Semi-Passive RIS-Aided Sequential Channel Estimation and Prediction

Mirza A. Haider [1], Yimin D. Zhang [1], Yanwu Ding [2], Dan Shen [3], Khanh Pham [4], and Genshe Chen [3]

[1] Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA
[2] Department of Electrical and Computer Engineering, Wichita State University, Wichita, KS 67260, USA
[3] Intelligent Fusion Technology, Inc., Germantown, MD 20876, USA
[4] Air Force Research Laboratory, Kirtland Air Force Base, NM 87117, USA

***Abstract*—The popularity of mmWave in 5G and future communications is hindered by challenging propagation environments, such as line-of-sight obstruction. Reconfigurable intelligent surfaces (RIS) address this issue by dynamically modifying wireless channels, thereby enhancing data rates, reducing latency, and improving reliability in non-line-of-sight scenarios. For high data-rate communication and precise mobile user localization, a large RIS is required, resulting in a high pilot overhead for channel estimation. To address this issue, we exploit a semi-passive RIS with sparsely distributed active RIS elements in lieu of fully passive RIS. This approach efficiently enables channel estimation both between the base station and the RIS as well as between the RIS and the mobile users. Structured covariance matrix interpolation optimally utilizes the array aperture from the sparsely placed active RIS elements. Recognizing the need for frequent channel estimation, we introduce a recurrent neural network-based model for sequential channel prediction, resulting in a significant reduction of the required training pilot signals. Simulation results affirm the capability and effectiveness of the proposed approach to enhance data transmission.**

Keywords: Channel estimation, reconfigurable intelligent surface, sparse array, neural network, sequential learning.

I. Introduction

Reconfigurable intelligent surfaces (RIS) have garnered significant attention for 5G and beyond, showcasing their capability to dynamically modify wireless channels between communicating units. They bring forth advantages such as enhanced spectral efficiency, lowered base station (BS) transmit power, deployment flexibility, improved mitigation of multiuser interference, and precise direction of arrival (DOA) estimation [1–4]. Wireless communication systems assisted by RIS achieve enhanced signal-to-noise ratio (SNR), reduced outage probability and symbol error rate, and higher ergodic capacity [5, 6].

In multiple-input multiple-output (MIMO) communication systems, the optimization of transmitter parameters, such as the transmit power, modulation constellation, data rate, antenna selection, and precoding codewords, depends on the knowledge of the channel state information (CSI) to achieve high efficiency and reliability [7]. On the other hand, when the line-of-sight (LOS) between the BS and a mobile user (MS) is obstructed, exploitation of RIS is found effective to create a virtual LoS link between them [3]. However, a MIMO system exploiting fully passive RIS can only estimate the cascade CSI linking the BS and MS via the RIS. An effective strategy to enable CSI estimation of both BS-RIS and MS-RIS channels is through the exploitation of a semi-passive RIS which includes a small number of active RIS elements with sensing capabilities [8, 9]. In particular, leveraging structural placement of sparse active RIS elements benefits from reduced hardware complexity and shortened training data for channel estimation [10–12].

For rapidly time-varying channels, the CSI needs to be frequently estimated, which, in turn, necessitates a high overhead of pilot sequences. However, using advanced machine learning techniques, it is possible to predict the MS-BS or MS-RIS uplink channel and anticipate upcoming channels. In [13], channel predictors based on a vector Kalman filter and on a multilayer perceptron (MLP) are found to have similar performance. In [14], convolutional neural networks combined with autoregressive models and recurrent neural networks (RNN) are used to estimate and predict the CSI in massive MIMO systems. RNNs are effective tools for CSI prediction as they serve as nonlinear approximators for mapping features within the data [15, 16]. Simple RNNs are often unable to effectively process distant data because of the vanishing gradient problem. To overcome this issue, advanced architectures, such as autoregressive networks with exogenous inputs (NARX) as well as evolved RNN models, e.g., long short-term memory (LSTM) and gated recurrent units (GRU), have been proposed [17, 18]. In [16, 19], it is found that the semi-passive RIS can perform channel estimation and prediction, but repetitive estimation is required for predicting each single channel.

In this paper, we introduce a hybrid model which is designed to efficiently estimate and predict time-varying MS-RIS uplink channels so that the pilot overhead is reduced. The proposed approach utilizes a limited number of active RIS elements, arranged as an L-shape sparse array, to estimate the azimuth and elevation angles of the mobile user. In this approach, the fast-fading multipath uplink channel from a mobile user to the semi-passive RIS is first estimated. We employ a structured matrix completion method to the received data observed at the sparse active RIS elements to enable enhanced estimation of the azimuth and elevation DOAs with an increased number of degrees of freedom. To properly pair between the multipath azimuth and elevation angles, we first estimate the signal covariance matrix based on the estimated elevation angles, followed by the estimation of the azimuth angles which minimize the error between the observed data covariance matrix and that based on the estimated model. The full channel is recovered after determining the path loss. Next, we predict the CSI over time by utilizing GRU-based evolved RNN models, thus avoiding the transmission of pilot signals from the mobile user during selected time frames. The predicted channels obtained in a time frame are further utilized as the input of the network to predict the subsequent channels. By using fewer pilot symbols, the proposed model reduces the channel estimation overhead, allowing for more efficient data transmission. Simulation results verify the effectiveness of the proposed GRU-based techniques for sequential prediction of the channels. The advantages of using GRU-based channel prediction over that based on a simple RNN network are also demonstrated.

Notations: Lower- and upper-case bold letters are used to denote vectors and matrices, respectively. Specifically, $\mathbf{I}_N$ represents the $N \times N$ identity matrix. We use $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ to denote the transpose

This material is based upon work supported by the Air Force Research Laboratory (AFRL) under Contract No. FA9453-22-C-A127. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the AFRL.

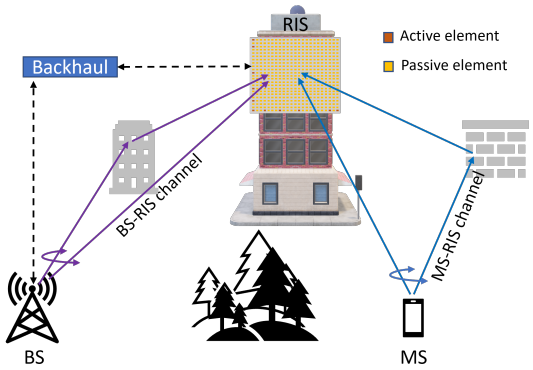


Fig. 1: Communication between a BS and user is facilitated by an RIS-assisted channel model.

and Hermitian operations of a matrix or a vector. The $(\cdot)^{\dagger}$ operator returns the pseudo-inverse of a matrix. The trace of a matrix is represented by $\mathrm{Tr}(\cdot)$, while $\mathrm{diag}(\cdot)$ translates a vector into a diagonal matrix. $\otimes$ denotes the Kronecker product. In addition, $\|.\|_*$ and $\|.\|_{\mathrm{F}}$ respectively represent the nuclear and Frobenius norms of a matrix. We use $\mathbb{E}[\cdot]$ to denote statistical expectation, and $\mathbb{C}^{M\times N}$ to represent the complex space with a dimension of $M \times N$.

II. System and Signal Models

A. Array Signal Models

Consider a downlink communication system where the BS has B_b antennas, while the MS is equipped with a single antenna. Due to the obstacles that block the LOS between the BS and the MS, wireless communication between them is established through a semi-passive RIS. The RIS is assumed to be a rectangular planar shape with $M = M_x \times M_z$ elements, which are uniformly spaced with distance d across both x and z axes. While most of the M RIS elements are passive, i.e., they only have the passive reflecting capability, a small number of $\bar{M}$ active elements, which can actively sense the impinging signal while maintaining their passive reflecting capability, are used to enable channel estimation between the MS and RIS. We place the active RIS elements in an L-shape, comprising two linear subarrays respectively with $\bar{M}_x$ and $\bar{M}_z$ elements in the x and z directions. Therefore, the total number of active RIS elements is given as $\bar{M} = \bar{M}_x + \bar{M}_z - 1$ because one active RIS element is shared by both subarrays.

Denoting θ as the elevation angle and ϕ as the azimuth angle with $0 \le \theta \le \pi/2$ and $-\pi/2 \le \phi \le \pi/2$, the mth element of the steering vector $\mathbf{a}_{\mathrm{RIS}}(\theta,\phi)$ is expressed as

$$[\mathbf{a}_{\mathrm{RIS}}(\theta,\phi)]_m = e^{-j\frac{2\pi}{\lambda}((m_x-1)d\sin(\theta)\sin(\phi)+(m_z-1)d\cos(\theta))} \quad (1)$$

for $m_x = 1, 2, \cdots, M_x$ and $m_z = 1, 2, \cdots, M_z$, where λ stands for the wavelength. The two-dimensional steering vector $\mathbf{a}_{\mathrm{RIS}}(\theta,\phi)$ can be decomposed into the Kronecker product of the horizontal and vertical subarrays $\mathbf{a}_x(\theta,\phi)$ and $\mathbf{a}_z(\theta)$, i.e.,

$$\mathbf{a}_{\mathrm{RIS}}(\theta,\phi) = \mathbf{a}_x(\theta,\phi) \otimes \mathbf{a}_z(\theta) \quad (2)$$

with

$$\mathbf{a}_x(\theta,\phi) = [1, e^{-j\frac{2\pi}{\lambda}d\sin(\theta)\sin(\phi)}, \ldots, e^{-j\frac{2\pi}{\lambda}(M_x-1)d\sin(\theta)\sin(\phi)}]^{\mathrm{T}}, \quad (3)$$

$$\mathbf{a}_z(\theta) = [1, e^{-j\frac{2\pi}{\lambda}d\cos(\theta)}, \ldots, e^{-j\frac{2\pi}{\lambda}(M_z-1)d\cos(\theta)}]^{\mathrm{T}}. \quad (4)$$

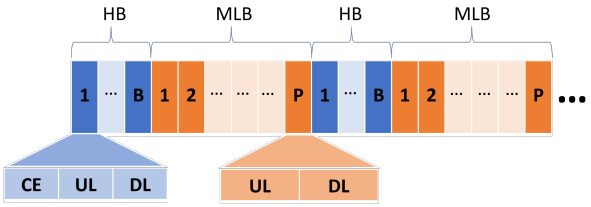


Fig. 2: The transmission frame structure.

The MS-RIS channel vector is the result of L superposed paths,

$$\mathbf{h} = \sum_{l=1}^{L} \beta_l \mathbf{a}_{\mathrm{RIS}}(\theta_l, \phi_l), \quad (5)$$

where β_l is path gain of the lth path arriving from (θ_l, ϕ_l) for $l = 1, 2, \cdots, L$.

B. Signal Structure

Fig. 2 shows the framework for channel detection and prediction. Because both BS and RIS have fixed positions and, as a result, the BS-RIS channel is relatively stable, our focus is on the estimation of the dynamic MS-RIS channel. Compared to traditional time division duplexing (TDD) schemes, we develop a machine learning (ML)-based TDD (ML-TDD) scheme that reduces channel estimation overhead within the frame structure and thus improves the efficiency of data transmission. The CSI is obtained using ML techniques, leveraging correlation among adjacent intervals [14].

As depicted in Fig. 2, each loop in the ML-TDD scheme consists of a head block (HB) and a ML block (MLB). After each loop, a new loop begins to track environmental variations through the HB and continue the ML block. An HB consists of B conventional TDD coherence intervals, where channels are estimated using the sparse RIS active elements, serving as uplink channel CSI data. An MLB comprises P coherence intervals without the channel estimation phase, i.e., no pilot sequences are transmitted from the MS. Instead, the MS-RIS channel is predicted using the ML-based channel predictor scheme during this block. The ML-based TDD process involves the estimation of the CSI in the HBs using the sparse array estimator, followed by the prediction of the CSI using a GRU-based network based on the previously estimated CSI.

Within each HB block, two modes are present: channel estimation (CE) and data transmission. In the CE mode, the MS transmits pilot signals so the RIS, using its active elements, can estimate the channel. Active RIS elements gather $\dot{T}$ samples per time frame within the HB block, forwarding them to the BS via a backhaul network to carry out the estimation of the MS-RIS channel $\mathbf{h}_T \in \mathbb{C}^{M\times 1}$. The data transmission mode encompasses both uplink and downlink phases. In this mode, the RIS switches to a reflection mode, utilizing all components to reflect signals from both uplink and downlink communications.

III. Proposed Channel Estimation and Prediction Method

We present the channel estimation and prediction steps in two respective subsections.

A. RIS-Assisted MS-RIS Channel Estimation

In each HB, the RIS estimates the MS-RIS channel using sparse RIS elements. We employ the hybrid optimized non-redundant array (ONRA) structure in an L-shaped configuration [12, 20]. The ONRA configurations offer optimized array design with redundancy-free

difference lags and a desired array aperture. The positions of the RIS subarray elements along the x- and z-axes are denoted as $\mathbb{X} = \{p_0, p_1, \cdots, p_{\bar{M}x-1}\}\lambda/2$ and $\mathbb{Z} = \{q_0, q_1, \cdots, q_{\bar{M}z-1}\}\lambda/2$, respectively, where p_i and q_i are integers for all i, and $p_0 = q_0 = 0$ is assumed. The steering vectors corresponding to these two RIS subarrays are respectively denoted as $\tilde{\mathbf{a}}_x(\theta_l, \phi_l)$ and $\tilde{\mathbf{a}}_z(\theta_l)$.

Denote the uplink pilot signal transmitted by the MS as $s_u(t) = \sqrt{\Gamma}s(t)$, where Γ is the transmitted power and $s(t)$ is the signal waveform with $\mathbb{E}(|s(t)|^2) = 1$. Then, the received signals at the x-axis and z-axis subarrays are respectively given as [10]

$$\mathbf{x}(t) = \sum_{l=1}^{L} \beta_l \tilde{\mathbf{a}}_x(\phi_l, \theta_l) s_u(t) + \mathbf{n}_x(t), \tag{6}$$

$$\mathbf{z}(t) = \sum_{l=1}^{L} \beta_l \tilde{\mathbf{a}}_z(\theta_l) s_u(t) + \mathbf{n}_z(t), \tag{7}$$

for $t = 1, 2, \cdots, \dot{T}$, where $\mathbf{n}_x(t)$ and $\mathbf{n}_z(t)$, both following $\mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{n}}^2 \mathbf{I}_N)$, represent the additive white Gaussian noise vectors.

We first perform DOA estimation of the received signals at the vertical subarray. The signal vectors observed at the $\dot{T}$ samples are stacked as matrix

$$\mathbf{Z} = [\mathbf{z}(1), \mathbf{z}(2), \cdots, \mathbf{z}(\dot{T})]. \tag{8}$$

Assuming that the noise elements are uncorrelated with the signals, we estimate the covariance matrix of $\mathbf{z}(\mathbf{t})$ as

$$\tilde{\mathbf{R}}_{z_{\mathrm{RIS}}} = \mathbf{Z}\mathbf{Z}^{\mathrm{H}} = \tilde{\mathbf{A}}_z \mathbf{R}_s \tilde{\mathbf{A}}_z^{\mathrm{H}} + \sigma_n^2 \mathbf{U}_z \mathbf{U}_z^{\mathrm{H}}, \tag{9}$$

where $\mathbf{R}_s$ is the source covariance matrix, $\tilde{\mathbf{A}}_z = [\tilde{\mathbf{a}}_z(\theta_1), \tilde{\mathbf{a}}_z(\theta_2), \cdots, \tilde{\mathbf{a}}_z(\theta_L)]$, and $\mathbf{U}_z$ is a binary diagonal mask matrix indicating the presence of sensors in the z-axis subarray, i.e., $[\mathbf{U}_z]_{g,g}$ is equal to 1 if $g\lambda/2 \in \mathbb{Z}$ and 0 otherwise. Ideally, $\mathbf{R}_s$ is given as $\mathbf{R}_s = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_L^2)$ with σ_l^2 denoting the power of the lth signal path.

Due to the sparse arrangement of active RIS elements, matrix $\tilde{\mathbf{R}}_{z_{\mathrm{RIS}}}$ contains missing elements when considered in the half-wavelength space. To address this, we employ matrix completion techniques to interpolate $\tilde{\mathbf{R}}_{z_{\mathrm{RIS}}}$, resulting in the covariance matrix for the virtual uniform linear array (ULA) along the z-axis, denoted as $\hat{\mathbf{R}}_{z_{\mathrm{RIS}}}$. Covariance matrix interpolation can be solved using a rank minimization or atomic norm minimization problem [21, 22].

Once the interpolated covariance matrix $\hat{\mathbf{R}}_{z_{\mathrm{RIS}}}$ is obtained, we use the MUSIC algorithm to estimate the elevation DOAs of the MS-RIS multipath signals. Subsequently, the array manifold matrix $\mathbf{A}_z(\hat{\theta}) \in \mathbb{C}^{M_z \times L}$ for the z-axis subarray is constructed based on these estimated elevation angles.

Azimuth angle estimates are required to align them with their corresponding elevation angles. Leveraging the interconnections among signal subspace components facilitates the estimation of $\mathbf{R_s}$ as follows [12]:

$$\hat{\mathbf{R}}_s = \mathbf{A}_z^{\dagger}(\hat{\theta}) \hat{\mathbf{V}}_{zs} (\hat{\mathbf{\Gamma}}_z - \sigma_{\mathrm{n}}^2 \mathbf{I}_L) \hat{\mathbf{V}}_{zs}^{\mathrm{H}} (\mathbf{A}_z^{\dagger}(\hat{\theta}))^{\mathrm{H}}, \tag{10}$$

where $\hat{\mathbf{V}}_{zs} \in \mathbb{C}^{M_x \times L}$ represents the estimated signal subspace for the z-axis subarray, while $\hat{\mathbf{\Gamma}}_{zs} \in \mathbb{C}^{L \times L}$ is a diagonal matrix containing the eigenvalues corresponding to the signal subspace.

Since both the azimuth and elevation subarrays observe the same signal power, the signal eigenvalues are interchangeable between the two subarrays. Thus, the eigenvalues of the covariance matrix for both subarrays follow the order $\gamma_1 > \gamma_2 > \cdots > \gamma_L$. Therefore, the azimuth angles can be determined from the subsequent optimization problem:

$$\hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} \|(\hat{\mathbf{R}}_{x_{\mathrm{RIS}}} - \sigma_n^2 \mathbf{I}_{M_x}) - \mathbf{A}_x(\boldsymbol{\phi}, \hat{\theta}) \hat{\mathbf{R}}_s \mathbf{A}_x^{\mathrm{H}}(\boldsymbol{\phi}, \hat{\boldsymbol{\theta}})\|, \tag{11}$$

where $\mathbf{A}_x(\hat{\phi}, \hat{\theta}) = [\mathbf{a}_x(\hat{\phi}_1, \hat{\theta}_1), \cdots, \mathbf{a}_x(\hat{\phi}_L, \hat{\theta}_L)] \in \mathbb{C}^{M_x \times L}$. Finally, after both azimuth and elevation angles of all paths are obtained, the MS-RIS channel gains β_l of each path can be separately estimated [12]. As a result, in the Tth time frame of an HB block, the MS-RIS multipath channel $\mathbf{h}_T$ is reconstructed based on (5).

B. RIS-Assisted MS-RIS Channel Prediction

As illustrated in Fig. 3(a), an HB contains B coherent blocks, and consecutive channels $\hat{\mathbf{h}}_{T-B+1}, \hat{\mathbf{h}}_{T-B+2}, \cdots, \hat{\mathbf{h}}_T$ are estimated at time T. A deep machine learning-based algorithm is then introduced that leverages a GRU network to predict channels $\mathbf{h}_{T+1}, \mathbf{h}_{T+2}, \cdots, \mathbf{h}_{T+P}$ in the MLB. In this sequential channel prediction scheme in the MLB block, these estimated channels are utilized to predict $\mathbf{h}_{T+1}$, which is then fed into the network to predict $\mathbf{h}_{T+2}$, and the process continues until the last time frame in the MLB.

The GRU network is selected for its ability to handle long-term dependencies, achieved through memory cells and multiplicative gates, as illustrated in Fig. 3(b). In each time step, the memory cell produces output by considering the estimated channel input and feedback from the hidden state of the previous time step. The GRU model has two gating signals: an update gate called $\mathbf{y}_T$ and a reset gate called $\mathbf{r}_T$. These gates can be fine-tuned based on the rate of channel variations, providing the GRU network with better adaptability and robustness compared to memoryless networks like MLPs.

For the activation vectors of gates, the current input channel vector $\hat{\mathbf{h}}_T$ and the memory from previous short-term $\mathbf{s}_{T-1}$ are fed into three distinct fully connected (FC) layers, the equations of the update and reset gates are respectively expressed as [18]

$$\mathbf{y}_T = \sigma_g(\mathbf{W}_y \hat{\mathbf{h}}_T + \mathbf{U}_y \mathbf{s}_{T-1} + \mathbf{b}_y), \tag{12}$$

$$\mathbf{r}_T = \sigma_g(\mathbf{W}_r \hat{\mathbf{h}}_T + \mathbf{U}_r \mathbf{s}_{T-1} + \mathbf{b}_r), \tag{13}$$

where subscripts y and r correspond to the update and reset gates, respectively. The weight matrices $\mathbf{W}$ and $\mathbf{U}$, along with appropriate subscripts, represent the weight matrices for the FC layers corresponding to the current input vector $\hat{\mathbf{h}}_T$ and the previous state $\mathbf{s}_T$. The bias vector $\mathbf{b}$, with an appropriate subscript, represents the bias vector. In addition, $\sigma_g(\cdot)$ denotes the sigmoid activation function.

Through the process of discarding old memories of the hidden state at the previous time step and integrating new ones the GRU network is updated as

$$\mathbf{s}_T = (1 - \mathbf{y}_T) \otimes \mathbf{s}_{T-1} + \mathbf{y}_T \otimes \tanh(\mathbf{W}_s \hat{\mathbf{h}}_T + \mathbf{U}_s(\mathbf{r}_T \otimes \mathbf{s}_{T-1}) + \mathbf{b}_s), \tag{14}$$

where $\tanh(\cdot)$ denotes the hyperbolic tangent function. Next, the new step $\mathbf{s}_T$ and output $\hat{\mathbf{h}}_{T+1}$ are then used as the input to predict $\hat{\mathbf{h}}_{T+2}$

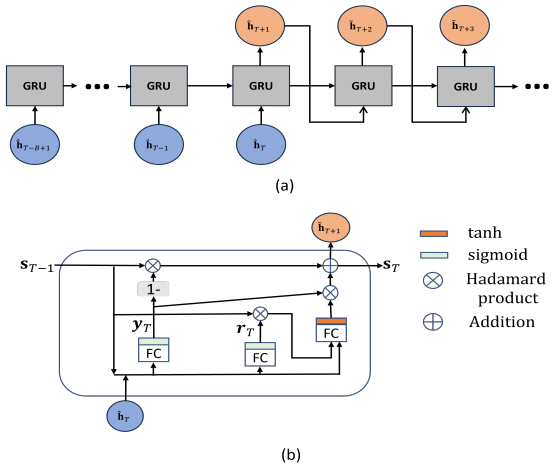


Fig. 3: GRU structure for channel prediction.

and so on. We used the general form of RNN [23] to compare with the GRU model.

The model's overall structure is a sequence-to-vector GRU. In the GRU training phase, the GRU layer processes an input sequence of length B. The input to the GRU network is a series of channel vectors $\hat{\mathbf{h}}_{T-B+1}, \hat{\mathbf{h}}_{T-B}, \cdots, \hat{\mathbf{h}}_T$ which are estimated in the HB. The output is the predicted channel vector $\mathbf{h}_{T+1}$ for the $(T+1)$-th time slot, which is generated by a dense layer. In order to use a real-valued GRU architecture, we decompose the complex-valued input vectors into real and imaginary parts, i.e., $\{\mathrm{Re}(\hat{\mathbf{h}}_{T-B+1}), \mathrm{Im}(\hat{\mathbf{h}}_{T-B+1}), \cdots, \mathrm{Re}(\hat{\mathbf{h}}_T), \mathrm{Im}(\hat{\mathbf{h}}_T)\}$. The output layer is configured with a dimension of $2M$, representing the real and imaginary parts of the channel vector $\{\mathrm{Re}(\mathbf{h}_{T+1}), \mathrm{Im}(\mathbf{h}_{T+1})\}$ at the $(T+1)$-th time slot. Using this predicted channel vector as input, $\{\mathrm{Re}(\mathbf{h}_{T+2}), \mathrm{Im}(\mathbf{h}_{T+2})\}$ can be predicted. This sequential process extends over P coherent intervals before initiating the subsequent loop.

IV. Simulation Results

The RIS is a square UPA that comprises $M = 23 \times 23 = 529$ elements. The carrier frequency used is $f_c = 28$ GHz. We consider two paths, denoted by $L = 2$, between the mobile user and the RIS. The path loss exponents for these paths are set to 2.2 and 2.1, respectively. There are $\bar{M} = 11$ active RIS elements which are configured in an L shape consisting of two linear subarrays, each having $\bar{M}_x = \bar{M}_z = 6$ active sensors which are placed based on the hybrid ONRA design. More specifically, these active sensors are positioned along the x- and z-axes at $\mathbb{X} = \mathbb{Z} = \{0, 3, 7, 12, 20, 22\}\lambda/2$. If we consider only the non-negative lags for both subarray, then $\mathbb{D}^X_{\text{self}} = \mathbb{D}^Z_{\text{self}} = \{0, 2, 3, 4, 5, 7, 8, 9, 10, 12, 13, 15, 17, 19, 20, 22\}\lambda/2$.

The RIS is positioned at Cartesian coordinates (30 m, 4 m, 10 m), with the user initially at (30 m, 34 m, 2 m). Training and testing datasets reflect the user's mobility, specifically their walking speeds of 1 m/s and 3 m/s relative to the RIS. For predictive modeling, $N_{\text{train}} = 11^3$ datasets are generated to capture various MS-RIS channel dynamics at these speeds. The ADAM algorithm, utilizing a batch size of 64, trains the model. During training, datasets for both speeds are merged. For testing, the system randomly selects 20% of the data for validation, ensuring the trials are independent

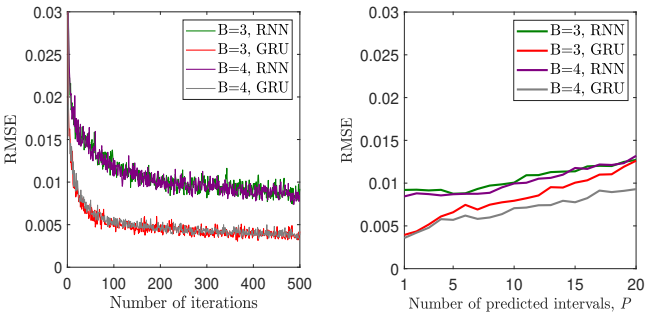
During the analysis of the results, we consider two types of HB frame spans with different numbers of coherent blocks: one with $B = 3$ coherent blocks, and the other with $B = 4$ coherent blocks. When estimating the channel, the user sends $T = 200$ uplink pilot samples, and the MU pilot signal's transmit power used for channel estimation is 25 dBm. The channel noise power remains fixed at $\sigma_n^2 = -80$ dBm. MinMax scaling is applied to each feature dimension independently using MinMaxScaler. The channel prediction performance is evaluated using the root-mean-square error (RMSE) between predicted and actual channel states, expressed as

$$\Delta_n = \sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left\|\hat{\mathbf{h}}_{n,q} - \mathbf{h}_{n,q}\right\|^2} \tag{15}$$

for $n = T+1, T+2, \cdots, T+P$, where Q denotes the number of independent trials.

Fig. 4(a) shows the RMSE variation with respect to the number of iterations for the initial predicted channel $\mathbf{h}_{T+1}$, utilizing RNN and GRU for different input channel orders B. The network undergoes training and testing stages with mixed user speeds (1 m/s and 3 m/s). As the iterations increase, there is a notable decrease in the RMSE of the predicted channels. The figure provides insights into the optimal number of input channels required for accurate predictions.

It is clear that when the input order is low ($B = 1$ and $B = 2$), the channel prediction may not be optimal [16]. Over successive iterations, GRU models outperform traditional RNNs. This advantage



(a) (b)

Fig. 4: RMSE performance of the predicted channels averaged over 5 independent trials. (a) Convergence of the predicted channels ($P = 1$). (b) RMSE performance curve of the predicted channels using RNN and GRU (after 500 iterations).

is attributed to the GRU's ability to maintain essential spatial features over extended transmissions [24]. GRUs are specifically designed to overcome the vanishing gradient problem, which hampers the ability of simple RNNs to recognize long-term dependencies in data sequences.

GRU exploits a more complex structure with memory cells and gates, allowing them to capture and maintain information over longer sequences. Despite RNN's simpler structure, characterized by fewer gates compared to GRU, leading to reduced matrix multiplication and increased speed, it may not capture long-term dependencies as effectively as GRU.

Fig. 4(b) displays the RMSE performance of the test datasets for the RNN and GRU networks, used for sequential time interval prediction. The GRU network Consistently outperforms the RNN networks, as previously described. As the interval between predictions increases, the prediction performance for both RNN and GRU networks decreases. Additionally, all networks achieve better performance when a higher order of the initial input is assumed, especially for the GRU case. As the number of predicted intervals rises, both GRU and RNN networks encounter declining RMSE performance. For the GRU model, the performance degrades more rapidly for the case of $B = 3$, implying that a higher-order model provides more accurate channel estimation and prediction, at the expense of higher computational complexity due to longer input sequences in the model. Depending on the requirements of the networks, we can either shorten or prolong the prediction intervals and start estimating the channels. Through the use of RNN models, we have uncovered that we can predict CSI for extended periods of time. Accurate prediction is crucial for maintaining reliable RIS-to-user connections. It also holds value in advancing sparse semi-passive RIS, offering the potential for establishing high data rate connectivity in future wireless systems.

V. Conclusion

We presented a new method to improve the quality of service in an RIS-assisted network through the estimation and sequential prediction of the channel states. The GRU network utilizes the CSI obtained at the previous time frames to predict subsequent channels. This approach reduces the pilot requirement for channel estimation. With the help of structured matrix completion and pair-matching algorithms, multipath channels are estimated with high precision. This semi-passive RIS configuration expands the array aperture and reduces computational load. Simulation results demonstrated the effectiveness and performance superiority of the proposed approach over other machine learning algorithms.

VI. References

[1] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE commun. Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.

[2] 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. 116 753–116 773, 2019.

[3] 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, 2021.

[4] M. Asif Haider and Y. D. Zhang, "RIS-aided integrated sensing and communication: A mini-review," *Frontiers Signal Process.*, vol. 3, pp. 1–8, May 2023.

[5] E. Björnson, Ö. Özdogan, and E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, 2019.

[6] L. Yang, Y. Yang, D. B. da Costa, and I. Trigui, "Outage probability and capacity scaling law of multiple RIS-aided networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 256–260, 2020.

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[8] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," *IEEE Access*, vol. 9, pp. 44 304–44 321, 2021.

[9] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1707–1725, 2022.

[10] X. Hu, R. Zhang, and C. Zhong, "Semi-passive elements assisted channel estimation for intelligent reflecting surface-aided communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1132–1142, 2021.

[11] M. A. Haider, M. W. T. Chowdhury, and Y. D. Zhang, "Sparse channel estimation for IRS-aided systems exploiting 2-D sparse arrays," in *Proc. IEEE Sensor Array and Multich. Signal Process. Workshop (SAM)*, Trondheim, Norway, June 2022, pp. 31–35.

[12] M. Asif Haider, Y. D. Zhang, and E. Aboutanios, "ISAC system assisted by RIS with sparse active elements," *EURASIP J. Adv. Signal Process.*, vol. 2023, no. 1, pp. 1–22, 2023.

[13] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, "Massive MIMO channel prediction: Kalman filtering vs. machine learning," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 518–528, 2020.

[14] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Machine learning-based channel prediction in massive MIMO with channel aging," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 2960–2973, 2020.

[15] W. Lyu, Z. Zhang, C. Jiao, K. Qin, and H. Zhang, "Performance evaluation of channel decoding with deep neural networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2018, pp. 1–6.

[16] M. A. Haider, S. R. Pavel, Y. D. Zhang, and E. Aboutanios, "Active IRS-assisted MIMO channel estimation and prediction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Rhodes Island, Greece, June 2023, pp. 1–5.

[17] W. Jiang and H. D. Schotten, "Deep learning for fading channel prediction," *IEEE Open J. Commun. Society*, vol. 1, pp. 320–332, 2020.

[18] ——, "Recurrent neural networks with long short-term memory for fading channel prediction," in *Proc. IEEE 91st Vehi. Techn. Conf. (VTC2020-Spring)*, 2020, pp. 1–5.

[19] M. A. Haider, Y. D. Zhang, and E. Aboutanios, "Channel estimation and prediction in wireless communications assisted by semi-passive RIS," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Seoul, Korea, April 2024, pp. 1–5.

[20] A. Ahmed and Y. D. Zhang, "Generalized non-redundant sparse array designs," *IEEE Trans. Signal Process.*, vol. 69, pp. 4580–4594, 2021.

[21] C. Zhou, Y. Gu, X. Fan, Z. Shi, G. Mao, and Y. D. Zhang, "Direction-of-arrival estimation for coprime array via virtual array interpolation," *IEEE Trans. Signal Process.*, vol. 66, no. 22, pp. 5956–5971, 2018.

[22] S. Liu, Z. Mao, Y. D. Zhang, and Y. Huang, "Rank minimization-based Toeplitz reconstruction for DoA estimation using coprime array," *IEEE Commun. Lett.*, vol. 25, no. 7, pp. 2265–2269, 2021.

[23] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[24] S. Yang, X. Yu, and Y. Zhou, "LSTM and GRU neural network performance comparison study: Taking yelp review dataset as an example," in *Proc. Int. Workshop Electron. Commun. Artificial Intelligence (IWECAI)*, Qingdao, China, June 2020, pp. 98–101.