Waveform design and performance analysis for full-duplex integrated sensing and communication," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1823–1837, Jun. 2022.
- [55] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *ICLR 2015*, San Diego, CA, USA, May 2015.



William D. Lukito received B.Sc. and M.Sc. degrees in Telecommunications Engineering from Bandung Institute of Technology (ITB), Bandung, Indonesia, in 2021 and 2022, respectively. He is currently working towards the Ph.D. degree with La Trobe University, Australia under the Australian SmartSat CRC Scholarship. His research interests include wireless communications, Internet of Things, integrated sensing and communications, machine learning, and digital signal processing.



Wei Xiang (S'00–M'04–SM'10) is Cisco Research Chair of AI and IoT, Director & Chief Scientist of the Australian Centre for AI and Medical Innovation, and Director of the Cisco-La Trobe Centre for AI and IoT at La Trobe University. Previously, he was Foundation Chair and Head of Discipline of IoT Engineering at James Cook University, Cairns, Australia. For the past five consecutive years (2020–2024), has consistently ranked among Stanford University World's Top 2% Scientists for both his single-year and career-long impacts. Due to his

instrumental leadership in establishing Australia's first accredited Internet of Things Engineering degree program, he was inducted into Pearcey Foundations Hall of Fame in October 2018. He is a TEDx speaker and an elected Fellow of the IET in UK and Engineers Australia. He received the TNQ Innovation Award in 2016, and Pearcey Entrepreneurship Award in 2017, and Engineers Australia Cairns Engineer of the Year in 2017. He was a co-recipient of four Best Paper Awards at WiSATS2019, WCSP2015, IEEE WCNC2011, and ICWMC2009. He has been awarded several prestigious fellowship titles. He was named a Queensland International Fellow (2010–2011) by the Queensland Government of Australia, an Endeavour Research Fellow (2012–2013) by the Commonwealth Government of Australia, a Smart Futures Fellow (2012–2015) by the Queensland Government of Australia, and a JSPS Invitational Fellow jointly by the Australian Academy of Science and Japanese Society for Promotion of Science (2014–2015). He was the Vice Chair of the IEEE Northern Australia Section from 2016–2020. He is currently an Associate Editor for IEEE COMMUNICATIONS SURVEYS & TUTORIALS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE INTERNET OF THINGS JOURNAL, IEEE ACCESS, and Nature journal of Scientific Reports. He has published over 450 peer-reviewed papers including 3 books and nearly 300 journal articles. He has severed in a large number of international conferences in the capacity of General Co-Chair, TPC Co-Chair, Symposium Chair, etc. His research interest includes the Internet of Things, wireless communications, machine learning for IoT data analytics, and computer vision.

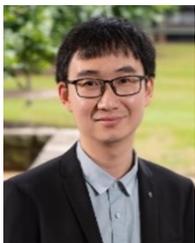


Phu Lai received his PhD degree in Information and Communications Technology from Swinburne University of Technology, Australia, in 2022. He is currently a postdoctoral research fellow at the Australian Centre for AI in Medical Innovation (ACAMI) and Cisco-La Trobe Centre for AI and IoT, where he has been working on several research and industry-sponsored projects in wireless communications, bioinformatics, smart cities, and so on.



Peng Cheng received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2006 and 2009, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2013. From 2014 to 2017, he was a Postdoctoral Research Scientist with CSIRO, Sydney, NSW, Australia. From 2017 to 2020, he was an ARC DECRA Fellow/Lecturer with the University of Sydney, Sydney. He is currently an ARC DECRA Fellow and a Senior Lecturer with the Department of Computer

Science and Information Technology, La Trobe University, Bundoora, VIC, Australia, and affiliated with the University of Sydney. He has authored or coauthored more than 100 research papers in leading international journals and conferences. His research interests include wireless AI, machine learning, IoT, mmWave communications, and compressive sensing.



Chang Liu (M'19) received the Ph.D. degree from Dalian University of Technology, China, in 2017. In 2015, he was a joint Ph.D. Scholar with the University of Tennessee, Knoxville, TN, USA. From 2017 to 2022, he was a Post-Doctoral Research Fellow with the University of Electronic Science and Technology of China, and a Research Fellow with the University of New South Wales, Sydney, Australia. Since 2022, he has been an Alexander von Humboldt Research Fellow with the Institute for Digital Communications, Friedrich Alexander University Erlangen-Nuremberg (FAU), Erlangen, Germany. From 2022 to 2023, he worked as an Assistant Professor (Research) at The Hong Kong Polytechnic University. He is currently a Lecturer (Assistant Professor in the U.S. system) with the Department of Computer Science and Information Technology, La Trobe University, Melbourne, Australia. To date, he has published more than 60 papers in leading international journals and conferences. He was a recipient of the World's Top 2% Scientists by Stanford University since 2022, the Best Paper Award at the IEEE International Symposium on Wireless Communication Systems in 2022, and an Exemplary Reviewer from IEEE TRANSACTIONS ON COMMUNICATIONS. He is a foundation member of IEEE ComSoc special interest group on orthogonal time frequency space (OTFS), an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS, Electronics, Future Internet, and Frontiers in Communications and Networks. His research interests include machine learning in communications, with a particular focus on integrated sensing and communication (ISAC), intelligent reflecting surface (IRS)-assisted communications, OTFS, the Internet of Things (IoT), and cognitive radio.

university Erlangen-Nuremberg (FAU), Erlangen, Germany. From 2022 to 2023, he worked as an Assistant Professor (Research) at The Hong Kong Polytechnic University. He is currently a Lecturer (Assistant Professor in the U.S. system) with the Department of Computer Science and Information Technology, La Trobe University, Melbourne, Australia. To date, he has published more than 60 papers in leading international journals and conferences. He was a recipient of the World's Top 2% Scientists by Stanford University since 2022, the Best Paper Award at the IEEE International Symposium on Wireless Communication Systems in 2022, and an Exemplary Reviewer from IEEE TRANSACTIONS ON COMMUNICATIONS. He is a foundation member of IEEE ComSoc special interest group on orthogonal time frequency space (OTFS), an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS, Electronics, Future Internet, and Frontiers in Communications and Networks. His research interests include machine learning in communications, with a particular focus on integrated sensing and communication (ISAC), intelligent reflecting surface (IRS)-assisted communications, OTFS, the Internet of Things (IoT), and cognitive radio.



Guoqiang Mao (S'98–M'02–SM'08–F'18) is a Chair Professor, Director of the Center for Smart Driving and Intelligent Transportation Systems, Southeast University. Between 2014 and 2019, he was a Leading Professor, Founding Director of the Research Institute of Smart Transportation, and Vice-Director of the ISN State Key Lab at Xidian University. Before that he was with University of Technology Sydney and the University of Sydney. He has published 300 papers in international conferences and journals that have been cited more than 15,000 times. His H-

index is 57 and is in the list of Top 2% most-cited Scientists Worldwide by Stanford University in 2022, 2023 and 2024 both by Single Year and by Career Impact. He serves as a Vice-Director of Smart Transportation Information Engineering Society, Chinese Institute of Electronics (2022-), and was a co-chair IEEE ITS Technical Committee on Communication Networks (2014–2017). He is an editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (since 2018), IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2014–2019), IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (2010–2020) and received "Top Editor" award for outstanding contributions to the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2011, 2014 and 2015. He has served as a chair, co-chair and TPC member in a number of international conferences. His research interest includes intelligent transport systems, Internet of Things, wireless localization techniques, mobile communication systems, and applied graph theory and its applications in telecommunications. He is a Fellow of IEEE, AAIA and IET.