Hybrid Near/Far-Field Channel Prediction for RIS-Aided LEO Satellite Networks

Jian Xiao, Ji Wang, *Senior Member, IEEE*, Xingwang Li, *Senior Member, IEEE*, Wenwu Xie, Nguyen Cong Luong, and Arumugam Nallanathan, *Fellow, IEEE*

***Abstract*—A hybrid near- and far- field cascaded channel prediction scheme is proposed for reconfigurable intelligent surface (RIS) assisted low earth orbit (LEO) satellite networks. In particular, an efficient neural network architecture, inspired by the intrinsic characteristics of wireless signals and termed the signal-informed network (SIN), is exploited to learn the precise mapping between historical uplink channels and future downlink channels. Specifically, in the proposed SIN, the time-domain autocorrelation modeling required by the channel prediction algorithm is converted into frequency-domain representation modeling, which aims to represent high-dimensional channels in terms of major frequency components. Furthermore, considering the specific non-linear phase information of hybrid-field channels, a multi-branch phase-aware module in SIN is developed to exhibit a physics-compliant channel semantic representation. Finally, a deep supervision-based encoder-decoder architecture with the auxiliary loss function is constructed as the network backbone. Simulation results demonstrate that compared to the state-of-art channel prediction models, the proposed SIN model exhibits superior channel prediction accuracy and convergence speed.**

***Index Terms*—Channel prediction, near-field communications, low earth orbit satellite, reconfigurable intelligent surface.**

I. Introduction

Due to severe signal propagation loss caused by long-distance satellite-to-ground communication links, conventional LEO satellites primarily support communication scenarios with stable line-of-sight (LOS) links [1]. However, in complex urban environments, the LOS transmission of the LEO satellite is probably obstructed by large buildings. To effectively enhance the signal coverage of LEO satellites, a new type of electromagnetic metasurface, i.e., reconfigurable intelligent surface (RIS), is expected to improve the service quality of LEO satellites at lower cost. In [2], an energy-efficient transmission scheme is proposed for stacked intelligent metasurface assisted LEO satellite systems.

In spite of the fact that various signal processing algorithms have been put forward for RIS enabled terrestrial networks [3], channel state information (CSI) acquisition in RIS-aided LEO satellite systems is highly challenging. Firstly, the passive RIS needs to integrate a large number of reflecting elements to improve passive beamforming gain, leading to high-dimensional cascaded channels. Secondly, the high mobility of LEO satellites introduces severe Doppler frequency shifts, exhibiting fast time-varying channels. Thirdly, the large propagation delay caused by the long-distance transmission link results in the outdated CSI. Moreover, as the number of RIS elements increases, the served users may be located in the near-field region of the RIS [4]. In near-field communications, the channel relates to both distances and angles, unlike the angle-only dependence in far-field cases [5], [6], making typical angular-domain sparsity-based far-field channel estimation and tracking algorithms unsuitable. Furthermore, due to the dynamic nature of terrestrial scatterers, a hybrid near- and far-field radiation scenario is suitable to RIS-aided LEO satellite networks [7], causing the complex overall channel distribution.

Recently, data-driven channel prediction algorithms have shown great potential, which are largely inspired by classic time-series forecast networks. In particular, the long short-term memory (LSTM) model and its derivative variants are popular network models [8]–[10]. However, the conventional LSTM-like models utilize a recursive forecast framework, inevitably leading to error propagation in multi-step channel predictions. In [11], a transformer-based parallel channel prediction model was proposed to avoid error propagation through the self-attention mechanism. Note that the computational complexity of the classical self-attention mechanism grows quadratically with the length of input, increasing the prediction complexity for the high-dimensional cascaded channel. Moreover, existing channel prediction models primarily consider far-field terrestrial communications with the small-scale RIS [8]–[11].

To address the hybrid-field cascaded channel acquisition for RIS-aided LEO satellite systems, we propose an efficient neural network architecture to precisely predict downlink multi-frame CSI from historical uplink CSI. We refer to the proposed channel prediction network as the signal-informed network (SIN), named for its architectural design that is tailored to the intrinsic structure and correlations of wireless signal. Specifically, in the proposed SIN, following Parseval's theorem of Fourier transform, we develop a frequency-domain modeling module to capture global periodicity and trends of

Jian Xiao and Ji Wang are with the Department of Electronics and Information Engineering, College of Physical Science and Technology, Central China Normal University, Wuhan 430079, China (e-mail: jianx@mails.ccnu.edu.cn; jiwang@ccnu.edu.cn).

Xingwang Li is with the School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China (e-mail: lixingwang@hpu.edu.cn).

Wenwu Xie is with the School of Information Science and Engineering, Hunan Institute of Science and Technology, Yueyang 414006, China (e-mail: gavinxie@hnist.edu.cn).

Nguyen Cong Luong is with the Faculty of Computer Science, Phenikaa University, Hanoi 12116, Vietnam (e-mail: luong.nguyencong@phenikaa-uni.edu.vn).

Arumugam Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, E1 4NS London, U.K., and also with the Department of Electronic Engineering, Kyung Hee University, Yongin-si 17104, South Korea (e-mail: a.nallanathan@qmul.ac.uk).

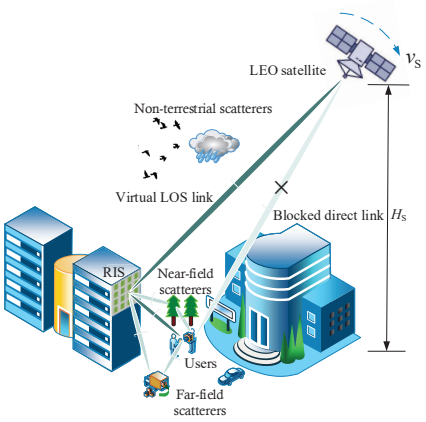

Fig. 1. RIS assisted LEO satellite communications.

historical CSI. Then, considering the non-linear phase and long-range dependency of hybrid-field channels, we exploit a multi-branch phase-aware module to decompose the feature tensor into amplitude and phase components. Moreover, we design a deep supervision strategy to enhance gradient flow in SIN. Numerical results demonstrate that the proposed SIN achieves superior channel prediction performance.

II. System Model and Problem Formulation

A. System Model

As illustrated in Fig. 1, we consider an RIS-aided LEO satellite system under frequency division duplex (FDD) mode in a typical urban environment, where the direct links between an LEO satellite and K single-antenna users are obstructed. The LEO satellite is equipped with $M = M_1 \times M_2$ uniform planar array (UPA) antennas. To enhance the signal coverage, an RIS configured with $N = N_1 \times N_2$ UPA reflecting elements is deployed to create the virtual LOS link. The heights of the LEO satellite, the RIS and users are H_S, H_R and H_U, respectively. The users are assumed to be located in a near-field region of the RIS. Following massive multiple-input multiple-output (MIMO) channel modeling in LEO satellite networks [12], the RIS→LEO satellite channel $\mathbf{G} \in \mathbb{C}^{M\times N}$ in time block t at uplink carrier frequency f^{u} can be expressed as

$$\mathbf{G}\left(f^{\mathrm{u}},t\right)=\sqrt{\frac{\kappa^{\mathrm{r}}L_t^{\mathrm{RS}}}{\kappa^{\mathrm{r}}+1}}e^{j\eta^{\mathrm{RS}}}\mathbf{b}(\phi_t^{\mathrm{S}},\varphi_t^{\mathrm{S}})\mathbf{a}^T(\phi_t^{\mathrm{RS}},\varphi_t^{\mathrm{RS}})e^{j2\pi\psi_t^{\mathrm{RS}}}+$$
$$\gamma^{\mathrm{RS}}\sum_{s=1}^{S^{\mathrm{RS}}}\beta_t^{\mathrm{RS},s}\sqrt{L_t^{\mathrm{RS},s}}\mathbf{b}(\phi_t^{\mathrm{S},s},\varphi_t^{\mathrm{S},s})\mathbf{a}^T(\phi_t^{\mathrm{RS},s},\varphi_t^{\mathrm{RS},s})e^{j2\pi\psi_t^{\mathrm{RS},s}}, \tag{1}$$

where L_t^{RS} denotes the basic path loss for the LOS link[1]. κ^{r} denotes the Rician factor. η^{RS} is the random initial phase for the LOS path, following the uniform distribution $\eta^{\mathrm{RS}} \sim \mathcal{U}[0, 2\pi]$. $\phi_t^{\mathrm{S}}(\varphi_t^{\mathrm{S}})$ and $\phi_t^{\mathrm{RS}}(\varphi_t^{\mathrm{RS}})$ denote azimuth (elevation) angles of arrival and departure, respectively. Parameter $\psi_t^{\mathrm{RS}} = tT_s\mathcal{D}_t^{\mathrm{RS}} - f^{\mathrm{u}}\tau_t^{\mathrm{RS}}$ signifies the fast time-varying and long time decay characteristics for the LOS link, in which T_s, τ_t^{RS} and $\mathcal{D}_t^{\mathrm{RS}}$ denote the sampling period, transmission delay, and Doppler frequency shift of LOS link, respectively. $\mathbf{a}(\phi_t^{\mathrm{RS}},\varphi_t^{\mathrm{RS}}) \in \mathbb{C}^{N\times 1}$ and $\mathbf{b}(\phi_t^{\mathrm{S}},\varphi_t^{\mathrm{S}}) \in \mathbb{C}^{M\times 1}$ represent the transmitting array response at the RIS and receiving array response at the LEO satellite, respectively. The far-field array response with respect to azimuth angle ϕ_t^i and elevation angle $\varphi_t^i (i \in \{\mathrm{S}, \mathrm{RS}\})$ can be expressed as

$$\mathbf{a}(\phi_t^i,\varphi_t^i)=\mathbf{a}_1(\phi_t^i,\varphi_t^i)\otimes\mathbf{a}_2(\varphi_t^i),$$
$$\mathbf{a}_1(\phi_t^i,\varphi_t^i)\triangleq\frac{1}{\sqrt{n_1}}\left[1,e^{j\varsigma\sin\phi_t^i\sin\varphi_t^i},\ldots,e^{j\varsigma(n_2-1)\sin\phi_t^i\sin\varphi_t^i}\right]^T,$$
$$\mathbf{a}_2\left(\varphi_t^i\right)\triangleq\frac{1}{\sqrt{n_2}}\left[1,e^{j\varsigma\cos\varphi_t^i},\ldots,e^{j\varsigma(n_2-1)\cos\varphi_t^i}\right]^T, \tag{2}$$

where $n_1 \in \{M_1, N_1\}$, $n_2 \in \{M_2, N_2\}$, $\varsigma = \frac{2\pi c}{f^{\mathrm{u}}}d$, c is the speed of light, and d denotes adjacent element spacing.

For the NLOS channel component in the RIS→LEO satellite link, S^{RS} represents the number of scatterers, and $\gamma^{\mathrm{RS}} = 1/\sqrt{(\kappa^{\mathrm{r}}+1)S^{\mathrm{RS}}}$ is a normalization factor. The complex gain $\beta_t^{\mathrm{RS},s}$ of scatterer path s is generated according to the exponential power delay profile (PDP)[2]. Due to space limitation, the other NLOS channel parameters for scatterer path s, e.g., $L_t^{\mathrm{RS},s}$, $\phi_t^{\mathrm{S},s}$, $\varphi_t^{\mathrm{S},s}$ and $\psi_t^{\mathrm{RS},s}$, are defined in accordance with those of the LOS link.

Considering S^{n} near-field and S^{f} far-field scatterers in the k-th user→RIS link, the hybrid-field channel $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ can be expressed as

$$\mathbf{h}_k\left(f^{\mathrm{u}},t\right)=\underbrace{\sqrt{\frac{\kappa_k^{\mathrm{r}}L_{k,t}^{\mathrm{UR}}}{\kappa_k^{\mathrm{r}}+1}}\eta^{\mathrm{UR}}\mathbf{c}(\phi_{k,t}^{\mathrm{UR}},\varphi_{k,t}^{\mathrm{UR}},d_{k,t}^{\mathrm{UR}})e^{j2\pi\psi_t^{\mathrm{UR}}}}+$$
$$\gamma^{\mathrm{UR}}\sum_{s=1}^{S^{\mathrm{UR}}}\beta_{k,t}^{\mathrm{UR},s}\sqrt{L_{k,t}^{\mathrm{UR},s}}\mathbf{c}(\phi_{k,t}^{\mathrm{UR},s},\varphi_{k,t}^{\mathrm{UR},s},d_{k,t}^{\mathrm{UR},s})e^{j2\pi\psi_t^{\mathrm{UR},s}}, \tag{3}$$

where $S^{\mathrm{UR}} = S^{\mathrm{n}}+S^{\mathrm{f}}$, $\gamma^{\mathrm{UR}} = 1/\sqrt{(\kappa_k^{\mathrm{r}}+1)S^{\mathrm{UR}}}$, $d_{k,t}^{\mathrm{UR}}$ denotes the distance from the k-th user to the center of RIS, and κ_k^{r} denotes the Rician factor for k-th user→RIS link. Similar to (1), $\psi_t^{\mathrm{UR}} = tT_s\mathcal{D}_{k,t}^{\mathrm{UR}} - f^{\mathrm{u}}\tau_{k,t}^{\mathrm{UR}}$ and $\psi_t^{\mathrm{UR},s} = tT_s\mathcal{D}_{k,t}^{\mathrm{UR},s} - f^{\mathrm{u}}\tau_{k,t}^{\mathrm{UR},s}$, while parameters $L_{k,t}^{\mathrm{UR}}(L_{k,t}^{\mathrm{UR},s})$, $\phi_{k,t}^{\mathrm{UR}}(\phi_{k,t}^{\mathrm{UR},s})$, $\varphi_{k,t}^{\mathrm{UR}}(\varphi_{k,t}^{\mathrm{UR},s})$, $\mathcal{D}_{k,t}^{\mathrm{UR}}(\mathcal{D}_{k,t}^{\mathrm{UR},s})$, and $\tau_{k,t}^{\mathrm{UR}}(\tau_{k,t}^{\mathrm{UR},s})$ denote the basic path loss, initial phase, complex gain, azimuth angle, elevation angle, Doppler frequency shift and transmission delay for LOS link (scatterer path s), respectively. $\mathbf{c}(\phi_{k,t}^{\mathrm{UR}},\varphi_{k,t}^{\mathrm{UR}},d_{k,t}^{\mathrm{UR}})$ denotes the near-field receiving array response at the RIS for the LOS link, which is given by

$$\mathbf{c}(\phi_{k,t}^{\mathrm{UR}},\varphi_{k,t}^{\mathrm{UR}},d_{k,t}^{\mathrm{UR}})=\mathbf{c}_1(\phi_{k,t}^{\mathrm{UR}},\varphi_{k,t}^{\mathrm{UR}},d_{k,t}^{\mathrm{UR}})\otimes\mathbf{c}_2(\phi_{k,t}^{\mathrm{UR}},\varphi_{k,t}^{\mathrm{UR}},d_{k,t}^{\mathrm{UR}}),$$
$$[\mathbf{c}_1]_{n_1}\triangleq e^{-j\frac{2\pi f^{\mathrm{u}}}{c}\left(-\bar{n}_1 d\cos\phi_{k,t}^{\mathrm{UR}}\sin\varphi_{k,t}^{\mathrm{UR}}+\frac{\bar{n}_1{}^2d^2\left(1-\left(\cos\phi_{k,t}^{\mathrm{UR}}\sin\varphi_{k,t}^{\mathrm{UR}}\right)^2\right)}{2d^{\mathrm{UR}}}\right)},$$
$$[\mathbf{c}_2]_{n_2}\triangleq e^{-j\frac{2\pi f^{\mathrm{u}}}{c}\left(-\bar{n}_2 d\cos\varphi_{k,t}^{\mathrm{UR}}+\frac{\bar{n}_2{}^2d^2\sin^2\varphi_{k,t}^{\mathrm{UR}}}{2d^{\mathrm{UR}}}\right)}, \tag{4}$$

where $\bar{n}_1 = n_1 - \frac{N_1+1}{2}$ and $\bar{n}_2 = n_2 - \frac{N_2+1}{2}$.

[1]According to 3GPP standard for non-terrestrial networks [13], the basic path loss in dB unit is given by $L_t^{\mathrm{RS}} = L_t^{\mathrm{FS}} + L_t^{\mathrm{CL}} + L_t^{\mathrm{SF}}$, which is composed of free space path loss L_t^{FS}, cluster loss L_t^{CL} and shadow fading L_t^{SF}.

[2]In general, the channel path gain $\beta_t^{\mathrm{RS},s}$ can be characterized by the complex Gaussian distribution [8], [12], i.e., $\beta_t^{\mathrm{RS},s} \sim \mathcal{CN}(0, \sigma_t^{\mathrm{RS},s})$ with $\sum_{s=0}^{S-1}\sigma_t^{\mathrm{RS},s} = 1$. Here, $\sigma_t^{\mathrm{RS},s}$ is referred to as the PDP and is calculated with the single slope exponential PDP.

For the NLOS component in (3), the definition of array response $\mathbf{c}(\phi_{k,t}^{\mathrm{UR},s}, \varphi_{k,t}^{\mathrm{UR},s}, d_{k,t}^{\mathrm{UR},s})$ relies on the distance $d_{k,t}^{\mathrm{UR},s}$ from scatterer s to the RIS. If $d_{k,t}^{\mathrm{UR},s} > R$, the expression of $\mathbf{c}(\phi_{k,t}^{\mathrm{UR},s}, \varphi_{k,t}^{\mathrm{UR},s}, d_{k,t}^{\mathrm{UR},s})$ in (4) will be simplified into $\mathbf{c}(\phi_{k,t}^{\mathrm{UR},s}, \varphi_{k,t}^{\mathrm{UR},s})$, following the far-field definition in (2). On the contrary, if $d_{k,t}^{\mathrm{UR},s} \leq R$, $\mathbf{c}(\phi_{k,t}^{\mathrm{UR},s}, \varphi_{k,t}^{\mathrm{UR},s}, d_{k,t}^{\mathrm{UR},s})$ will be expressed as (4). The near-field array response is not only related to angles $\phi_{k,t}^{\mathrm{UR},s}$ and $\varphi_{k,t}^{\mathrm{UR},s}$ but also relied on distance $d_{k,t}^{\mathrm{UR},s}$. In particular, the phase of far-field array response is a linear function, while the near-field array response is nonlinear due to the presence of the second-order terms in (4).

Let $\boldsymbol{\theta} = [e^{j\theta_1}, e^{j\theta_2}, \cdots, e^{j\theta_N}]^T \in \mathbb{C}^{N\times 1}$ represent the phase shift vector of the RIS. The received pilot signal $\mathbf{y}_q \in \mathbb{C}^{M\times 1}$ in the q-th slot at the LEO satellite is given by

$$\mathbf{y}_q = \sum_{k=1}^{K} \mathbf{G}\mathrm{diag}(\boldsymbol{\theta}_q)\mathbf{h}_k s_{k,q} + \mathbf{w}_{q,k}, \tag{5}$$

where $\mathrm{diag}(\cdot)$ denotes the vector diagonalization operation, $s_{k,q}$ is the transmit signal of the k-th user, and $\mathbf{w}_q \sim \mathcal{CN}(0, \sigma_n^2\mathbf{I}_M)$ is the Gaussian noise with variance σ_n^2. We denote $\mathbf{H}_k = \mathbf{G}\mathrm{diag}(\mathbf{h}_k) \in \mathbb{C}^{M\times N}$ as the cascaded channel.

B. Problem Formulation

In this work, we focus on the downlink channel prediction of multiple frames according to the historical uplink CSI. Without loss of generality, we adopt the minimum variance unbiased estimator estimator with multi-user orthogonal pilots to acquire the uplink CSI due to its simplification and popularity. Note that the proposed channel prediction scheme is applicable to existing low-overhead channel estimation algorithms [14]. Let $\widetilde{\mathbf{Y}}_k \in \mathbb{C}^{M\times Q}$ denote received pilot signal for the k-th user after Q slots, the estimated uplink CSI is given by [15]

$$\widehat{\mathbf{H}}_k^{\mathrm{u}} = \widetilde{\mathbf{Y}}_k \boldsymbol{\Theta}^{\mathbf{H}} \left(\boldsymbol{\Theta}\boldsymbol{\Theta}^{\mathbf{H}}\right)^{-1}, \tag{6}$$

where $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_Q] \in \mathbb{C}^{N\times Q}$, and the discrete Fourier transform (DFT)-based reflection protocol is adopted to design $\boldsymbol{\Theta}$, i.e., the (p_1, p_2)-th element of $\boldsymbol{\Theta}$ satisfies $\boldsymbol{\Theta}_{p_1,p_2} = e^{-j(2\pi(p_1-1)(p_2-1)/Q)}, (1 \leq p_1, p_2 \leq Q)$.

According to the estimated historical cascaded channel $\widehat{\mathbf{H}}_k^{\mathrm{u}} = [\widehat{\mathbf{H}}_{k,1}^{\mathrm{u}}, \widehat{\mathbf{H}}_{k,2}^{\mathrm{u}}, \cdots, \widehat{\mathbf{H}}_{k,P^{\mathrm{u}}}^{\mathrm{u}}] \in \mathbb{C}^{M\times N\times P^{\mathrm{u}}}$ from the past P^{u} frames, we construct a channel prediction framework to realize the precise prediction of downlink cascaded channel $\mathbf{H}_k^{\mathrm{d}} = [\mathbf{H}_{k,1}^{\mathrm{d}}, \mathbf{H}_{k,2}^{\mathrm{d}}, \cdots, \mathbf{H}_{k,P^{\mathrm{d}}}^{\mathrm{d}}] \in \mathbb{C}^{M\times N\times P^{\mathrm{d}}}$ in the next P^{d} consecutive frames. In the neural network, the input tensor is designed as $\overline{\mathbf{H}}_k^{\mathrm{u}} = [\Re(\widehat{\mathbf{H}}_k^{\mathrm{u}}), \Im(\widehat{\mathbf{H}}_k^{\mathrm{u}})] \in \mathbb{C}^{M\times N\times 2P^{\mathrm{u}}}$, and the data label is $\overline{\mathbf{H}}_k^{\mathrm{d}} = [\Re(\mathbf{H}_k^{\mathrm{d}})\Im(\mathbf{H}_k^{\mathrm{d}})] \in \mathbb{C}^{M\times N\times 2P^{\mathrm{d}}}$, where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of a complex number, respectively. Let $\widehat{\mathbf{H}}_k^{\mathrm{d}} \in \mathbb{C}^{M\times N\times 2P^{\mathrm{d}}}$ denote the output of the channel prediction network model, and ℓ_1 norm-based loss function is defined as

$$\mathcal{L}\left(\overline{\mathbf{H}}_k^{\mathrm{d}}, \widehat{\mathbf{H}}_k^{\mathrm{d}}\right) = \frac{1}{MNP^{\mathrm{d}}} \sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{p=1}^{P^{\mathrm{d}}} \left\| \overline{\mathbf{H}}_{k,m,n,p}^{\mathrm{d}} - \widehat{\mathbf{H}}_{k,m,n,p}^{\mathrm{d}} \right\|. \tag{7}$$

Finally, we carry out the real-to-complex operation to obtain the prediction channel $\widetilde{\mathbf{H}}_{k,m,n,p}^{\mathrm{d}} = \widehat{\mathbf{H}}_{k,m,n,p}^{\mathrm{d}} + j\widehat{\mathbf{H}}_{k,m,n,p+P^{\mathrm{d}}}^{\mathrm{d}}$.

III. Proposed Channel Prediction Network

The proposed channel prediction network, i.e., the SIN architecture, is composed of the frequency-domain multilayer perceptron (MLP) module and the multi-branch phase-aware module, aiming to construct the mapping from $\overline{\mathbf{H}}_k^{\mathrm{u}}$ to $\overline{\mathbf{H}}_k^{\mathrm{d}}$.

A. Frequency-Domain MLP Module

The existing channel prediction works commonly carry out a series of dedicated operations in the time domain of the historical CSI, focusing on the variations in amplitude over time. In this work, we exploit a frequency-domain MLP module to capture the global dependencies of the cascaded channel, transforming time-domain autocorrelation modeling into frequency-domain representation[3]. The Parseval's theorem states that the total energy of a signal can be measured equivalently in time or frequency domains as [16]

$$\int_{-\infty}^{\infty} |\mathbf{H}(t)|^2 \mathrm{d}t = \int_{-\infty}^{\infty} |\mathcal{H}(f)|^2 \mathrm{d}f, \tag{8}$$

where $\mathbf{H}(t)$ and $\mathcal{H}(f)$ denote the time-domain and frequency-domain tensors, respectively. We observe that $\mathbf{H}(t)$ can be represented by utilizing the frequency components $\mathcal{H}(f)$ that concentrates the majority of energy, especially for the cascaded channel with limited scatterers. In the proposed SIN, the DFT in the time domain is employed to convert input tensor $\overline{\mathbf{H}}_k^{\mathrm{u}}$ as

$$\overline{\mathcal{H}}^{\mathrm{T}}[v] = \mathcal{F}^{\mathrm{T}}(\overline{\mathbf{H}}_k^{\mathrm{u}}) = \sum_{p=0}^{P^{\mathrm{u}}-1} \overline{H}_k^{\mathrm{u}}[p] e^{-j2\pi\frac{vp}{P^{\mathrm{u}}}}. \tag{9}$$

According to the conjugate symmetry property of DFT, only half of $\overline{\mathcal{H}}^{\mathrm{T}} \in \mathbb{C}^{M\times N\times P^{\mathrm{u}}}$ is sufficient to represent overall frequency domain information. Then, a specialized frequency domain MLP module is designed, where the output feature of the l-th layer can be expressed as

$$\mathcal{Y}^l = \sigma^l\left(\mathcal{Y}^{l-1}\mathcal{W}^l + \mathcal{B}^l\right), \tag{10}$$

where $\mathcal{Y}^0 = \overline{\mathcal{H}}^{\mathrm{T}}$ for $l = 0$. $\sigma^l(\cdot)$, $\mathcal{W}^l = \overline{\mathbf{W}}_r^l + j\overline{\mathbf{W}}_i^l$ and $\mathcal{B}^l = \overline{\mathbf{b}}_r^l + j\overline{\mathbf{b}}_i^l$ denote the activation function, complex-valued weights and biases of the l-th layer in MLP, respectively. Following the principle of complex-value multiplication, (10) can be rewritten as

$$\begin{aligned}\mathcal{Y}^l =& \sigma^l\left(\Re\left(\mathcal{Y}^{l-1}\right)\overline{\mathbf{W}}_r^l - \Im\left(\mathcal{Y}^{l-1}\right)\overline{\mathbf{W}}_i^l + \overline{\mathbf{b}}_r^l\right) \\ &+ j\sigma^l\left(\Re\left(\mathcal{Y}^{l-1}\right)\overline{\mathbf{W}}_i^l + \Im\left(\mathcal{Y}^{l-1}\right)\overline{\mathbf{W}}_r^l + \overline{\mathbf{b}}^l\right).\end{aligned} \tag{11}$$

Finally, the inverse discrete Fourier transform (IDFT) is employed to restore the frequency domain characteristics to the time domain, and a feature ascending dimension operation is carried out that projects $\mathbf{F}$ into $\mathbf{F}^{\mathrm{i}} \in \mathbb{R}^{M\times N\times C}$.

[3]This paper considers the narrowband temporal-spatial channel prediction for MIMO systems, where the concept of domain transformation pertains to feature tensor operations within the neural network, rather than to the time-frequency domain channels in multi-carrier systems. In other words, the frequency-domain representation is defined as the feature tensor obtained by performing DFT on the input tensor.

Algorithm 1 Multi-branch phase-aware module

Input: Feature tensor $\mathbf{F}^{\mathrm{i}} \in \mathbb{R}^{M\times N\times C}$, Splitting factor $\mathcal{S} = C/C^s$
Output: Feature tensor $\mathbf{F}^{\mathrm{o}} \in \mathbb{R}^{M\times N\times C}$

Segment splitting:
  $\mathbf{F}^{\mathrm{s}} = [\mathbf{F}^{\mathrm{i}}_1, \mathbf{F}^{\mathrm{i}}_2, \ldots, \mathbf{F}^{\mathrm{i}}_{\mathcal{S}}] \in \mathbb{R}^{M\times N\times C^s\mathcal{S}}$
  satisfying $\mathbf{F}_s \in \mathbb{R}^{H\times W\times D}(1 \le s \le \mathcal{S})$
Height-Channel permutation in the first branch:
  Rearrange $\mathbf{F}^{\mathrm{s}}$ as $\mathbf{F}^{\mathrm{s}}_1 \in \mathbb{R}^{C^s\times N\times M\mathcal{S}}$
  Pass the phase-aware module $\mathbf{F}^{\mathrm{p}}_1 = \delta(\mathbf{F}^{\mathrm{s}}_1) \in \mathbb{R}^{C^s\times N\times M\mathcal{S}}$
  Rearrange $\mathbf{F}^{\mathrm{p}}_1$ as $\mathbf{F}^{\mathrm{o}}_1 \in \mathbb{R}^{M\times N\times C^s\mathcal{S}}$
Width-Channel permutation in the second branch:
  Rearrange $\mathbf{F}^{\mathrm{s}}$ as $\mathbf{F}^{\mathrm{s}}_2 \in \mathbb{R}^{C^s\times M\times N\mathcal{S}}$
  Pass the phase-aware module $\mathbf{F}^{\mathrm{p}}_2 = \delta(\mathbf{F}^{\mathrm{s}}_2) \in \mathbb{R}^{C^s\times M\times N\mathcal{S}}$
  Rearrange $\mathbf{F}^{\mathrm{p}}_2$ as $\mathbf{F}^{\mathrm{o}}_2 \in \mathbb{R}^{M\times N\times C^s\mathcal{S}}$
Linear mapping in the third branch: $\mathbf{F}^{\mathrm{o}}_3 = \mathbf{F}^{\mathrm{i}}\mathbf{W}^{\mathrm{c}} + \mathbf{b}^{\mathrm{c}}$
Spatial DFT in the fourth branch:
  2D DFT: $\mathbf{F}^{\mathrm{s}}_4[u,v] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathbf{F}^{\mathrm{i}}[m,n]e^{-j2\pi(\frac{um}{2M}+\frac{vn}{N})}$
  Modulate spectrum with learnable filter $\mathbf{\Psi}$: $\mathbf{F}^{\mathrm{p}}_4 = \mathbf{\Psi}\odot\mathbf{F}^{\mathrm{s}}_4$
  2D IDFT in the spatial domain: $\mathbf{F}^{\mathrm{o}}_4 = \mathcal{F}^{\mathrm{S},-1}\left(\mathbf{F}^{\mathrm{p}}_4\right)$
Multi-branch feature aggregation:
  Aggregation with adaptive weights $\boldsymbol{\alpha}_l$: $\mathbf{F}^{\mathrm{o}}_5 = \sum_{l=1}^{3}\mathbf{F}^{\mathrm{o}}_l \odot \boldsymbol{\alpha}_l$
  Aggregation with feature concatenation: $\mathbf{F}^{\mathrm{o}}_6 = [\mathbf{F}^{\mathrm{o}}_4, \mathbf{F}^{\mathrm{o}}_5]$
  Aggregation with skip connection: $\widehat{\mathbf{F}}^{\mathrm{o}} = \mathbf{F}^{\mathrm{o}}_6\mathbf{W}^{\mathrm{o}} + \mathbf{F}^{\mathrm{i}}$

B. Multi-Branch Phase-Aware Module

It is well known that a general physic form of wireless signal/channel vector can be represented as $\mathbf{z} = |\mathbf{z}| \odot e^{j\boldsymbol{\vartheta}}$, where $|\mathbf{z}|$ and $\boldsymbol{\vartheta}$ denote amplitude and phase of $\mathbf{z}$, respectively. According to Euler's formula, it can be unfolded as

$$\mathbf{z} = |\mathbf{z}| \odot \cos\boldsymbol{\vartheta} + j|\mathbf{z}| \odot \sin\boldsymbol{\vartheta}. \tag{12}$$

In the proposed phase-aware module, given an input tensor $\bar{\mathbf{z}} \in \mathbb{R}^{M}$, its amplitude tensor $|\bar{\mathbf{z}}|$ and phase tensor $\bar{\boldsymbol{\vartheta}}$ are constructed by utilizing linear layers, i.e., $|\bar{\mathbf{z}}| = \bar{\mathbf{z}}\overline{\mathbf{W}}^{\mathrm{a}} + \bar{\mathbf{b}}^{\mathrm{a}}$ and $\bar{\boldsymbol{\vartheta}} = \bar{\mathbf{z}}\overline{\mathbf{W}}^{\mathrm{p}} + \bar{\mathbf{b}}^{\mathrm{p}}$, where $\overline{\mathbf{W}}^{i} \in \mathbb{R}^{M\times M}$ and $\bar{\mathbf{b}}^{i} \in \mathbb{R}^{M}(i \in \{\mathrm{a},\mathrm{p}\})$ are the weights and bias of linear layers, respectively. To preserve diverse information, we utilize the concatenate operation to replace the summation operation in (12), i.e., $\bar{\mathbf{z}}^{\mathrm{c}} = [|\bar{\mathbf{z}}| \odot \cos\bar{\boldsymbol{\vartheta}}, j|\bar{\mathbf{z}}| \odot \sin\bar{\boldsymbol{\vartheta}}] \in \mathbb{R}^{2M}$. By this way, the feature tensor $\mathbf{F}^{\mathrm{i}}$ is decomposed into the amplitude and phase parts.

To characterize the spatial correlations of the cascaded channel, we further propose a multi-branch phase-aware module. As illustrated in **Algorithm 1**, the multi-branch phase-aware module comprises four feature extraction branches along different dimensions of the input tensor $\mathbf{F}^{\mathrm{i}}$, i.e., the height-channel permutation branch, the width-channel permutation branch, the linear mapping branch, and the spatial DFT branch. Additionally, the phase-aware module $\delta(\cdot)$ is integrated into selected branches, enhancing the efficient feature learning for the hybrid-field channel[4]. Finally, the feature tensors $\mathbf{F}^{\mathrm{o}}_i(i \in \{1,\ldots,4\})$ are aggregated to generate the output feature $\mathbf{F}^{\mathrm{o}}$ with various feature fusion strategies.

[4]In RIS-aided LEO satellite systems, the high mobility of the LEO satellite induces severe Doppler frequency shifts, leading to dramatic phase variations in the cascaded channel. Particularly, for the considered hybrid-field channel, the non-linear phase of the array response imposes greater challenges for channel prediction. Hence, we employ the phase-aware module to focus on the specific phase information of feature tensors.

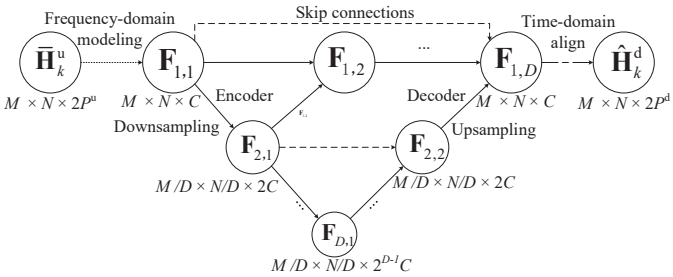

Fig. 2. Deep supervision-based encoder-decoder architecture.

C. Deep Supervision-Based Encoder-Decoder Architecture

Fig. 2 presents the node graph for deep supervision-based encoder-decoder architecture, which takes the frequency-domain MLP and the multi-branch phase-aware module as basic components. After frequency-domain modeling, the tensor operation for feature map $\mathbf{F}_{i,j}(1 \le i,j \le D)$ is given by

$$\mathbf{F}_{i,j} = \begin{cases} \mathcal{E}^{\mathrm{e}}\left(\mathbf{F}_{i,j}\right), & j = 1, \\ \mathcal{E}^{\mathrm{e}}\left(\left[\left[\mathbf{F}_{i,s}\right]_{s=0}^{j}, \mathcal{E}^{\mathrm{d}}(\cdot)\left(\mathbf{F}_{i+1,j}\right)\right]\right), & j > 1, \end{cases} \tag{13}$$

where $\mathcal{E}^{\mathrm{e}}(\cdot)$ and $\mathcal{E}^{\mathrm{d}}(\cdot)$ denote an encoder and a decoder operation, respectively. Each encoder has a multi-branch phase-aware module and a convolutional layer with stride 2, while each decoder consists of a multi-branch phase-aware module and a nearest interpolation-based upsampling layer.

To improve gradients flow and parameters update in the node graph, we design a deep supervision training strategy to reconstruct the loss function in (14), which is given by

$$\overline{\mathcal{L}} = \sum_{i=1}^{D-1}\overline{\boldsymbol{\alpha}}_i\mathcal{L}\left(\overline{\mathbf{H}}^{\mathrm{d}}_k, \widehat{\mathbf{H}}^{\mathrm{d}}_{k,i}\right) + (1-\overline{\boldsymbol{\alpha}}_i)\mathcal{L}\left(\mathcal{F}^{\mathrm{T}}(\overline{\mathbf{H}}^{\mathrm{d}}_k), \mathcal{F}^{\mathrm{T}}(\widehat{\mathbf{H}}^{\mathrm{d}}_{k,i})\right), \tag{14}$$

where $0 \le \overline{\boldsymbol{\alpha}}_i \le 1$ is a pre-defined weight, and an auxiliary loss term $\mathcal{L}\left(\mathcal{F}^{\mathrm{T}}(\overline{\mathbf{H}}^{\mathrm{d}}_k), \mathcal{F}^{\mathrm{T}}(\widehat{\mathbf{H}}^{\mathrm{d}}_{k,l})\right)$ is introduced to align the predicted CSI and true CSI in the frequency domain.

IV. Numerical Results

In simulation setups, we set $M = 4\times 8$, $N = 4\times 64$, $K = 4$, $P^{\mathrm{u}} = 10$, $P^{\mathrm{d}} = 5$, $S^{\mathrm{n}} = S^{\mathrm{f}} = 3$, $\kappa^{\mathrm{r}} = \kappa^{\mathrm{r}}_k = 10$ dB, $H_{\mathrm{S}} = 500$ km, $H_{\mathrm{R}} = 20$ m, $H_{\mathrm{U}} = 1$ m, $v_{\mathrm{S}} = 7000$ m/s, $v_{\mathrm{U}} = 1$ m/s, $f^{\mathrm{u}} = 9.9$ GHz and downlink carrier frequency $f^{\mathrm{d}} = 10.1$ GHz. The parameters of path loss and shadow fading follow the 3GPP TR 38.811 protocol [13]. In the hyper-parameter setups of the proposed SIN, we set $\mathcal{S} = 2$, $D = 3$, $C = 48$, $\overline{\boldsymbol{\alpha}}_l = 0.6, i \in \{1,2\}$, and the number of training epochs is $\mathcal{P} = 50$. We compare the proposed SIN with state-of-the-art benchmarks, i.e., LSTM [8], CNN-LSTM [9], TCN [10], and Transformer [11]. The normalized mean squared error (NMSE) is employed as the performance metric, i.e., $\mathrm{NMSE} = \mathbb{E}\{||\widetilde{\mathbf{H}}^{\mathrm{d}}_k - \mathbf{H}^{\mathrm{d}}_k||^2_F / ||\mathbf{H}^{\mathrm{d}}_k||^2_F\}$.

In Fig. 3, we present NMSE performance of different channel prediction schemes as the uplink SNR. Due to the high-dimensional characteristics of cascaded channel, the serialization operation-based traditional LSTM model is hard to accomplish the desired channel prediction task. Although

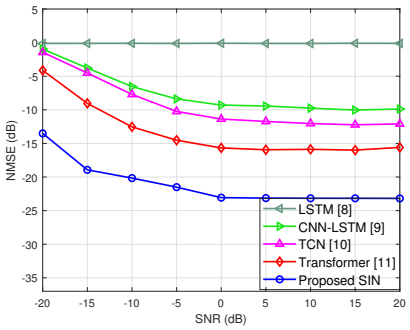

Fig. 3. NMSE vs. SNR for different schemes.

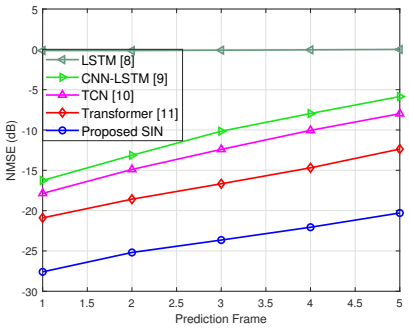

Fig. 4. NMSE vs. prediction frame.

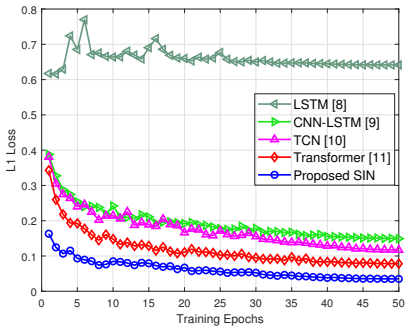

Fig. 5. Loss value vs. number of training epochs.

TABLE I
TRAINING OVERHEAD FOR DIFFERENT NETWORK MODELS

Networks	Parameters (M)	FLOPs (G)
LSTM	290.2	4.063
CNN-LSTM	296.8	4.701
TCN	10.242	3.955
Transformer	1.310	15.61
Proposed SIN	**3.315**	**7.859**

existing enhanced approaches, e.g., CNN-LSTM, TCN and Transformer, provide certain performance improvements, the proposed SIN exhibits superior channel prediction accuracy across the entire range of SNR. In Fig. 4, we provide NMSE performance of various schemes in the p-th prediction frame ($p = \{1, 2, \ldots, P^{\mathrm{d}}\}$), where the uplink SNR is set to 0 dB in the test stage. As the number of prediction frames increases, the NMSE performance of all channel prediction schemes declines due to weaker correlations. However, compared to existing benchmarks, the proposed SIN demonstrates superior accuracy and robustness. Moreover, Fig. 5 shows the convergence performance of different schemes during the network training stage, in which the proposed SIN can achieve the faster convergence than other schemes.

Table I summarizes the number of parameters and floating point of operations (FLOPs) of different network models. Compared to the classic channel prediction models, e.g., LSTM, CNN-LSTM and TCN, the proposed SIN requires fewer parameters. Additionally, it incurs lower FLOPs than the Transformer. Overall, the proposed SIN achieves a superior balance between the spatial and computational complexities while also delivering enhanced accuracy in channel prediction.

V. CONCLUSIONS

This letter proposed an efficient high-dimensional hybrid-field cascade channel prediction model for the RIS-aided FDD satellite. Specifically, we designed frequency-domain MLP and multi-branch phase-aware modules to realize global time-domain and spatial-domain correlation modeling of cascaded channels, respectively. A deep supervision-based encoder-decoder architecture was developed to exploit the latent presentations of high-dimensional hybrid-field channels. Numerical results demonstrate that the proposed SIN exhibits superior channel prediction accuracy compared to existing benchmarks. In the future works, on the basis of the proposed SIN, we will further develop the specialized CSI acquisition scheme for multi-hop RIS empowered non-terrestrial networks.

REFERENCES

[1] Z. Xiao *et al.*, "LEO satellite access network (LEO-SAN) toward 6G: Challenges and approaches," *IEEE Wireless Commun.*, vol. 31, no. 2, pp. 89–96, Apr. 2024.

[2] S. Lin, J. An, L. Gan, M. Debbah, and C. Yuen, "Stacked intelligent metasurface enabled LEO satellite communications relying on statistical CSI," *IEEE Wireless Commun. Lett.*, vol. 13, no. 5, pp. 1295–1299, May 2024.

[3] C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu, Y. Pan, H. Ren, M. D. Renzo, A. Lee Swindlehurst, R. Zhang, and A. Y. Zhang, "An overview of signal processing techniques for RIS/IRS-aided wireless systems," *IEEE J. Sel. Top. Signal Process.*, vol. 16, no. 5, pp. 883–917, Aug. 2022.

[4] J. An, C. Yuen, L. Dai, M. Di Renzo, M. Debbah, and L. Hanzo, "Near-field communications: Research advances, potential, and challenges," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 100–107, Jun. 2024.

[5] X. Mu, J. Xu, Y. Liu, and L. Hanzo, "Reconfigurable intelligent surface-aided near-field communications for 6G: Opportunities and challenges," *IEEE Veh. Technol. Mag.*, vol. 19, no. 1, pp. 65–74, Mar. 2024.

[6] J. Wang, J. Xiao, Y. Zou, W. Xie, and Y. Liu, "Wideband beamforming for RIS assisted near-field communications," *IEEE Trans. Wireless Commun.*, 2024.

[7] J. Xiao, J. Wang, Z. Wang, J. Wang, W. Xie, and Y. Liu, "Multi-task learning for near/far field channel estimation in STAR-RIS networks," *IEEE Trans. Commun.*, vol. 72, no. 10, pp. 6344–6359, Oct. 2024.

[8] W. Xu, J. An, Y. Xu, C. Huang, L. Gan, and C. Yuen, "Time-varying channel prediction for RIS-assisted MU-MISO networks via deep learning," *IEEE Trans. Cognit. Commun. Networking*, vol. 8, no. 4, pp. 1802–1815, Dec. 2022.

[9] C. Nguyen, T. M. Hoang, and A. A. Cheema, "Channel estimation using CNN-LSTM in RIS-NOMA assisted 6G network," *IEEE Trans. Mach. Learn. Commun. Networking*, vol. 1, pp. 43–60, May 2023.

[10] T. Gong, S. Zhang, M. Xu, Z. Li, and O. A. Dobre, "Lightweight TCN-based spatial-temporal channel extrapolation for RIS-aided communication," *IEEE Trans. Veh. Technol.*, 2024.

[11] G. Xia, H. Liu, and K. Long, "Transformer-empowered parallel channel prediction for fast-paced and dynamic RIS-aided wireless communication systems," *IEEE Commun. Lett.*, vol. 28, no. 6, pp. 1347–1351, Jun. 2024.

[12] L. You, K.-X. Li, J. Wang, X. Gao, X.-G. Xia, and B. Ottersten, "Massive MIMO transmission for LEO satellite communications," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1851–1865, Aug. 2020.

[13] G. T. 38.811, "Study on new radio (NR) to support non-terrestrial networks (release 15)," *document TR 38.811, V15.4.0, 3GPP*, 2020.

[14] 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, Mar. 2022.

[15] T. L. Jensen and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *Proc.IEEE ICASSP*. IEEE, 2020, pp. 5000–5004.

[16] K. Gröchenig, *Foundations of time-frequency analysis*. Springer Science & Business Media, 2013.