Dimension-Independent Channel Estimation for RIS-Assisted Communications: From Cascaded to Separate

Chao Jiang, *Graduate Student Member, IEEE*, Tianyue Zheng, *Graduate Student Member, IEEE*, Jieao Zhu, *Graduate Student Member, IEEE*, Shenheng Xu, *Member, IEEE*, and Linglong Dai, *Fellow, IEEE*

***Abstract*—Channel estimation in reconfigurable intelligent surface (RIS) assisted communications requires high pilot overhead due to numerous RIS elements incapable of signal processing. Recently, research on sensing RIS has provided a dimension-independent channel estimation scheme with merely three pilots. Nevertheless, it assumes that the BS-RIS channel is perfectly known to the RIS and remains invariant over a prolonged period, while inducing high hardware and power consumption. To address these issues, this paper introduces a generalized approximate message passing (GAMP) based channel estimation framework to achieve dimension-independent estimation of separate channels, without the assumption of known BS-RIS channel. Specifically, we first formulate the channel estimation problem in RIS assisted communications as a compressive phase retrieval problem. Based on the phaseless power observations, we leverage the GAMP algorithm to retrieve original sparse signals, which inherently supports the sparse-sampling sensing RIS architecture. Furthermore, by exploiting the intrinsic mapping between the sparse representations of the channels and the power observations in the angular domain, we propose a learned GAMP network to enhance the convergence stability and estimation accuracy. Finally, simulation results demonstrate that the proposed approach can efficiently estimate both the BS-RIS and UE-RIS channels with four pilots, while eliminating the requirement for prior knowledge of the BS-RIS channel and significantly reducing hardware and power consumption.**

***Index Terms*—Reconfigurable intelligent surface (RIS), channel estimation, generalized approximate message passing (GAMP), deep learning.**

I. Introduction

RECONFIGURABLE intelligent surface (RIS) has been envisioned as a promising technology due to its capability to reconfigure wireless channels [1], [2]. By properly manipulating the phase-shifts of RIS elements, the reflection signals can be reconfigured to desired directions with high beamforming gain [3]. Fully exploiting this high beamforming gain requires accurate channel state information (CSI), making channel estimation a critical procedure in RIS-aided communication systems [4]. Nevertheless, traditional RIS with no radio frequency (RF) chains can only reflect the incident signals passively. Therefore, the BS-RIS and RIS-UE channels cannot be separately acquired and only the cascaded UE-RIS-BS or BS-RIS-UE channel can be estimated at the BS or the UE side [5], [6]. The corresponding dimension of the cascaded channel, given by the product of the number of BS antennas M and the number of RIS elements N in practice, is N times that of the BS-UE channel. To achieve high beamforming gain, RIS is generally fabricated with extremely numerous elements, and the dimension of the cascaded channel increases sharply, resulting in high pilot overhead for channel estimation [7], [8].

A. Prior Works

To tackle the above challenge, numerous studies have exploited the structure properties of the cascaded channel and developed various channel estimation approaches with reduced pilot overhead. Generally, these approaches can be classified into three categories.

The first category leverages the spatial correlation among adjacent RIS elements. For example, the authors in [9] divided RIS elements into several groups, where each group consisted of adjacent elements and was configured with common reflection coefficients. Based on this method, the cascaded channel of each group was estimated sequentially with reduced pilot overhead by switching off the other groups. Further, an analytical framework for the optimal group size was presented in [10], [11] to maximize the achievable rate of the communications.

The second solution to reducing pilot overhead is to utilize the sparsity of channels in the transform domain. For example, researchers in [12] and [13] exploited the sparsity in the angular domain, and formulated the cascaded channel estimation problem as a compressive sensing (CS) problem, which could be solved by the orthogonal matching pursuit (OMP) algorithm [14]. Meanwhile, approximate message passing (AMP) approaches were utilized in [15], [16], [17], and [18] by iteratively performing inference on the factor graph. To avoid grid mismatch in the above schemes and

Received 28 September 2025; revised 12 March 2026 and 18 May 2026; accepted 22 May 2026. Date of current version 28 May 2026. This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFB3811503, in part by the National Science Fund for Distinguished Young Scholars under Grant 62325106, and in part by the National Natural Science Foundation of China under Grant 62031019. The associate editor coordinating the review of this article and approving it for publication was K. Adachi. *(Corresponding author: Linglong Dai.)*

The authors are with the Department of Electronic Engineering and the State Key Laboratory of Space Network and Communications, Tsinghua University, Beijing 100084, China (e-mail: jiangc24@mails.tsinghua.edu.cn; zhengty22@mails.tsinghua.edu.cn; zja21@mails.tsinghua.edu.cn; shxu@tsinghua.edu.cn; daill@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TWC.2026.3697234

pursuit super-resolution estimation, [19] divided the cascaded channel estimation problem into two CS sub-problems and applied atomic norm minimization (ANM) [20] to sequentially estimate continuous channel parameters. Moreover, several other off-grid channel estimation algorithms were proposed to improve estimation accuracy through optimized OMP [21], deep learning [22], and sparse Bayesian learning [23], [24].

The third category exploits the explicit structure properties that the BS-RIS channel is shared among all users. Based on this multi-user correlation, several ingenious channel estimation schemes have been proposed with low pilot overhead. Specifically, [25], [26], [27] first estimated the cascaded channel of a typical user, and then the cascaded channels of the other users are estimated with low pilot overhead based on their strong correlation with that of the typical user. Moreover, the authors in [28] revealed that the angular representations of the cascaded channels exhibit entirely common non-zero rows and partially common non-zero columns among all users, which could be leveraged to further reduce the pilot overhead. In [29], the shared column-block sparsity of the cascaded channels among different users was exploited to estimate the multi-user channels in an iterative manner.

B. Motivations

Despite substantial research on low-overhead RIS channel estimation schemes, all the above approaches are designed to accurately estimate the cascaded channel, and the corresponding pilot overhead is inevitably proportional to the number of RIS elements. Considering that the number of RIS elements in practical systems is generally extremely large (e.g. 1100 elements [30], 2304 elements [31], and 10240 elements [32]), the aforementioned schemes with *dimension-dependent* pilot overhead remain high.

To further reduce pilot overhead in RIS assisted wireless communications, we have previously conceived a *dimension-independent* channel estimation approach based on interference random field (IRF) [33]. Specifically, inspired by optical interference where phase information can be converted to power information, the signals from the BS and UE are required to be simultaneously transmitted to the RIS. Consequently, a similar interference phenomenon named IRF occurs, the power information of which can be exploited to retrieve channel phase information. To capture this power information, a novel hardware architecture named sensing RIS is designed, where each RIS element additionally integrates a low-cost power detector. Based on the received power observations, we further proposed a channel estimation framework with *dimension-independent* pilot overhead [33].

However, the channel estimation approach in [33] assumes that the perfect CSI of the BS-RIS channel is known to RIS and remains invariant over a prolonged period. Although this assumption is reasonable in static scenarios, it may not be valid in highly dynamic propagation environments. For example, when moving scatterers such as vehicles, autonomous aerial vehicles (AAVs), or pedestrians exist between the BS and the RIS, the CSI of the BS–RIS channel changes rapidly and becomes unknown to the RIS, thus rendering the above channel estimation scheme inapplicable. Meanwhile, sensing RIS achieves reduced pilot overhead at the cost of integrating additional power detectors. In the existing hardware architecture, each sensing RIS element necessitates a dedicated power detector to calculate optimal phase shifts, thereby incurring high power consumption and hardware overhead.

C. Our Contributions

To resolve the above problems, we propose generalized AMP (GAMP) and learned GAMP based channel estimation frameworks to achieve dimension-independent estimation of separate channels with sparse-sampling power observations. Specifically, the contributions of this paper are summarized as follows.[1]

- Inspired by the four-step phase-shift technique in holographic imaging, and with the assumption of channel sparsity in the angular domain, we formulate the channel estimation problem in sensing RIS assisted communications as a compressive phase retrieval problem. Compared to existing CS-based channel estimation schemes that require both amplitude and phase information of the received signals, our proposed approach solely relies on the phaseless power observations. Moreover, by formulating the problem with a block-structured sensing matrix, the BS–RIS and UE–RIS channels contribute to the power observations in a distinguishable manner, thus enabling individual estimation of the two separate channels.
- To resolve the formulated compressive phase retrieval problem, we leverage the GAMP algorithm to develop a practical channel estimation scheme. This scheme enables computationally efficient and accurate reconstruction for the sparse angular-domain representations of the BS–RIS and the UE–RIS channels. It is worth noting that, this approach does not require the assumption that the BS-RIS channel is known, and utilizes phaseless power observations of only four pilots to achieve accurate channel estimation.
- To reduce the power consumption of sensing RIS, we propose two sparse-sampling sensing RIS architectures, which is equivalent to performing row sampling on the sensing matrix. In these architectures, only a small portion of RIS elements are equipped with power detectors. According to whether these RIS elements are uniformly distributed, we develop the regular architecture and the irregular architecture as two possible realizations of the sparse-sampling sensing RIS.
- Furthermore, we propose an enhanced learned GAMP algorithm that capitalizes on the sparsity properties of both the channels and the interference random field in the angular domain. By exploiting the intrinsic mapping between sparse representations of the channels and the interference random fields, we develop a simple initial-value network for parameter initialization and transform the conventional fixed empirical damping coefficient into trainable parameters. Such improvement can significantly

[1]Simulation codes will be provided to reproduce the results in this paper: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html

enhance the algorithm's convergence stability and estimation accuracy.

- Simulation results demonstrate that the proposed GAMP algorithm-based channel estimation approach achieves separate estimation of both the BS-RIS and the UE-RIS channels with sparse-sampling power observations. This approach requires merely four pilots, which is independent of the number of RIS elements. Moreover, the developed learned GAMP algorithm significantly enhances conventional GAMP in convergence stability, iteration efficiency, and signal recovery precision.

D. Organization and Notation

1) Organization: The rest of this paper is organized as follows. In Section II, we present the system model, the proposed hardware architecture of sparse-sampling sensing RIS, and the formulation of channel estimation problem. Subsequently, in Section III, GAMP algorithm and our proposed channel estimation scheme are elaborated, followed by an extension to multi-antenna transceivers. Then, we point out the limitations of traditional GAMP algorithm, and further propose a learned GAMP framework based on deep learning in Section IV. The computational complexity of the proposed scheme is also analyzed in this section. Finally, simulation results are presented in Section V, followed by the conclusion of this paper in Section VI.

2) Notation: $(\cdot)^{-1}$, $(\cdot)^{*}$, $(\cdot)^{T}$, and $(\cdot)^{H}$ denote the inverse, conjugate, transpose, and conjugate transpose operations, respectively; $\{L\}$ represents the set of integers $\{0, 1, \cdots, L-1\}$; $\|\cdot\|_2$ is the $\mathcal{L}_2$-norm of its argument; $\mathcal{CN}(\mu, \nu)$ represents the complex Gaussian distribution with mean μ and variance ν; $\left((n)\right)_N$ represents the integer n modulo N within range $\{N\}$; $\otimes$ ($\odot$) denotes the Kronecker (Hadamard) product.

II. System Model for Sensing RIS

The system model of the sensing RIS assisted communications will be presented in this section. Subsection II-A will firstly specify the channel model, while Subsection II-B will elaborate the signal model for transmission and channel estimation. Then, we will describe the hardware architecture of the proposed sparse-sampling sensing RIS in Subsection II-C. Finally, the corresponding channel estimation problem will be formulated as a compressive phase retrieval problem in Subsection II-D.

A. Channel Model

This work considers a RIS-aided wireless communication system, in which one user equipment (UE) is served by a base station (BS) with the assistance of RIS. To simplify the design and analysis of the algorithms in this paper, it is assumed that the direct BS-UE link is completely blocked and thus severely attenuated, as shown in Fig. 1. To counteract the above blockage, a RIS is deployed between the BS and the UE to enable the establishment of a reflective link. Actually, the existence of the BS-UE channel has no influence on the problem formulation and the proposed channel estimation framework. Moreover, even if the direct BS-UE link is not blocked, it can still be estimated independently using traditional methods by simply turning off the RIS during a separate pilot phase [34].

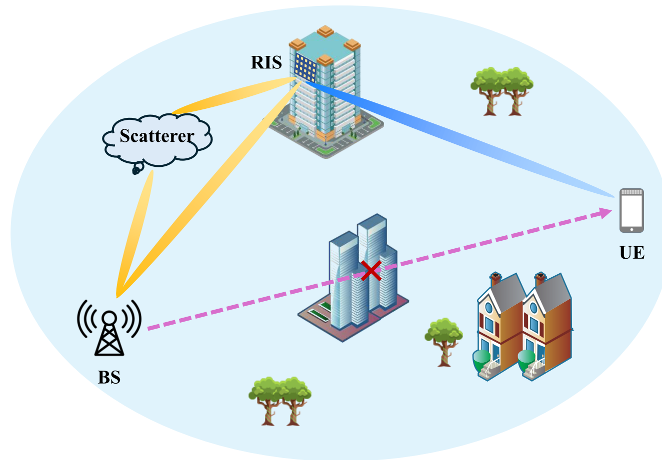


Fig. 1. A schematic diagram of RIS assisted communications.

The RIS is equipped with a uniform planar array (UPA) of N elements, and both the BS and the UE have a single-antenna.[2] Let $\mathbf{g} \in \mathbb{C}^N$ ($\mathbf{f} \in \mathbb{C}^N$) denote the channel from the BS (UE) to the RIS, respectively. The multipath Saleh-Valenzuela channel model, as described in [35], can be employed to characterize $\mathbf{g}$ as[3]

$$\mathbf{g} = \sqrt{\frac{N}{L_g}} \sum_{\ell_1=0}^{L_g-1} \alpha_{\ell_1}^{g} \mathbf{a}(\theta_{\ell_1}^{g}, \phi_{\ell_1}^{g}), \tag{1}$$

where L_g denotes the number of paths between the RIS and the BS, while $\alpha_{\ell_1}^{g}$ and $\theta_{\ell_1}^{g}$ ($\phi_{\ell_1}^{g}$) denotes the complex path gain and the azimuth (elevation) angle at the RIS for the ℓ_1-th path, respectively. Similarly, the UE-RIS channel $\mathbf{f}$ is given by

$$\mathbf{f} = \sqrt{\frac{N}{L_f}} \sum_{\ell_2=0}^{L_f-1} \alpha_{\ell_2}^{f} \mathbf{a}(\theta_{\ell_2}^{f}, \phi_{\ell_2}^{f}), \tag{2}$$

where L_f denotes the number of paths between the RIS and the UE, and $\alpha_{\ell_2}^{f}$, $\theta_{\ell_2}^{f}$ ($\phi_{\ell_2}^{f}$) denotes the complex path gain, the azimuth (elevation) angle at the RIS for the ℓ_2-th path. $\mathbf{a}(\theta, \phi) \in \mathbb{C}^N$ represents the normalized array steering vector for the RIS. Given a typical UPA with $N = N_1 \times N_2$ elements, $\mathbf{a}(\theta, \phi)$ can be represented as

$$\mathbf{a}(\theta, \phi) = \frac{1}{\sqrt{N}} \left[e^{-j2\pi d \sin(\theta)\cos(\phi)\mathbf{n}_1/\lambda} \right] \otimes \left[e^{-j2\pi d \sin(\phi)\mathbf{n}_2/\lambda} \right], \tag{3}$$

where $\mathbf{n}_1 = [0, 1, \cdots, N_1-1]^T$ and $\mathbf{n}_2 = [0, 1, \cdots, N_2-1]^T$, λ denotes the carrier wavelength, and d is the antenna spacing satisfying $d = \lambda/2$.

Furthermore, since both the BS-RIS channel $\mathbf{g}$ and the UE-RIS channel $\mathbf{f}$ are sparse in the angular domain, the spatial

[2]This configuration is adopted for simplicity of exposition. Our channel estimation scheme can be easily extended to multi-antenna transceivers, which will be described in Subsection III-D briefly.

[3]For analytical simplicity, the angular spread of each scatterer is assumed to be infinitesimal. In practical scenarios with non-negligible angular spreads, the proposed channel estimation approach remains effective.

domain channels $\mathbf{g}$ and $\mathbf{f}$ can be represented in the angular domain as

$$\mathbf{g} = \mathbf{U}\tilde{\mathbf{g}}, \tag{4}$$

$$\mathbf{f} = \mathbf{U}\tilde{\mathbf{f}}, \tag{5}$$

where $\tilde{\mathbf{g}} \in \mathbb{C}^N$ and $\tilde{\mathbf{f}} \in \mathbb{C}^N$ are sparse angular-domain representations of the BS-RIS and UE-RIS channels, respectively. Besides, the codebook $\mathbf{U} \in \mathbb{C}^{N\times N}$ is defined to be

$$\begin{aligned}\mathbf{U} = \big[&\bar{\mathbf{a}}(\bar{\theta}_1, \bar{\phi}_1) \cdots, \bar{\mathbf{a}}(\bar{\theta}_1, \bar{\phi}_{N_2}), \cdots, \\ &\bar{\mathbf{a}}(\bar{\theta}_{N_1}, \bar{\phi}_1) \cdots, \bar{\mathbf{a}}(\bar{\theta}_{N_1}, \bar{\phi}_{N_2})\big],\end{aligned} \tag{6}$$

where the spatial angles of azimuth and elevation are represented as $\bar{\theta}_n = \frac{2}{N_1}\left(n - \frac{N_1+1}{2}\right), n = 1, 2, \cdots, N_1$ and $\bar{\phi}_n = \frac{2}{N_2}\left(n - \frac{N_2+1}{2}\right), n = 1, 2, \cdots, N_2$. Similar to (3), the codebook atoms $\bar{\mathbf{a}}(\bar{\theta}, \bar{\phi})$ can be presented by

$$\bar{\mathbf{a}}(\bar{\theta}, \bar{\phi}) = \frac{1}{\sqrt{N}}\left[e^{-j\pi\bar{\theta}\mathbf{n}_1}\right] \otimes \left[e^{-j\pi\bar{\phi}\mathbf{n}_2}\right]. \tag{7}$$

With the assumption of infinitesimal angular spread, according to the multipath channel model (1) and (2), the sparse vectors $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{f}}$ have no more than L_g and L_f non-zero elements, respectively. However, when the angular spread of each scatterer is non-negligible, the sizes of the support sets of $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{f}}$ can be larger than L_g and L_f [36].

Moreover, it should be noted that the mutual coupling effect is ignored for analysis simplification, as commonly done in most existing literature on RIS channel estimation. Actually, for rectangular arrays with $d = \lambda/2$ antenna spacing, the antenna spacing along the diagonal directions is $\lambda/\sqrt{2}$, which may render the mutual coupling effect non-negligible [37], [38]. With the mutual coupling matrix $\mathbf{C}_r \in \mathbb{C}^{N\times N}$ considered, the equivalent channels can be represented as $\mathbf{g}' = \mathbf{C}_r\mathbf{g}$ and $\mathbf{f}' = \mathbf{C}_r\mathbf{f}$, which still exhibit sparsity under the equivalent codebook $\mathbf{U}' = \mathbf{C}_r\mathbf{U}$. In this context, the proposed channel estimation framework remains applicable.

B. Signal Model

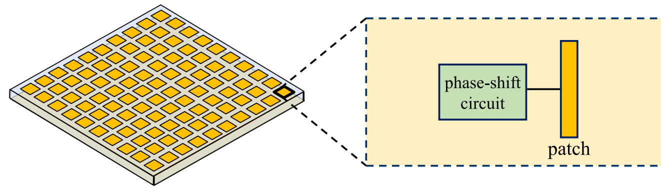
Given the channel model, the received signal of the UE can be represented as

$$y_{\mathrm{UE}} = \sqrt{P_{\mathrm{BS}}}\mathbf{f}^H\mathbf{\Psi}\mathbf{g}s + v, \tag{8}$$

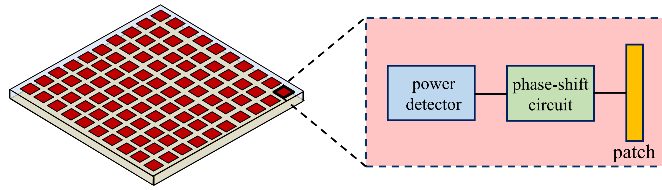
where P_{BS} is the transmitting power of the BS; s denotes the normalized BS transmitting symbol satisfying $\mathbb{E}[|s|^2] = 1$; $\mathbf{\Psi} = \mathrm{diag}(\boldsymbol{\psi}) = \mathrm{diag}\left([\psi_1, \cdots, \psi_N]^T\right)$ is the RIS phase-shift matrix with $\psi_n \in \mathbb{C}$, $|\psi_n| = 1$ representing the phase-shift of the n-th RIS element ($n \in \{N\}$); $v \sim \mathcal{CN}(0, \nu^v)$ is the received thermal noise introduced at the UE. If the channels $\mathbf{g}$ and $\mathbf{f}$ are known, then we can easily obtain the RIS phase-shift matrix $\mathbf{\Psi}$ as

$$\psi_n^{\mathrm{opt}} = \exp\left(-j\arg(g_n f_n^*)\right), \ \forall n \in \{N\}. \tag{9}$$

However, due to the rapid changing of the wireless environment, these channel vectors require frequent re-estimation, causing unaffordable pilot overhead in traditional RIS assisted communications. Intuitively speaking, due to the single-antenna BS and passively reflecting RIS elements, only one complex number can be obtained during each uplink channel estimation time slot. This severe lack of information[4] causes the inefficiency of the traditional uplink channel estimation.

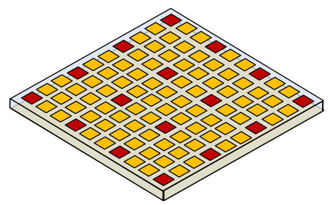


(a) Traditional RIS.

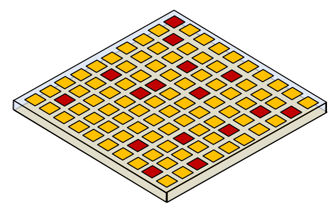


(b) Sensing RIS.

Fig. 2. Comparison of different hardware architectures.



(a) regular architecture.



(b) irregular architecture.

Fig. 3. Proposed sparse-sampling sensing RIS architecture.

To address this problem, we replace the conventional passive RIS in Fig. 1 with a sensing RIS [33]. Compared with traditional RIS structure as a phase-shift circuit shown in Fig. 2(a), sensing RIS elements additionally integrate power detectors for capturing the power information of the IRF signal, which is exhibited in Fig. 2(b). The detected IRF signal $\mathbf{p} \in \mathbb{R}_{\geq 0}^N$ can be represented as

$$\mathbf{p} = \left|\sqrt{P_{\mathrm{BS}}}\mathbf{g}s + \sqrt{P_{\mathrm{UE}}}\mathbf{f}s' + \mathbf{w}\right|^2, \tag{10}$$

where $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \nu^w\mathbf{I}_N)$ represents the noise introduced in the power detectors, P_{UE} denotes the transmitting power of the UE, and s' denotes the normalized UE transmitting symbol satisfying $\mathbb{E}[|s'|^2] = 1$.

C. Sparse Sampling Sensing RIS Architecture

As discussed in Subsection I-A, one of the main challenges for the sensing RIS-aided communications lies in the high hardware and power consumption. To overcome this challenge and facilitate the practical operation of sensing RIS, we propose a new architecture based on sparse-distributed

[4]It is worth noting that, deploying a multi-antenna BS with M antennas cannot mitigate this shortage of data, since the dimension of the BS-RIS channel scales linearly with M.

power detectors. Compared with the existing sensing RIS architecture, in which each RIS element integrates a dedicated power detector to capture the power of IRF, in the proposed sparse sensing RIS, only a limited subset of the RIS elements is equipped with power detectors. Depending on whether these RIS elements are uniformly distributed, we develop the regular architecture and the irregular architecture, as illustrated in Fig. 3, respectively. In this manner, the quantity of power detectors decreases greatly for low hardware and power consumption, at the price of a limited number of power observations.

Suppose that the sparse-sampling sensing RIS employs N_s power detectors scattered among all the N reflecting elements. The positions of these detectors are indexed by a selection matrix $\mathcal{H} \in \{0,1\}^{N_1 \times N_2}$, which is defined as

$$\mathcal{H}_{k,\ell} = \begin{cases} 1, & \text{power detector present at } (k,\ell) \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Then, the detected IRF signal $\mathbf{p}$ in (10) can be rewritten as

$$\mathbf{p} = \left| \sqrt{P_{\mathrm{BS}}}\mathbf{g}s + \sqrt{P_{\mathrm{UE}}}\mathbf{f}s' + \mathbf{w} \right|^2 \odot \mathrm{vec}(\mathcal{H}). \tag{12}$$

Given our exclusive focus on the N_s non-zero entries of the IRF signal, the vector $\mathbf{p}$ is compressed into $\mathbf{p}_{\mathcal{H}} \in \mathbb{R}_{\geq 0}^{N_s}$, where the non-zero entries are gathered together.

D. Problem Formulation

The objective of this paper is to estimate the separate channels of BS-RIS and UE-RIS from the power observations $\mathbf{p}$ defined in (12). The separate channels in the angular domain is denoted as $\mathbf{x} \in \mathbb{C}^{2N}$,

$$\mathbf{x} = \left[\sqrt{P_{\mathrm{BS}}}\tilde{\mathbf{f}}^T, \sqrt{P_{\mathrm{UE}}}\tilde{\mathbf{g}}^T \right]^T. \tag{13}$$

To achieve this, we necessitate fully exploiting both the sparse prior knowledge and the physical properties of the interference process. As delineated in the preceding channel model II-A, the sparsity assumption lies in the angular domain, thereby requiring a codebook $\mathbf{A}$ to represent the non-sparse observations $\mathbf{Ax}$ by the underlying sparse vector $\mathbf{x}$. Moreover, due to the phaseless characteristic of the interference process, the actual observed signal is $|\mathbf{Ax}|$, which is the magnitude of the complex signal $\mathbf{Ax}$.

Inspired by the four-step phase-shift methodology in holographic imaging [39], which can extract phase information from $Q = 4$ exposures implemented by imposing different phase-shifts to one optical path, the codebook $\mathbf{A} \in \mathbb{C}^{4N_s \times 2N}$ can be defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{U}_{\mathcal{H}} & \mathbf{O}_{N_s \times N} \\ \mathbf{O}_{N_s \times N} & \mathbf{U}_{\mathcal{H}} \\ \mathbf{U}_{\mathcal{H}} & \jmath\mathbf{U}_{\mathcal{H}} \\ \mathbf{U}_{\mathcal{H}} & \mathbf{U}_{\mathcal{H}} \end{bmatrix}, \tag{14}$$

where $\mathbf{O}_{N_s \times N}$ denotes the all-zero matrix of size $N_s \times N$, and $\mathbf{U}_{\mathcal{H}}$ characterizes the row-wise selection mechanism determined by the non-zero entries of $\mathrm{vec}(\mathcal{H})$.

Since the estimated sparse signal $\mathbf{x}$ needs to simultaneously satisfy the IRF observation equation $\mathbf{y} = \sqrt{\mathbf{p}}$ and the sparsity prior, an optimization problem can be formulated as

$$\min_{\mathbf{x}} \ \mathcal{L}(\mathbf{x}) = \frac{1}{2}\left\| \mathbf{y} - |\mathbf{Ax}| \right\|_2^2 + \lambda \left\| \mathbf{x} \right\|_1, \tag{15}$$

where $\lambda > 0$ denotes the regularization factor balancing prior knowledge on sparsity and error of solution, and the received magnitude signal vector $\mathbf{y} \in \mathbb{R}_{\geq 0}^{4N_s}$ can be expressed as

$$\mathbf{y} = \left[\sqrt{\mathbf{p}}_{1,\mathcal{H}}^T, \sqrt{\mathbf{p}}_{2,\mathcal{H}}^T, \sqrt{\mathbf{p}}_{3,\mathcal{H}}^T, \sqrt{\mathbf{p}}_{4,\mathcal{H}}^T \right]^T. \tag{16}$$

If the noncoherent sparse recovery problem (15) can be solved efficiently, then the separate channels $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}$ can be acquired accordingly. Unfortunately, due to the absolute term $|\mathbf{Ax}|$, the problem is highly non-convex.

III. Channel Estimation via GAMP Algorithm

This section will introduce a GAMP algorithm based channel estimation scheme to solve the problem formulated in (15). Specifically, in Subsection III-A, a brief introduction to the GAMP algorithm will be presented. Then, we will utilize the GAMP algorithm to recover full channels from sampled IRF signals in Subsection III-B. After that, we will investigate inherent ambiguity in the GAMP based channel estimation scheme and discuss how to eliminate it in Subsection III-C. Finally, in Subsection III-D, the proposed channel estimation scheme's extension to multi-antenna transceivers will be illustrated briefly.

A. Preliminaries of GAMP Algorithm

Various methods can be applied to resolve the noncoherent sparse recovery problem formulated in Subsection II-D. The aforementioned type of problem is typically categorized as a sparse phase retrieval problem, characterized by the reconstruction of the underlying sparse signal vector $\mathbf{x}$ exclusively from phaseless observations $\mathbf{y}$. Among these methods, GAMP algorithm [40] has emerged as a prominent solution, demonstrating remarkable efficiency and robustness in both theoretical analysis and practical implementation scenarios.

GAMP is an iterative Bayesian inference algorithm designed to well-resolve high-dimensional linear inverse problems. As an extension of the standard AMP framework [41], GAMP algorithm can incorporate arbitrary input and output distributions, and is applicable to linear as well as nonlinear observation models. Moreover, GAMP algorithm exhibits superior computational efficiency, owing to each iteration exclusively involving scalar operations rather than complicated matrix manipulations.

Focusing on problem (15), if the posterior distribution $p(\mathbf{x}|\mathbf{y})$ can be available, we can implement maximum a posteriori (MAP), as well as minimum mean square error (MMSE) estimation, which can be obtained as

$$(\hat{\mathbf{x}})_{\mathrm{MAP}} = \operatorname*{argmax}_{\mathbf{x}} \ p(\mathbf{x}|\mathbf{y}), \tag{17}$$

$$\begin{aligned} (\hat{\mathbf{x}})_{\mathrm{MMSE}} &= \operatorname*{argmax}_{\hat{\mathbf{x}}} \ \mathbb{E}\left[\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 | \mathbf{y} \right] \\ &= \mathbb{E}\left[\mathbf{x}|\mathbf{y} \right]. \end{aligned} \tag{18}$$

Algorithm 1 Sum-Product GAMP Algorithm

Input: Sparse sensing matrix $\mathbf{A} \in \mathbb{C}^{4N_s \times 2N}$, observed amplitude signal $\mathbf{y} \in \mathbb{R}_{\geq 0}^{4N_s}$, precision tolerance ϵ, maximum number of iterations $T_{\max}$.

Output: Estimated sparse signal $\mathbf{x} \in \mathbb{C}^{2N}$.

1: $\mathbf{B} = \mathrm{abs}(\mathbf{A})^2$
2: $t \leftarrow 1$
3: **repeat**
4: $\boldsymbol{\nu}^p(t) = \beta \mathbf{B} \boldsymbol{\nu}^x(t) + (1-\beta)\boldsymbol{\nu}^x(t-1)$
5: $\alpha(t) = \mathrm{mean}(\boldsymbol{\nu}^p(t))$
6: $\hat{\mathbf{p}}(t) = \mathbf{A}\hat{\mathbf{x}}(t) - \frac{1}{\alpha(t)}\boldsymbol{\nu}^p(t) \odot \hat{\mathbf{s}}(t-1)$
7: $\hat{\mathbf{s}}(t) = \beta g_{\mathrm{out}}(\mathbf{y}, \hat{\mathbf{p}}(t), \boldsymbol{\nu}^p(t)) + (1-\beta)\hat{\mathbf{s}}(t-1)$
8: $\boldsymbol{\nu}^s(t) = (1-\beta)\boldsymbol{\nu}^s(t-1) - \beta\alpha(t) g'_{\mathrm{out}}(\mathbf{y}, \hat{\mathbf{p}}(t), \boldsymbol{\nu}^p(t))$
9: $\boldsymbol{\nu}^r(t) = \left(\mathbf{B}^T \boldsymbol{\nu}^s(t)\right)^{-1}$
10: $\hat{\mathbf{r}}(t) = \hat{\mathbf{x}}(t) + \boldsymbol{\nu}^r(t) \odot \left(\mathbf{A}^H \hat{\mathbf{s}}(t)\right)$
11: $\bar{\mathbf{x}}(t) = \beta\hat{\mathbf{x}}(t) + (1-\beta)\bar{\mathbf{x}}(t-1)$
12: $\boldsymbol{\nu}^x(t+1) = \alpha(t)\boldsymbol{\nu}^r(t) \odot g'_{\mathrm{in}}(\hat{\mathbf{r}}(t), \alpha(t)\boldsymbol{\nu}^r(t))$
13: $\hat{\mathbf{x}}(t+1) = g_{\mathrm{in}}(\hat{\mathbf{r}}(t), \alpha(t)\boldsymbol{\nu}^r(t))$
14: $t \leftarrow t+1$
15: **until** $t > T_{\max}$ or $\|\hat{\mathbf{x}}(t+1) - \bar{\mathbf{x}}(t)\|^2 < \epsilon$

However, the true posterior distribution $p(\mathbf{x}|\mathbf{y})$ is analytically intractable and computationally prohibitive. To address this problem, the GAMP algorithm performs loopy belief propagation (BP) on a factor graph. Specifically, assuming that the components of the unknown vector $\mathbf{x}$ are statistically independent, we can factorize the posterior distribution as

$$\begin{aligned} p(\mathbf{x}|\mathbf{y}) &\propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \\ &= \prod_m p(y_m|z_m) \prod_n p(x_n), \end{aligned} \tag{19}$$

where $\mathbf{z} \triangleq \mathbf{A}\mathbf{x}$ denotes a linear transform. Then the factors of $p(\mathbf{x}|\mathbf{y})$ are represented as factor nodes and each x_n constitutes variable nodes in the factor graph. Through message passing between these nodes, the estimation of the variable $\mathbf{x}$ is iteratively updated until all nodes converge towards a consensus. Moreover, by employing the central limit theorem (CLT) in conjunction with quadratic approximations, the message passing between the variable nodes and the factor nodes can be significantly simplified.

It is worth noting that the GAMP algorithm provides computationally efficient approximations of both max-sum and sum-product loopy BP. The former corresponds to MAP estimation, while the latter is the implementation of MMSE estimation. In this paper, we concentrate exclusively on the sum-product form of the GAMP algorithm, for the MMSE estimation of the sparse signal $\mathbf{x}$.

B. Recover Full Channels From Sampled IRF

This subsection will present how to leverage the sum-product GAMP algorithm to recover sparse signal $\mathbf{x}$ from phaseless observations $\mathbf{y}$, as summarized in Algorithm 1.

The output nonlinear functions $g_{\mathrm{out}}(\cdot)$ and $g'_{\mathrm{out}}(\cdot)$ in lines 7 and 8 of Algorithm 1 are defined as

$$g_{\mathrm{out}}(y_m, \hat{p}_m, \nu_m^p) = \frac{\mathbb{E}\left[z_m|y_m; \hat{p}_m, \nu_m^p\right] - \hat{p}_m}{\nu_m^p}, \tag{20}$$

$$g'_{\mathrm{out}}(y_m, \hat{p}_m, \nu_m^p) = \frac{\mathrm{var}\left[z_m|y_m; \hat{p}_m, \nu_m^p\right] - \nu_m^p}{(\nu_m^p)^2}. \tag{21}$$

Since $\mathbf{y} = |\mathbf{z} + \mathbf{w}|$, the distribution of y_m conditioned on z_m follows a Rician distribution

$$p(y_m|z_m) = \frac{2y_m}{\nu^w} \exp\left(-\frac{y_m^2 + |z_m|^2}{\nu^w}\right) I_0\left(\frac{2y_m|z_m|}{\nu^w}\right), \tag{22}$$

where $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind. Moreover, by utilizing the CLT, GAMP algorithm approximates z_m following a Gaussian distribution with mean $\hat{p}_m$ and variance ν_m^p. Then the marginal posterior distribution $p(z_m|y_m; \hat{p}_m, \nu_m^p)$ can be approximated as

$$p(z_m|y_m; \hat{p}_m, \nu_m^p) = \frac{p(y_m|z_m)\mathcal{CN}(z_m; \hat{p}_m, \nu_m^p)}{\int p(y_m|z_m)\mathcal{CN}(z_m; \hat{p}_m, \nu_m^p) dz_m}. \tag{23}$$

Thus, the conditional mean and variance of $z_m|y_m$ are readily available, as derived in [42]

$$\begin{aligned} &\mathbb{E}\left[z_m|y_m; \hat{p}_m, \nu_m^p\right] \\ &= \left(\frac{y_m}{1 + \nu^w/\nu_m^p} R_0(\varrho_m) + \frac{|\hat{p}_m|}{1 + \nu_m^p/\nu^w}\right) \frac{\hat{p}_m}{|\hat{p}_m|}, \end{aligned} \tag{24}$$

$$\begin{aligned} &\mathrm{var}\left[z_m|y_m; \hat{p}_m, \nu_m^p\right] \\ &= \frac{y_m^2}{(1 + \nu^w/\nu_m^p)^2} + \frac{|\hat{p}_m|^2}{(1 + \nu_m^p/\nu^w)^2} + \frac{1 + \varrho_m R_0(\varrho_m)}{1/\nu^w + 1/\nu_m^p} \\ &\quad - \left|\mathbb{E}\left[z_m|y_m; \hat{p}_m, \nu_m^p\right]\right|^2, \end{aligned} \tag{25}$$

where

$$R_0(\varrho_m) \triangleq \frac{I_1(\varrho_m)}{I_0(\varrho_m)} \quad \text{and} \quad \varrho_m \triangleq \frac{2y_m|\hat{p}_m|}{\nu^w + \nu_m^p}. \tag{26}$$

As mentioned above, the GAMP algorithm approximates z_m following the Gaussian distribution $\mathcal{CN}(z_m; \hat{p}_m, \nu_m^p)$. With the iterations proceeding, the variance ν_m^p, also representing the uncertainty of the estimation of z_m, gradually decreases. Therefore, under high SNR conditions, ϱ_m becomes large and the complicated computation of $R_0(\varrho_m)$ in (24) and (25) can be approximated by [43]

$$R_0(\varrho_m) \approx \frac{8\varrho_m - 3}{8\varrho_m + 1}, \qquad \varrho_m > 10 \tag{27}$$

In contrast, under low SNR conditions, ϱ_m is generally small and $R_0(\varrho_m)$ can also be simplified as

$$R_0(\varrho_m) \approx \frac{\varrho_m}{2}, \qquad 0 < \varrho_m < 0.2 \tag{28}$$

Consequently, the computational complexity can be significantly mitigated per iteration without compromising accuracy.

To derive the specific expressions of the input nonlinear functions $g_{\mathrm{in}}(\cdot)$ and $g'_{\mathrm{in}}(\cdot)$, we assume that $\mathbf{x}$ is known to exhibit k-sparsity but lacks prior information on its support. Then, the Bernoulli-Gaussian (BG) model is conventionally adopted

$$p(x_n) = (1-\kappa)\delta(x_n) + \kappa\mathcal{CN}(x_n; 0, \sigma_x^2), \tag{29}$$

where $\kappa \triangleq k/N$ denotes the sparsity rate and σ_x^2 is the variance. For this BG prior, as available in [42] and [44], the input nonlinear functions can be represented as

$$g_{\text{in}}(\hat{r}_n, \nu_n^r) = \mathbb{E}\left[x_n|\hat{r}_n; \nu_n^r\right] = \frac{\xi_n}{1+\eta_n e^{-\zeta_n|\hat{r}_n|^2}}, \quad (30)$$

$$g'_{\text{in}}(\hat{r}_n, \nu_n^r) = \text{var}\left[x_n|\hat{r}_n; \nu_n^r\right] = \frac{\nu_n^r \xi_n(1+\zeta_n|\hat{r_n}|^2)}{1+\eta_n e^{-\zeta_n|\hat{r_n}|^2}} - \left|\mathbb{E}\left[x_n|\hat{r}_n; \nu_n^r\right]\right|^2, \quad (31)$$

where

$$\xi_n \triangleq \frac{\sigma_x^2}{\nu_n^r + \sigma_x^2}, \quad (32)$$

$$\eta_n \triangleq \frac{1-\kappa}{\kappa}\frac{\nu_n^r + \sigma_x^2}{\nu_n^r}, \quad (33)$$

$$\zeta_n \triangleq \frac{\sigma_x^2}{\nu_n^r(\nu_n^r + \sigma_x^2)}. \quad (34)$$

It is worth illustrating that, to enhance the convergence of the considered GAMP algorithm, a damping factor $\beta \in (0, 1]$ is incorporated to regulate the iterative variables updates, leading to a trade-off between guaranteed convergence and reduced convergence speed. Note that through numerous experiments, β is typically configured to be 0.25 and exhibits superior performances for phase retrieval [42].

C. Eliminate Inherent Ambiguity

Considering phase retrieval problems in holographic imaging, computational imaging, and signal processing, the phaseless measurements are typically obtained via the Discrete Fourier Transform (DFT). Due to the loss of phase information and the properties of DFT, these problems are inherently ill-posed and admit three unavoidable phase ambiguities. In this Subsection, we will introduce three different categories of phase ambiguities, their impacts on the problem, and the corresponding solutions.

1) Global Phase Ambiguity: If the vector $\mathbf{x}$ constitutes a solution, then $\mathbf{x}' = \mathbf{x}e^{j\varphi}$ also satisfies the solution requirements for any global phase-shift $\varphi \in \mathbb{R}$. Fortunately, according to (9), the global phase ambiguity present in the estimated channels $\hat{\tilde{\mathbf{g}}}$ and $\hat{\tilde{\mathbf{f}}}$ has no impact on RIS's phase-shift configuration, thus eliminating the need for explicit ambiguity removal.

2) Conjugate Reflection Ambiguity: Regarding the conjugate reflection of $\mathbf{x}$ (i.e., $x'_n = x_{N-n-1}, \forall n \in \{N\}$), the DFT of $\mathbf{x}'$ is represented as

$$\{\text{DFT}(\mathbf{x}')\}_k = e^{j2\pi k/N}\{\text{DFT}(\mathbf{x})\}_k^*. \quad (35)$$

Given merely magnitude information, we cannot distinguish between the DFT of $\mathbf{x}$ and that of $\mathbf{x}'$. However, in our phase retrieval problem, the vector $\mathbf{x}$ defined in (13) composed of two concatenated signals $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}$. Noticing the measurement matrix $\mathbf{A}$ given in (14), different phase-shifts are imposed on the DFTs of the two signals, resulting in $|\mathbf{A}\mathbf{x}| \neq |\mathbf{A}\mathbf{x}'|$. Therefore, the phase retrieval problem in this paper does not exhibit the conjugate reflection ambiguity.

3) Circular Shift Ambiguity: The circular shift version of $\mathbf{x}$ is defined as $x'_n = x_{((n-n_0))_N}, \forall n_0 \in \mathbb{Z}, n \in \{N\}$. Then the DFT of $\mathbf{x}'$ can be given by

$$\{\text{DFT}(\mathbf{x}')\}_k = e^{-j2\pi k n_0/N}\{\text{DFT}(\mathbf{x})\}_k. \quad (36)$$

It is straightforward to observe that the phaseless observation $|\mathbf{A}\mathbf{x}|$ is equal to $|\mathbf{A}\mathbf{x}'|$. Thus, we need to precisely estimate the circular shift n_0 in order to eliminate this inherent ambiguity.

It is worth noting that, in RIS assisted wireless communication system, the BS and the RIS are typically deployed at fixed locations with an unobstructed line-of-sight (LoS) channel between them. Hence, the maximum-value index of the sparse angular-domain representations for the BS-RIS channel $\tilde{\mathbf{g}}$, remains unchanged and can be predetermined as a priori. Comparing the maximum-value index with that of the reconstructed signal $\hat{\tilde{\mathbf{g}}}$, we can easily acquire the circular shift n_0 and eliminate the ambiguity.

D. Extension to Multi-Antenna Transceivers

As described previously, this paper merely considers single-antenna transceiver communication systems for expository simplicity. Here we will briefly demonstrate that the aforementioned channel estimation scheme remains applicable to multi-antenna transceivers.

Assume that an M-antenna BS serves a single-antenna UE with the assistance of an N-element sensing RIS. In order to acquire CSI, the BS and the UE should transmit known pilot symbols to the BS over T_s instants. In the t-th instant, the detected IRF signal can be represented as

$$\mathbf{p}_t = \left|\mathbf{G}\boldsymbol{\omega}_t s_t + \mathbf{f}\omega'_t s'_t + \mathbf{w}_t\right|^2, \quad (37)$$

where $\boldsymbol{\omega}_t \in \mathbb{C}^M$ denotes the precoding vector at the BS, with power constraint $\|\boldsymbol{\omega}_t\|_2^2 \leq P_{\text{BS}}$, while $|\omega'|^2 \leq P_{\text{UE}}$ is the UE precoding scalar introduced for the sake of formula symmetry. Letting $s_t = s'_t = 1$ and combining the received IRF signals of T_s instants, under the channel sparsity assumption, the concatenated signal $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_{T_s}]$ is written as

$$\mathbf{P} = \left|\mathbf{U}_N\widetilde{\mathbf{G}}\mathbf{U}_M^H\boldsymbol{\Omega} + \mathbf{U}_N\tilde{\mathbf{f}}\boldsymbol{\Omega}' + \mathbf{W}\right|^2, \quad (38)$$

where $\mathbf{U}_N$ denotes the $N \times N$ DFT matrix, and $\boldsymbol{\Omega} = [\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \cdots, \boldsymbol{\omega}_{T_s}]$, $\boldsymbol{\Omega}' = [\omega'_1, \omega'_2, \cdots, \omega'_{T_s}]$, $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{T_s}]$ are concatenated matrices consistent with $\mathbf{P}$. Vectorize $\mathbf{P}$ and then (38) can be rewritten as

$$\mathbf{p} = \text{vec}(\mathbf{P}) = |\underbrace{(\boldsymbol{\Omega}^T\mathbf{U}_M^*) \otimes \mathbf{U}_N}_{\mathbf{A}_1} \text{vec}(\widetilde{\mathbf{G}}) + \underbrace{\boldsymbol{\Omega}'^T \otimes \mathbf{U}_N}_{\mathbf{A}_2} \tilde{\mathbf{f}} + \text{vec}(\mathbf{W})|^2. \quad (39)$$

Intuitively, the channel estimation problem can also be formulated as a sparse signal phase retrieval problem, which remains addressable via the aforementioned GAMP algorithm. Compared with single-antenna transceiver system, the multi-antenna extension principally involves modifications to the measurement matrix $\mathbf{A}$, and introduces pilot overhead proportional to the number of transceiving antennas.

IV. Proposed Learned GAMP Network for Channel Estimation

In this section, we will firstly identify the inherent limitations of the traditional GAMP algorithm in Subsection IV-A. Then, we will perform spectrum analysis for the IRF signal and investigate its sparsity characteristic in the angular domain in Subsection IV-B. Subsequently, based on the spectrum analysis result, a deep learning enabled GAMP network will be proposed in Subsection IV-C. Finally, we will provide the computational complexity comparison in Subsection IV-D.

A. Challenges of GAMP Algorithm

GAMP algorithm achieves remarkable efficiency and robustness in addressing high-dimensional sparse signal phase retrieval problems, making it particularly attractive for signal processing across various fields. However, its practical applications face inherent limitations, particularly severe in our considered channel estimation for sensing RIS assisted system. Two principal drawbacks concerning convergence and iteration speed will be systematically described as follows.

On the one hand, as well investigated in [45], [46], sparse signal phase retrieval problems are inherently plagued by spurious local minima. GAMP algorithm is susceptible to convergence at suboptimal local solutions, leading to intolerable reconstruction errors, especially when the sensing matrix is ill-conditioned or the sparsity of the signal to be retrieved is not guaranteed strictly. To avoid bad local minima, the most typically adopted solution is implementing multiple random initializations, restarting the algorithm repeatedly, and selecting the reconstructed result with a minimum error as the final output [42]. Although this method can enhance convergence and reduce errors of the channel estimator through increasing repeated experiments, the resultant time complexity is fundamentally incompatible with low-latency requirements and prohibitive for practical applications in future wireless communication system.

The other critical drawback is the configuration of the damping factor β. Specifically, the theoretical determination of an optimal damping factor is challenging in existing algorithm architecture. An excessive β compromises convergence stability, whereas an insufficient value substantially decelerates convergence speed. Therefore, the damping factor is conventionally configured at an empirical value 0.25, rather than an optimized setting for the specific problem to be resolved. Additionally, the uniform configuration of identical damping factor across all variable updates in **Algorithm** 1 constitutes a significant methodological limitation, which disregards the heterogeneous convergence dynamics of different variables.

Consequently, these inherent limitations critically degrade the performance of GAMP based channel estimation scheme, particularly convergence stability and estimation errors. To address the above drawbacks, we develop a learned GAMP algorithm that capitalizes on the characteristics of IRF below.

B. IRF Spectrum Analysis

In order to retrieve the channel matrix from the IRF power observations $\mathbf{p}$, we need to exploit the structural properties of $\mathbf{p}$. Due to the sparsity of multipath channel in the angular domain as defined in (4) and (5), the IRF signals also exhibit sparsity in the angular domain. More importantly, the indices of non-zero elements of $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{f}}$ have a relationship with those of the sparse representation of $\mathbf{p}$.

Here we firstly consider the full-sampling architecture, and derive the relationship between the power signal $\mathbf{p}$ and the original phased $\mathbf{q} \triangleq \sqrt{P_{\mathrm{BS}}}\mathbf{g}s + \sqrt{P_{\mathrm{UE}}}\mathbf{f}s' + \mathbf{v}$. We construct $\mathbf{Q} \in \mathbb{C}^{N_1 \times N_2}$, $\mathbf{P} \in \mathbb{R}_{\geq 0}^{N_1 \times N_2}$, satisfying $\mathbf{q} = \mathrm{vec}(\mathbf{Q})$, $\mathbf{p} = \mathrm{vec}(\mathbf{P})$. Since $\mathbf{P} = \mathbf{Q} \odot \mathbf{Q}^*$, it follows that

$$\begin{aligned}\hat{\mathbf{P}}_{k1,k2} &= \frac{1}{N}\left(\hat{\mathbf{Q}} \star \widehat{\mathbf{Q}^*}\right)_{k_1,k_2} \\ &= \frac{1}{N}\sum_{n_2=0}^{N_2-1}\sum_{n_1=0}^{N_1-1}\hat{\mathbf{Q}}_{n_1,n_2}\hat{\mathbf{Q}}^*_{\left((n_1-k_1)\right)_{N_1},\left((n_2-k_2)\right)_{N_2}} \\ &= \frac{1}{N}\left(\mathbf{R}_{\hat{\mathbf{Q}},\hat{\mathbf{Q}}}\right)_{k_1,k_2}, \end{aligned} \tag{40}$$

where $\hat{\mathbf{P}}$ denotes the 2D-DFT of $\mathbf{P}$, $\star$ denotes circular convolution operation, and $\mathbf{R}_{\mathbf{A},\mathbf{B}}$ represents the correlation matrix between $\mathbf{A}$ and $\mathbf{B}$. Equation (40) demonstrates that the 2D-DFT of the non-coherent IRF signal $\mathbf{P}$ is the autocorrelation matrix of the spectrum of the coherent signal $\mathbf{q}$. To simplify analysis, we consider that both the BS-RIS channel and UE-RIS channel are single-path channels firstly, and then extend the analysis to multi-path scenario. By assuming an on-grid incident angle, the received coherent signal $\mathbf{q}$ is given by

$$\mathbf{q} = \sqrt{P_{\mathrm{BS}}}\alpha^g \mathbf{f}_{N_1,\ell_1} \otimes \mathbf{f}_{N_2,\ell_2} + \sqrt{P_{\mathrm{UE}}}\alpha^f \mathbf{f}_{N_1,\ell_1'} \otimes \mathbf{f}_{N_2,\ell_2'}, \tag{41}$$

where $\mathbf{f}_{N,\ell}$ denotes the ℓ-th column of the $N \times N$ DFT matrix. Thus, the spectrum of $\mathbf{Q} = \mathrm{unvec}(\mathbf{q})$ is given by

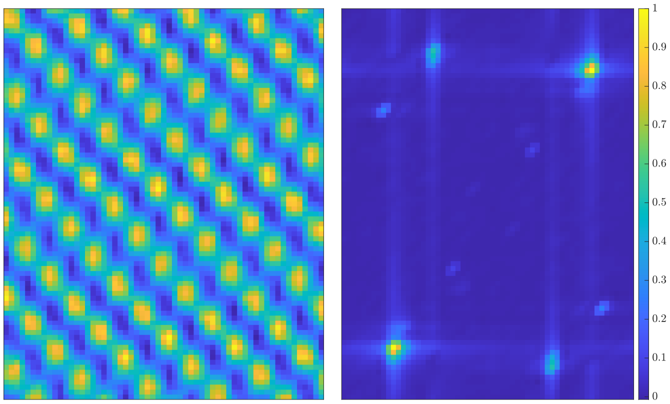
$$\hat{\mathbf{Q}} = \sqrt{P_{\mathrm{BS}}}\alpha^g \mathbf{e}_{\ell_1}\mathbf{e}_{\ell_2}^T + \sqrt{P_{\mathrm{UE}}}\alpha^f \mathbf{e}_{\ell_1'}\mathbf{e}_{\ell_2'}^T \triangleq \mathbf{Q}_1 + \mathbf{Q}_2, \tag{42}$$

where $\mathbf{e}_\ell$ denotes the unit vector with the only non-zero element in its ℓ-th entry. According to (40), the autocorrelation matrix of $\hat{\mathbf{Q}}$ is then given by

$$\begin{aligned}\mathbf{R}_{\hat{\mathbf{Q}},\hat{\mathbf{Q}}} &= \mathbf{R}_{\hat{\mathbf{Q}}_1,\hat{\mathbf{Q}}_1} + \mathbf{R}_{\hat{\mathbf{Q}}_1,\hat{\mathbf{Q}}_2} + \mathbf{R}_{\hat{\mathbf{Q}}_2,\hat{\mathbf{Q}}_1} + \mathbf{R}_{\hat{\mathbf{Q}}_2,\hat{\mathbf{Q}}_2} \\ &= (P_{\mathrm{BS}}|\alpha^g|^2 + P_{\mathrm{UE}}|\alpha^f|^2)\mathbf{e}_0\mathbf{e}_0^T \\ &\quad + \sqrt{P_{\mathrm{BS}}P_{\mathrm{UE}}}\alpha^g\alpha^{f*}\mathbf{e}_{\left((\ell_1-\ell_1')\right)_{N_1}}\mathbf{e}^T_{\left((\ell_2-\ell_2')\right)_{N_2}} \\ &\quad + \sqrt{P_{\mathrm{BS}}P_{\mathrm{UE}}}\alpha^{g*}\alpha^f\mathbf{e}_{\left((\ell_1'-\ell_1)\right)_{N_1}}\mathbf{e}^T_{\left((\ell_2'-\ell_2)\right)_{N_2}}. \end{aligned} \tag{43}$$

Therefore, the spectrum of the IRF signal exhibits a triple-peak pattern, with one peak $\mathbf{e}_0\mathbf{e}_0^T$ located at the zero-frequency origin, and another two conjugate-symmetric peaks located at the difference-frequency. The zero-frequency peak contains information about the BS-RIS and UE-RIS path gains, while the difference-frequency components correspond to the difference of incident angles of UE-RIS and BS-RIS beams. Thus, the indices of these two difference-frequency peaks are related to the indices of non-zero entries in $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{f}}$.

When extended to the multi-path scenario, the L_g BS-RIS paths and the L_f UE-RIS paths interfere pairwise, and the angular-domain IRF exhibits $(L_g+L_f)(L_g+L_f-1)+1$ peaks, as shown in Fig. 4. Similarly, these peaks also contain the information of $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{f}}$. It is worth noting that, the interference



(a) Spatial domain. (b) Angular domain.

Fig. 4. Normalized IRF pattern in the spatial (angular) domain. Notably, the zero-frequency peak of the angular-domain IRF is intentionally ignored.

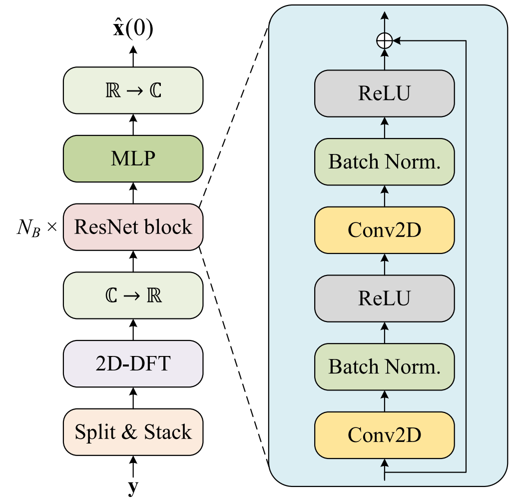


Fig. 5. Specific network structure of the designed initializer.

among different paths causes the accumulation of the zero-frequency component from all path pairs. Consequently, the zero-frequency component is determined jointly by multiple path gains, which complicates the extraction of individual path gain information compared with the difference-frequency components. Moreover, the accumulated zero-frequency peak is typically much larger than the difference-frequency peaks, potentially obscuring the informative peaks that contain angular information. Therefore, the zero-frequency component is intentionally ignored in the proposed channel estimation framework.

Then, we consider extending the above spectrum analysis to the sparse-sampling sensing RIS architecture. As known to us, sampling a sequence in the observation domain is equivalent to periodically repeating the spectrum in the transform domain. Consequently, the regular sampling architecture induces uncertainty of the angular coordinates, resulting in multiple globally optimal solutions that severely impede algorithmic convergence. Regarding the irregular sampling architecture, it preserves globally optimal solution uniqueness and we can still extract peaks from the spectrum, while at the cost of deteriorated spectral signal-to-noise ratio (SNR). Therefore, subsequent simulation experiments will exclusively concentrate on the irregular sampling architecture.

C. Deep Learning Based Model Design

In the previous Subsection IV-A, it is presented that the traditional GAMP algorithm exhibits two inherent limitations: its convergence highly depends on the selection of initial values, and the damping factor is uniformly fixed and remains constant throughout iterations.

As discussed in Subsection IV-B, the IRF spectrum pattern is highly connected to the sparse channel vector $\mathbf{x}$. Thus, we aim to strategically leverage the insight to well-resolve the former drawback. Specifically, the IRF signal demonstrates sparsity in the angular domain, with its non-zero indices strongly coupled to the support set of $\mathbf{x}$, which can be represented by

$$\mathcal{M}: \ \mathrm{supp}(\hat{\mathbf{P}}) \rightarrow \mathrm{supp}(\mathbf{x}), \tag{44}$$

where $\mathrm{supp}(\cdot)$ denotes the support set of a sparse vector or matrix. Therefore, it is feasible to exploit the properties of $\mathbf{x}$ and IRF in the angular domain, to provide effective initialization and enhance convergence for the GAMP algorithm. However, the aforementioned mapping $\mathcal{M}$ lacks exact mathematical characterization in multipath scenarios, particularly when the number of paths is too large. To overcome this challenge, deep learning techniques can be leveraged to extract features and fit the intricate nonlinear mapping through data-driven approach. Hence, we utilize neural network to acquire an effective initialization for enhanced algorithmic convergence, with its specific structure depicted in Fig. 5.

The network of the designed initializer takes as input the IRF signal $\mathbf{y}$, which is constructed by concatenating $Q = 4$ pilot observations and undergoes an initial pre-processing step labeled as "Split & Stack". This operation reorganizes the input into a format suitable for further transformation. Subsequently, to capture angular domain characteristics and fit the mapping $\mathcal{M}$, we apply 2D-DFT to transform the spatial observations into the angular domain. To facilitate compatibility with conventional real-valued neural network architectures, the complex-valued output of the 2D-DFT is decomposed into separate real and imaginary components, which are regarded as different channels. Further, the resulting real-valued signal is passed through N_B cascaded residual blocks to extract multi-scale features. Each residual block adopts the ResNet architecture [47] and comprises two convolutional layers, followed by batch normalization and ReLU activation function, without incurring significant performance loss while reducing the number of parameters. After the sequence of residual blocks, the multi-scale feature maps are processed via a multilayer perceptron (MLP), which serves as a global feature integrator. The MLP output is subsequently converted back to the complex domain, and the complex-valued representation $\hat{\mathbf{x}}(0)$ is regarded as an efficient initial value of the GAMP algorithm.

In order to address the latter limitation, the classical GAMP algorithm is deep unfolded into a neural network architecture,

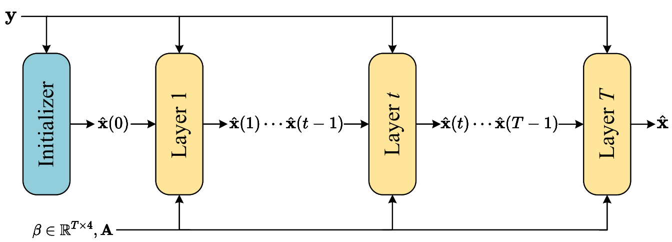


Fig. 6. Overall structure of the proposed learned GAMP network.

enabling a data-driven approach to parameter optimization. The damping factors, traditionally regarded as empirically chosen hyperparameters, are redefined as trainable parameters within this framework. This configuration enables the model to learn optimal damping factors directly from data during the training process, thereby enhancing adaptability and convergence performance. In particular, the damping factors are not uniform across layers, such that each layer maintains its own distinct set of damping factors. Additionally, within each unrolled layer, distinct damping factors are assigned to different variable updates, establishing an update-specific modulation mechanism. This layer-wise and update-specific design allows for a more flexible and fine-grained control of the inference dynamics, leading to improved reconstruction accuracy and stability compared to the conventional GAMP algorithm.

To construct a fully learnable GAMP architecture, the initializer is integrated into the deep unfolded GAMP neural network, resulting in a unified framework wherein both components are jointly optimized, which is shown as Fig. 6. The loss function is defined in (15), enabling end-to-end training of all parameters through gradient-based optimization. This cohesive design allows the model to adaptively learn both the initial estimates and the iterative refinement process in a data-driven manner. Consequently, the proposed learnable GAMP algorithm achieves notable improvements in both convergence and reconstruction accuracy compared to the conventional GAMP algorithm that relies on random initializations and empirical parameter settings.

D. Computational Complexity

Since the training process is conducted offline and does not affect the online estimation complexity, the focus of complexity analysis is primarily on the online estimation. Considering that the trainable damping factors do not contribute to the algorithmic complexity, the computational complexity of the deep unfolding part of the proposed learned GAMP algorithm remains equivalent to that of the traditional GAMP algorithm, i.e., $\mathcal{O}(TN_sN)$, where T denotes the number of iterations or layers. The computational complexity of the additional initialization network is mainly attributed to the fully connected layer in the MLP module used for features integration, with a complexity of $\mathcal{O}(N^2)$, and to the 2D-DFT operation, the complexity of which is reduced to $\mathcal{O}(N\log N)$ through the Fast Fourier Transform (FFT). Moreover, the computational complexity of the N_B cascaded ResNet blocks is dominated by the convolutional layers, and can be given by $\mathcal{O}(N_BC_{in}C_{out}K^2N)$, where C_{in} (C_{out}) denotes the number of the input (output) channels and K represents the size of the convolution kernel. Therefore, the overall computational complexity of the learned GAMP algorithm can be expressed as $\mathcal{O}(TN_sN)+\mathcal{O}(N^2)+\mathcal{O}(N\log N)+\mathcal{O}(N_BC_{in}C_{out}K^2N)$.

TABLE I

Simulation Parameters

Parameter	Value
RIS elements $N=N_1\times N_2$	$512=32\times 16$
BS transmitted power P_{BS}	1 W
UE transmitted power P_{UE}	300 mW
number of BS-RIS paths L_g	4
number of UE-RIS paths L_f	4
carrier frequency f_c	3.5 GHz
subcarrier bandwidth BW	180 kHz
power sensor noise factor F_p	$21\sim 41$ dB
thermal noise level n_0	-174 dBm/Hz
power sensor noise ν^w	$100\text{MHz}\times n_0F_p$
receiver thermal noise ν^v	$\text{BW}\times n_0$

V. Simulation Results and Discussion

In this section, we will present simulation results to evaluate the proposed channel estimators. Subsection V-A will introduce the detailed simulation setups. In Subsection V-B and Subsection V-C, CSI acquisition accuracy and achievable spectral efficiency will be shown to prove the feasibility and effectiveness of GAMP and learned GAMP based channel estimation schemes. The enhancement of the proposed learned GAMP in convergence will be exhibited in Subsection V-D.

A. Simulation Setup

In our simulation experiments, the system model described in Section II is considered, with its primary parameter configurations shown in Table I. For the Saleh-Valenzuela channel model in (1) and (2), the numbers of BS-RIS paths and UE-RIS paths are configured to $L_g=L_f=4$. The corresponding spatial angles of azimuth and elevation are assumed to be on-grid and randomly sampled from the grids. The complex channel gains $\alpha^g_{\ell_1}\sim\mathcal{CN}\left(0,\lambda^2/(4\pi R_{\text{BS}})^2\right),\forall\ell_1\in\{L_g\}$ and $\alpha^f_{\ell_2}\sim\mathcal{CN}\left(0,\lambda^2/(4\pi R_{\text{UE}})^2\right),\forall\ell_2\in\{L_f\}$, where $R_{\text{BS}}(R_{\text{UE}})$ represents the distance between the BS (UE) and the RIS, respectively. In this work, we configure $R_{\text{BS}}=50$ m and $R_{\text{UE}}=100$ m. To quantify the signal quality at the RIS and compare performances of different channel estimation schemes, we define the interferential SNR as

$$\bar{\gamma}\triangleq\left(\frac{P_{\text{BS}}\lambda^2}{(4\pi R_{\text{BS}})^2}+\frac{P_{\text{UE}}\lambda^2}{(4\pi R_{\text{UE}})^2}\right)\frac{1}{\nu^w}. \tag{45}$$

Additionally, we adopt the irregular sampling architecture, in which 50 percent of the RIS elements are equipped with power detectors to balance channel estimation precision and hardware overhead. Consequently, the energy consumption (mainly attributed to the power detectors in this work and proportional to their number) can be significantly reduced.

To train and evaluate the learned GAMP network, we construct a dataset comprising 10,000 training samples and 800 validation samples under the above experimental configuration, respectively. The entire network is trained in an unsupervised manner to minimize the loss function defined in (15). To improve training stability and avoid gradient vanishing in deeper layers, a layer-wise training strategy is employed, where the initializer is trained first, followed by adding one unfolded GAMP layer at a time and joint training. The network parameters are optimized utilizing the Adam optimizer, and the mini-batch size of per optimization is 256. The learning rate for the initializer optimization is configured as 10^{-4}, while 10^{-3} for the trainable damping factor updates. The number of training epochs is 1000 and we adopt early stopping with validation-based scheduling. Moreover, the implementation of the proposed learned GAMP network is based on the PyTorch library, and all simulations are performed on Geforce RTX 4090 24 GB GPUs.

For performance comparison, three baseline channel estimation schemes are considered:

1) SPARTA (SPARse Truncated Amplitude Flow): An effective truncated amplitude flow algorithm designed for sparse phase retrieval [48]. Compared with classical phase retrieval methods, SPARTA incorporates sparsity-aware initialization and truncated gradient iterations to improve recovery performance. Therefore, it serves as a strong baseline and we adapt it directly to the problem formulated in this work.

2) Phase-Oracle DnCNN (Denoising Convolutional Neural Network): A deep learning-based approach for sparse channel recovery [49]. The phase information of the IRF signals is assumed to be perfectly known. In this baseline, a DnCNN network is employed to denoise the angular-domain representations and identify the support set of the sparse channels. Subsequently, the least squares (LS) estimator is utilized to estimate the spatial-domain channels.

3) SuppPh-Oracle LS (Support Set and Phase Oracle): The support set of the sparse channels and the phase information of the IRF signals are assumed to be perfectly known. Then the LS estimation is performed to acquire the channel estimate.

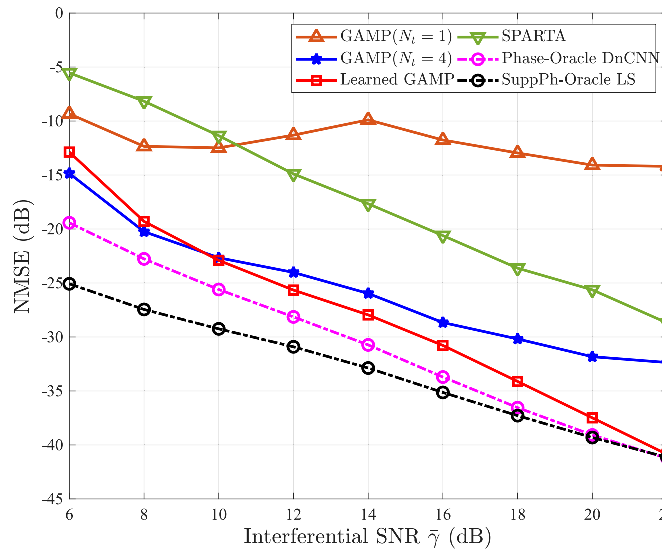


Fig. 7. NMSE of the BS-RIS channel estimators.

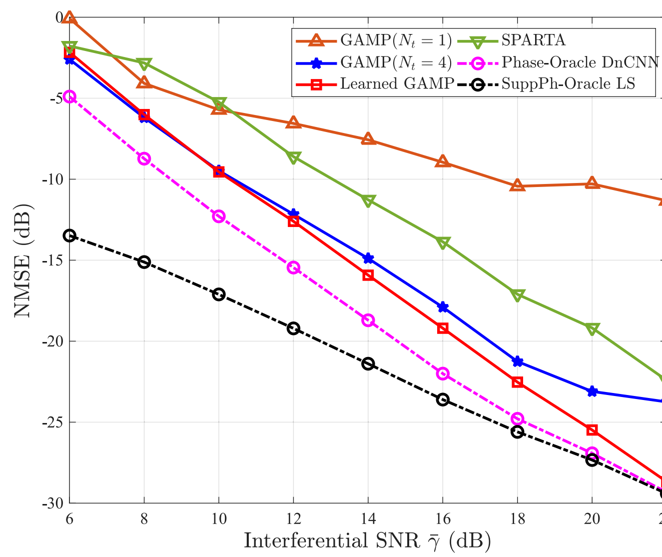


Fig. 8. NMSE of the UE-RIS channel estimators.

B. Accuracy of Acquired CSI

This subsection will evaluate the CSI acquisition accuracy of several channel estimation schemes. The normalized mean square error (NMSE) is utilized to quantify the channel estimation precision, which is represented as

$$\mathrm{NMSE}(\mathbf{h}) = \mathbb{E}\left[\frac{\|\hat{\mathbf{h}} - \mathbf{h}\|_2^2}{\|\mathbf{h}\|_2^2}\right], \tag{46}$$

where $\mathbf{h}$ denotes the ground-truth channel (the BS-RIS channel $\mathbf{g}$ or the UE-RIS channel $\mathbf{f}$), and $\hat{\mathbf{h}}$ is the estimated channel, respectively. Fig. 7 and Fig. 8 exhibit the NMSE performance curves against the interferential SNR $\bar{\gamma}$. In the simulation of the proposed GAMP-based method, the number of random initializations N_t is set to 1 and 4.

From the simulation results in Fig. 7 and Fig. 8, it is well verified that the proposed GAMP-based framework enables accurate separate channel estimation with merely 4 pilot symbols, regardless of the quantity of RIS elements, thus achieving dimension-independent separate channel estimation. As multiple initializations mitigate convergence to local minima, the GAMP-based channel estimation error exhibits a decreasing trend with increasing numbers of random initializations. It can also be observed that the GAMP-based method with $N_t = 4$ random initializations outperforms SPARTA across the entire interferential SNR range. Furthermore, the proposed learned GAMP algorithm achieves better performance than both GAMP and SPARTA over almost all interferential SNR levels. Compared with the "Phase-Oracle DnCNN" baseline, which assumes perfectly known phase information of the IRF signals, the learned GAMP algorithm exhibits only slight performance degradation, and approaches the "SuppPh-Oracle LS" baseline at high interferential SNR regions. Benefiting from a deep learning-based initialization network and learnable damping factors, the learned GAMP network effectively overcomes divergency and local minima, demonstrating the advantage of

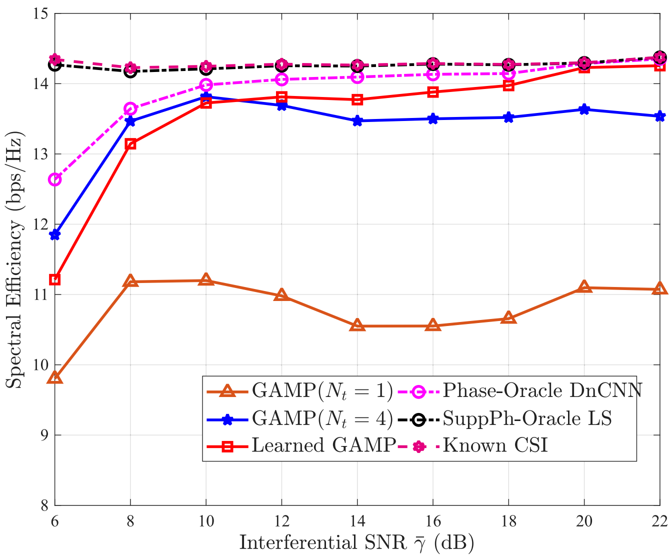


Fig. 9. Performance curve of spectral efficiency.

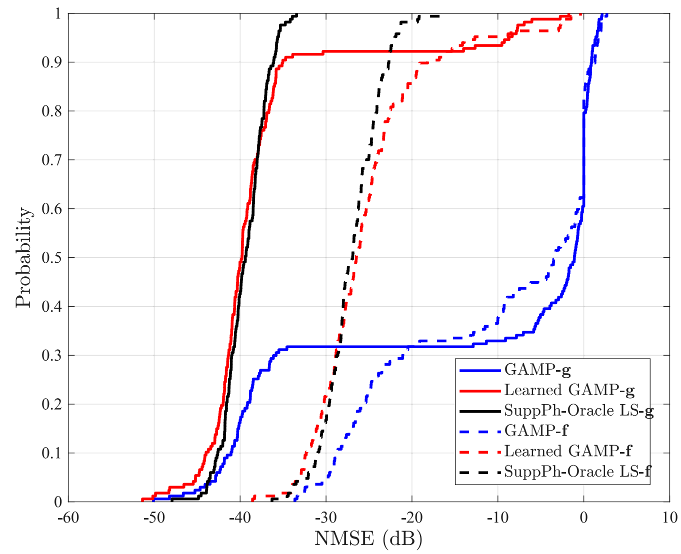


Fig. 10. CDF of NMSE. The interferential SNR $\bar{\gamma} = 20$ dB.

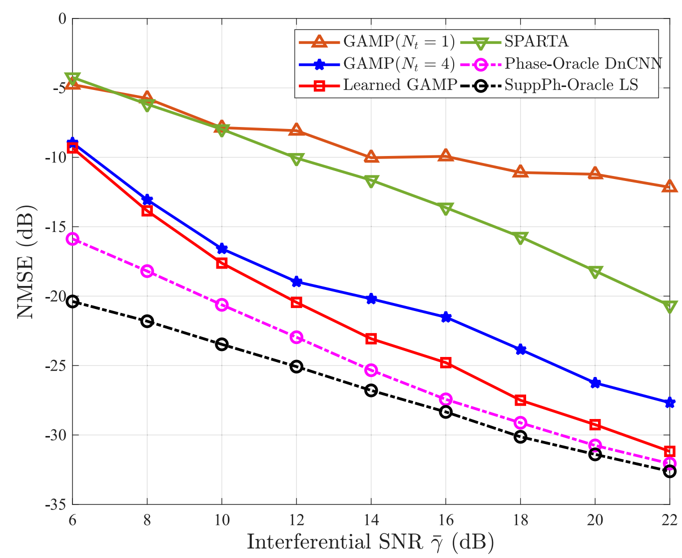


Fig. 11. NMSE performance of different BS-RIS channel estimation schemes under the 3GPP CDL-D channel model.

data-driven optimization. It is worth noting that, the NMSE of the GAMP algorithm does not decrease monotonically with increasing SNR, primarily because the algorithm exhibits an oscillatory convergence behavior against interferential SNR. This can be due to that, in moderate-to-high interferential SNR regions, the presence of non i.i.d Gaussian measurement matrix exacerbates the failure of "Onsager correction" term in GAMP algorithm, leading to frequent divergence [50], [51].

C. Achievable Spectral Efficiency

With the estimated CSI, the RIS phase-shift configuration $\mathbf{\Psi}$ can be calculated according to (9). Subsequently, achievable spectral efficiency is evaluated to quantify the influence of channel estimation accuracy on communication system performances, which can be represented as

$$R = \log_2\left(1 + \frac{P_{\mathrm{BS}}|\mathbf{f}^H\mathbf{\Psi}\mathbf{g}|^2}{\nu^v}\right). \tag{47}$$

As illustrated in Fig. 9, the spectral efficiency curves exhibit trends similar to the NMSE performances. At low-to-moderate interferential SNR regions, the channel estimation quality significantly impacts the achievable spectral efficiency. The proposed GAMP-based channel estimation method achieves an outstanding performance with a large number of random initializations, while at the cost of computational complexity and latency. Furthermore, the proposed learned GAMP scheme demonstrates excellent performance across a broad range of the interferential SNR, and is even close to the "SuppPh-Oracle LS" and "Known CSI" schemes at high SNR regions.

D. Convergence Comparison

The convergence of the learned GAMP algorithm is compared with the conventional GAMP using the cumulative distribution function (CDF) of NMSE, as shown in Fig. 10. The CDF curves are drawn through 400 independent trials under constant interferential SNR $\bar{\gamma} = 20$ dB. The learned GAMP network converges to a NMSE($\mathbf{g}$) below -30 dB in nearly 90 percent and a NMSE($\mathbf{f}$) below -20 dB in 85 percent of the trials, while the traditional GAMP achieves the same NMSE in only 30 percent of the trials. These results demonstrate the reliable convergence advantage of the learned GAMP algorithm. This improvement can be attributed to: 1) The learnable initializer provides an intial value close to the global optimum based on the intrinsic connection between sparse representations of the channels and the IRF in the angular domain; 2) layer-wise trainable damping factors adaptively regulate message passing and adjust variable update dynamics to avoid poor local minima. Therefore, the learned GAMP framework mitigates convergence to poor local minima and even divergency, making it a practical and scalable solution to real-time and high-accuracy channel estimation.

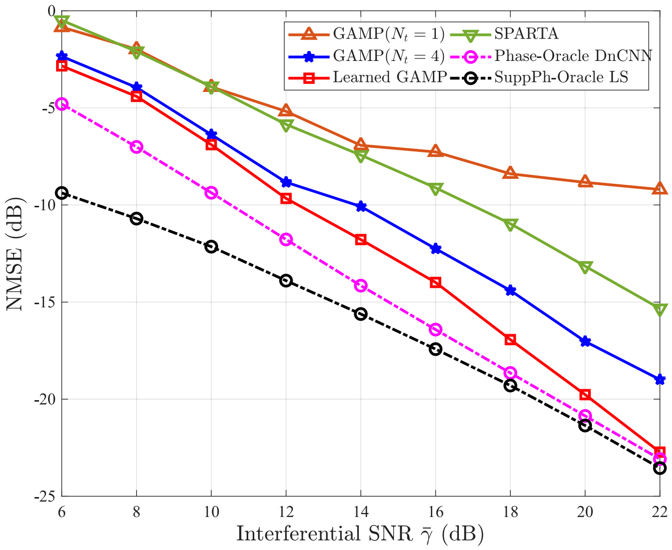


Fig. 12. NMSE performance of different UE-RIS channel estimation schemes under the 3GPP CDL-E channel model.

E. Other Simulation Results

In this subsection, we provide the channel estimation performance comparison of different schemes in practical communication systems. We conduct multi-subcarrier transmission and generate the 3rd Generation Partnership Project (3GPP) standard clustered delay line (CDL) channels via the MATLAB 5G Toolbox. The BS-RIS channel is generated using the 3GPP CDL-D channel model, while the UE-RIS channel is specified by the 3GPP CDL-E channel model. As illustrated in Fig. 11 and Fig. 12, the proposed GAMP-based method with $N_t = 4$ random initializations consistently outperforms SPARTA over the entire interferential SNR range. Moreover, the proposed learned GAMP algorithm still achieves the best performance among all non-oracle methods, and even approaches the "Phase-Oracle DnCNN" and "SuppPh-Oracle LS" schemes at high interferential SNR regions.

VI. Conclusion

In this study, the channel estimation problem in RIS assisted systems has been formulated as a compressive phase retrieval problem, and a novel channel estimation framework based on GAMP algorithm has been proposed. Unlike conventional schemes that estimate the high-dimensional cascaded channel with pilot overhead proportional to the number of RIS elements, the proposed method achieves dimension-independent estimation of separate BS-RIS and UE-RIS channels utilizing only four pilots. Furtherly, by exploiting the intrinsic connection between the sparse representations of the channels and the IRF in the angular domain, we have designed a learned GAMP network, which incorporates a neural initializer and layer-specific trainable damping factors, to enhance the conventional algorithm's convergence and estimation accuracy. Simulation results show that the GAMP-based channel estimation scheme achieves separate dimension-independent channel estimation. The learned GAMP algorithm significantly enhances conventional GAMP in convergence stability, iteration efficiency, and signal recovery precision.

Future work will focus on concretely investigating the impact of hardware impairments, including mutual coupling effect in rectangular arrays, power amplifier non-linearity and synchronization errors. Furthermore, it is promising to conduct prototype validation of the proposed framework based on our laboratory's hardware platform described in [52].

References

[1] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[2] K. Liu, Z. Zhang, L. Dai, and L. Hanzo, "Compact user-specific reconfigurable intelligent surfaces for uplink transmission," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 680–692, Jan. 2022.

[3] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[4] X. Wei, D. Shen, and L. Dai, "Channel estimation for RIS assisted wireless communications—Part I: Fundamentals, solutions, and future opportunities," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1398–1402, May 2021.

[5] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021.

[6] Y. Zhu, Y. Liu, Q. Wu, C. You, and Q. Shi, "Channel estimation by transmitting pilots from reconfigurable intelligent surface," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 3328–3343, Apr. 2024.

[7] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[8] J. Yang, Z. Lu, H. Yu, L. Wei, and C. Huang, "Holographic-inspired channel estimation for reconfigurable intelligent surface systems," *IEEE Wireless Commun. Lett.*, vol. 14, no. 12, pp. 4047–4051, Dec. 2025.

[9] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522–4535, Jul. 2020.

[10] A. Zappone, M. Di Renzo, X. Xi, and M. Debbah, "On the optimal number of reflecting elements for reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 464–468, Mar. 2021.

[11] N. K. Kundu, Z. Li, J. Rao, S. Shen, M. R. McKay, and R. Murch, "Optimal grouping strategy for reconfigurable intelligent surface assisted wireless communications," *IEEE Wireless Commun. Lett.*, vol. 11, no. 5, pp. 1082–1086, May 2022.

[12] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Process. Lett.*, vol. 27, pp. 905–909, 2020.

[13] T. Lin, X. Yu, Y. Zhu, and R. Schober, "Channel estimation for IRS-assisted millimeter-wave MIMO systems: Sparsity-inspired approaches," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4078–4092, Jun. 2022.

[14] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[15] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 210–214, Feb. 2020.

[16] H. Liu, X. Yuan, and Y.-J.-A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2621–2636, Nov. 2020.

[17] Y. Guo et al., "Efficient channel estimation for RIS-aided MIMO communications with unitary approximate message passing," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1403–1416, Feb. 2023.

[18] L. Zhou, J. Dai, and W. Xu, "Joint multi-user channel estimation for RIS-assisted massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 13993–14006, Oct. 2024.

[19] J. He, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmWave MIMO systems via atomic norm minimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5786–5797, Sep. 2021.

[20] Y. Chi and M. Ferreira Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Process. Mag.*, vol. 37, no. 2, pp. 39–57, Mar. 2020.

[21] Y. You, Y. Xue, L. Zhang, X. You, and C. Zhang, "Channel estimation for RIS assisted millimeter wave systems via OMP with optimization," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16783–16787, Dec. 2023.

[22] D. Dampahalage, K. B. Shashika Manosha, N. Rajatheva, and M. Latva-Aho, "Supervised learning based sparse channel estimation for RIS aided communications," in *Proc. ICASSP - IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2022, pp. 8827–8831.

[23] M. Jian and Y. Zhao, "A modified off-grid SBL channel estimation and transmission strategy for RIS-assisted wireless communication systems," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2020, pp. 1848–1853.

[24] X. Xu, S. Zhang, F. Gao, and J. Wang, "Sparse Bayesian learning based channel extrapolation for RIS assisted MIMO-OFDM," *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5498–5513, Aug. 2022.

[25] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607–6620, Oct. 2020.

[26] G. Zhou, C. Pan, H. Ren, P. Popovski, and A. L. Swindlehurst, "Channel estimation for RIS-aided multiuser millimeter-wave systems," *IEEE Trans. Signal Process.*, vol. 70, pp. 1478–1492, 2022.

[27] Z. Peng et al., "Channel estimation for RIS-aided multi-user mmWave systems with uniform planar arrays," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8105–8122, Dec. 2022.

[28] X. Wei, D. Shen, and L. Dai, "Channel estimation for RIS assisted wireless communications—Part II: An improved solution based on double-structured sparsity," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1403–1407, May 2021.

[29] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6853–6869, Oct. 2023.

[30] X. Pei et al., "RIS-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8627–8640, Dec. 2021.

[31] M. Cui, Z. Wu, Y. Chen, S. Xu, F. Yang, and L. Dai, "Demo: Low-power communications based on RIS and AI for 6G," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 1–2.

[32] F. Yang et al., "Reconfigurable reflectarrays and transmitarrays: From antenna designs to system applications," in *Proc. 12th Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2018, pp. 1–4.

[33] J. Zhu, K. Liu, Z. Wan, L. Dai, T. J. Cui, and H. V. Poor, "Sensing RISs: Enabling dimension-independent CSI acquisition for beamforming," *IEEE Trans. Inf. Theory*, vol. 69, no. 6, pp. 3795–3813, Jun. 2023.

[34] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, Jun. 2021.

[35] R. W. HeathJr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[36] Y. Guo, X. Guo, Y. Chen, and Y. Wang, "Statistical channel estimation for holographic MIMO exploiting the clustered sparsity," *IEEE Trans. Veh. Technol.*, vol. 74, no. 8, pp. 13161–13166, Aug. 2025.

[37] P. Hannan, "The element-gain paradox for a phased-array antenna," *IEEE Trans. Antennas Propag.*, vol. AP-12, no. 4, pp. 423–433, Jul. 1964.

[38] Y. Guo, X. Guo, and Y. Wang, "Mutual coupling-aware hybrid beamforming for densely packed MIMO metasurface," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 12630–12645, 2026.

[39] E.-H. Kim, J. Hahn, H. Kim, and B. Lee, "Profilometry without phase unwrapping using multi-frequency and four-step phase-shift sinusoidal fringe projection," *Opt. Exp.*, vol. 17, no. 10, pp. 7818–7830, May 2009.

[40] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2011, pp. 2168–2172.

[41] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

[42] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1043–1055, Feb. 2015.

[43] Y. Guo, X. Guo, Y. Wang, M. Di Renzo, and P. Zhang, "High SNR at a cost: The information-theoretic dilemma of magnitude-only atomic receivers," *IEEE Wireless Commun. Lett.*, vol. 14, no. 12, pp. 3997–4001, Dec. 2025.

[44] P. Schniter, "Turbo reconstruction of structured sparse signals," in *Proc. 44th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2010, pp. 1–6.

[45] S. Mukherjee and C. S. Seelamantula, "Fienup algorithm with sparsity constraints: Application to frequency-domain optical-coherence tomography," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4659–4672, Sep. 2014.

[46] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4814–4826, Sep. 2015.

[47] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[48] G. Wang, L. Zhang, G. B. Giannakis, M. Akçakaya, and J. Chen, "Sparse phase retrieval via truncated amplitude flow," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 479–491, Jan. 2018.

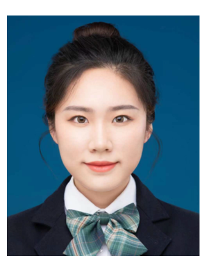
[49] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "RIS-aided mmWave MIMO channel estimation using deep learning and compressive sensing," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3503–3521, May 2023.

[50] J. Vila, P. Schniter, S. Rangan, F. Krzakala, and L. Zdeborová, "Adaptive damping and mean removal for the generalized approximate message passing algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 2021–2025.

[51] S. Rangan, P. Schniter, A. K. Fletcher, and S. Sarkar, "On the convergence of approximate message passing with arbitrary matrices," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5339–5351, Sep. 2019.

[52] J. Zhu, Z. Gu, Q. Ma, L. Dai, and T. J. Cui, "A self-controlled reconfigurable intelligent surface inspired by optical holography," *Nature Electron.*, vol. 8, no. 11, pp. 1108–1118, Oct. 2025.

Chao Jiang (Graduate Student Member, IEEE) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2024, where he is currently pursuing the M.S. degree with the Department of Electronic Engineering. His research interests include neural receivers and machine learning for wireless communications.



Tianyue Zheng (Graduate Student Member, IEEE) received the B.E. degree in information engineering from Southeast University, Nanjing, China, in 2022. She is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, Tsinghua University, Beijing, China. Her research interests include extremely large-scale MIMO (XL-MIMO), CSI acquisition, and AI for communications. She received the National Scholarship in 2019 and the Excellent Student of Jiangsu Province in 2021.

Jieao Zhu (Graduate Student Member, IEEE) received the B.E. degree in electronic engineering and the B.S. degree in applied mathematics from Tsinghua University, Beijing, China, in 2021, where he is currently pursuing the Ph.D. degree with the Department of Electronic Engineering. His research interests include electromagnetic information theory (EIT), reconfigurable intelligent surface (RIS), and Rydberg atomic receivers. He received the National Scholarship for undergraduates in 2018 and 2020 and for doctoral students in 2025. He was elected an Excellent Graduate of Beijing in 2021.

Shenheng Xu (Member, IEEE) received the B.S. degree in electronic information engineering and the M.S. degree in electromagnetic field and microwave technology from Southeast University, Nanjing, China, in 2001 and 2004, respectively, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles (UCLA), Los Angeles, CA, USA, in 2009.

From 2000 to 2004, he was a Research Assistant with the State Key Laboratory of Millimeter Waves, Southeast University. From 2004 to 2011, he was a Graduate Student Researcher and later a Post-Doctoral Researcher with the Antenna Research, Analysis, and Measurement Laboratory, UCLA. In 2012, he joined the Department of Electronic Engineering, Tsinghua University, Beijing, China, as an Associate Professor. His research interests include novel designs of high-gain antennas for advanced applications, artificial electromagnetic structures, and electromagnetic and antenna theories.

Linglong Dai (Fellow, IEEE) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2003, the M.S. degree (Hons.) from China Academy of Telecommunications Technology, Beijing, China, in 2006, and the Ph.D. degree (Hons.) from Tsinghua University, Beijing, in 2011.

From 2011 to 2013, he was a Post-Doctoral Research Fellow with the Department of Electronic Engineering, Tsinghua University, where he was an Assistant Professor from 2013 to 2016, an Associate Professor from 2016 to 2022, and has been a Professor since 2022. He has co-authored the book *MmWave Massive MIMO: A Paradigm for 5G* (Academic Press, 2016). He has authored or co-authored over 90 IEEE journal articles and over 50 IEEE conference papers. He also holds over 20 granted patents. His current research interests include massive MIMO, reconfigurable intelligent surface (RIS), millimeter-wave and terahertz communications, machine learning for wireless communications, near-field communications, and quantum information technologies. He has received five IEEE best paper awards at the IEEE ICC 2013, the IEEE ICC 2014, the IEEE ICC 2017, the IEEE VTC 2017-Fall, and the IEEE ICC 2018. He has also received the Tsinghua University Outstanding Ph.D. Graduate Award in 2011, Beijing Excellent Doctoral Dissertation Award in 2012, China National Excellent Doctoral Dissertation Nomination Award in 2013, the URSI Young Scientist Award in 2014, IEEE Transactions on Broadcasting Best Paper Award in 2015, the Electronics Letters Best Paper Award in 2016, the National Natural Science Foundation of China for Outstanding Young Scholars in 2017, the IEEE ComSoc Asia–Pacific Outstanding Young Researcher Award in 2017, the IEEE ComSoc Asia–Pacific Outstanding Paper Award in 2018, China Communications Best Paper Award in 2019, IEEE Access Best Multimedia Award in 2020, the IEEE Communications Society Leonard G. Abraham Prize in 2020, the IEEE ICC Outstanding Demo Award in 2022, the National Science Foundation for Distinguished Young Scholars in 2023, and the IEEE ComSoc Stephen O. Rice Prize in 2025. He was listed as a Highly Cited Researcher by Clarivate Analytics from 2020 to 2025.