Receive Polarization-Combined Space Shift Keying with Att-LSTM Detector for RIS-Aided Communications

Yuyan Liu[†1], Chaorong Zhang[†2], Benjamin K. Ng[†3*], and Chan-Tong Lam[†4]
[†]*Faculty of Applied Sciences, Macao Polytechnic University, Macao SAR, China*
[1]p2311866@mpu.edu.mo, [2]chaorong.zhang@mpu.edu.mo, [3]bng@mpu.edu.mo, [4]ctlam@mpu.edu.mo
[*]Corresponding author

***Abstract*—Reconfigurable Intelligent Surfaces (RIS) has been gaining recognition as a key technology for future wireless communication systems, thanks to their ability to significantly improve coverage and spectral efficiency by dynamically altering the propagation environment. In this paper, we propose a RIS-aided receive polarization-combined space shift keying (RIS-RPSSK) scheme, where multiple receive antennas with additional polarized states can be chosen to convey information bits without processing of constellation symbols, thus reducing some complexity. To further enhance this scheme, we propose a Long Short-Term Memory (LSTM)-based detector with attention mechanism (Att-LSTM) at the receiver. Simulation results demonstrate that the superiority of the RIS-RPSSK scheme over conventional schemes in terms of bit-error-rate (BER) performance under various conditions. Besides, as a promising technique for future wireless communication, the Att-LSTM detector can achieve satisfactory BER performance with less complexity compared to the traditional maximum likelihood detector.**

***Index Terms*—Reconfigurable intelligent surface, dual polarization, space shift keying, BER, spectral efficiency**

I. Introduction

The accelerated evolution of wireless communication technologies necessitates the development of innovative solutions to meet the increasing demand for higher data rates, enhanced coverage, and improved spectral efficiency (SE). Reconfigurable intelligent surface (RIS) has emerged as a cutting-edge technology, which is capable of dynamically altering the propagation environment to optimize signal transmission through adjusting phase shifts to perform passive beamforming. This novel technology is suitable for extending signal coverage, optimizing radio link quality, and improving positioning accuracy [1]–[3]. RIS is proven in [4] to have the capability to reflect dual-polarized (DP) signals and independently adjust the phases of signals with different polarization states. [5] explores the performance of RIS-assisted mobile communication system under DP setting. Utilizing DP reflecting elements, RIS can control two orthogonal polarizations instantaneously.

Space shift keying (SSK) has garnered attention for its ability to exploit spatial diversity without requiring multiple radio frequency chains. The integration of RIS with SSK (RIS-SSK) [6] highlights the potential to improve BER performance and balance spectral efficiency with energy consumption. In [16], a novel RIS-differential transmitted spatial modulation (DTSM) scheme are demonstrated good BER performance across various scenarios, including Nakagami-fading, which also indicates its potential for practical applications. To fully investigate the additional polarization degrees of freedom, [7] proposes a polarization shift keying to obtain a higher SE. [8] proposes space-polarization shift keying (SPSK) scheme leveraging DP antennas, where the space and polarization index are used to transmit the information. Integrating dual-polarized antennas with RIS-SSK presents an opportunity to further enhance system performance and balances spectral efficiency with energy consumption by leveraging polarization diversity.

As communication scenarios becoming increasingly complex, the improvement of efficient and reliable signal detector is the main issue that we need to focus on. The emerging development of deep learning (DL) techniques provide new opportunities to address signal detection issues in complex communication environments while enhancing the overall performance of wireless systems. [9] proposes data-driven DL detector and model-driven DL detectors to detect the RIS-SSK systems. [10] investigates the use of block deep neural networks (B-DNN) detector in RIS-aided received-SSK (RSSK) systems. The long short-term memory (LSTM) are proposed to achieve intelligent end-to-end transmissions in [11]. The LSTM, a type of recurrent neural network (RNN), have the ability to handle sequences and capture long-term dependencies. Given the characteristics of signal, we seek to leverage LSTM-based detector with attention mechanism (Att-LSTM detector) to enhance RIS-aided communicaton systems.

In this paper, we propose a RIS-aided receive polarization-combined space shift keying (RIS-RPSSK). Through comprehensive modeling and simulation, we demonstrate that the proposed RIS-RPSSK system significantly outperforms traditional SSK systems in terms of BER and SE under the same conditions. Moreover, a novel Att-LSTM detector is proposed to reduce complexity effectively. Meanwhile, the BER performance of proposed Att-LSTM detector is compared with that of maximum likelihood (ML) and greedy detector. Our work provides implications for the development of next-generation wireless communication networks, giving a

This work is funded by The Science and Technology Development Fund, Macau SAR. (File no. 0044/2022/A1)

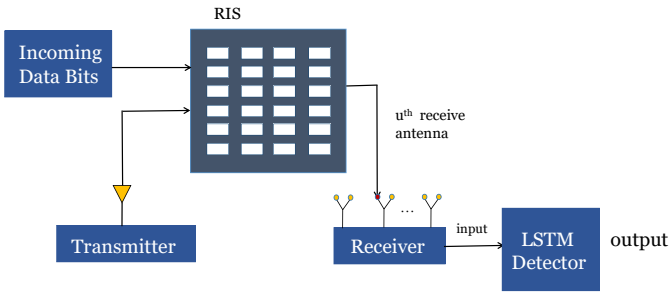

Fig. 1: System model of RIS-RPSSK scheme.

foundational framework for future exploration and innovation in intelligent communication systems.

The paper is organized as follows: Section II details the RIS-aided RPSSK schemes. In Section III, we presents the specific structure and configuration of Att-LSTM detector. Then, Section IV shows the effectiveness of the proposed RIS-aided RPSSK compared to RIS-RSSK/RSM and the proposed Att-LSTM detector compared to the traditional ML and greedy detector. Finally, Section V concludes the paper.

II. RIS-EMPOWERED RPSSK

A. System Model

Fig. 1 shows the diagram of the RIS-aided RPSSK system, which includes a transmitter with DP antenna and a receiver equipped with N_r DP antennas. A RIS with N passive reflector elements is leveraged to reflect signals from transmitter. We assume the direct link of transceiver is blocked and the channel state information (CSI) between RIS and receiver can be known at the RIS controller and solved by [12].

During the transmission of the RIS-RPSSK system, the information bits b mapped into the chosen indices of receive antennas with $L_1 = \log_2 N_r$ bits and the chosen polarization state of receive antennas respectively with $L_2 = \log_2 N_p$ where $N_p = 2$ stands for the left-handed circular polarization (LHCP) and right-handed circular polarization (RHCP) states. Thus, for the RIS-RPSSK, the total number of transmitted bits in any particular time slot is given by

$$B = \log_2 L_1 + \log_2 L_2, \tag{1}$$

The received signal $\mathbf{y} \in \mathbb{C}^{2N_r \times 1}$ can be expressed as

$$\mathbf{y} = N\sqrt{E_s}\zeta \mathbf{G}_2 \mathbf{\Theta} \mathbf{G}_1 + \mathbf{n} \tag{2}$$

$$= \mathbf{S} + \mathbf{n}, \tag{3}$$

where $\mathbf{n} \in \mathbb{C}^{2Nr \times 1}$ is the additive noise vector whose entries follow a complex Gaussian distribution with $\mathcal{CN}(0, \sigma^2)$. $\mathbf{y} = [[y_{1_{\mathrm{L}}}, y_{1_{\mathrm{R}}}], \cdots, [y_{N_{\mathrm{L}}}, y_{N_{\mathrm{R}}}]]^T$, where $y_{n_{\mathrm{L}}}$ and $y_{n_{\mathrm{R}}}$ represent the received signal at the LHCP and RHCP of the n-th DP receive antennas. $\mathbf{G}_2 \in \mathbb{C}^{2N_r \times 2N}$ denotes the channel matrix between RIS and receiver. $\mathbf{G}_1 \in \mathbb{C}^{2N \times 1}$ is the channel matrix between transmitter and RIS. $\mathbf{\Theta} \in \mathbb{C}^{2N \times 2N}$ represents the diagonal matrix of adjusted phase shifts of RIS. $\mathbf{S} = N\sqrt{E_s}\zeta \mathbf{G}_2 \mathbf{\Theta} \mathbf{G}_1$ represents the whole channel matrix via transceiver. $\zeta = \zeta_1 \zeta_2$ represents whole path loss where ζ_1 is path loss between transmitter and RIS and ζ_2 is path loss between RIS and receiver. E_s is the transmitted energy from source.

The channel between the polarized transmitter and nth reflecting element in [13] can be expressed as

$$\mathbf{g}_{1_n} = \mathbf{F}_{ris} \tilde{\mathbf{Q}}_i \mathbf{m}_t(\alpha_t), \tag{4}$$

$\mathbf{F}_{ris}$ is the deterministic polarization matrix of elements which has orthonormal sub-elements as

$$\mathbf{F}_{ris} = \frac{1}{\sqrt{1+\kappa_s}} \begin{bmatrix} 1 & \sqrt{\kappa_s} e^{j\omega_s} \\ -\sqrt{\kappa_s} e^{j\omega_s} & 1 \end{bmatrix}, \tag{5}$$

where κ_s^{-1} and ω_s are the XPI factor and the phase offset of the elements. $\mathbf{m}_t(\alpha_t)$ is the unit-norm polarization vector of the transmitter as

$$\mathbf{m}_t(\alpha_t) = \frac{1}{\sqrt{1+\kappa_t}} \begin{bmatrix} \cos(\alpha_t) & -\sin(\alpha_t) \\ \sin(\alpha_t) & \cos(\alpha_t) \end{bmatrix} \mathbf{F}_t, \tag{6}$$

$$\mathbf{F}_t = \begin{bmatrix} 1 & \sqrt{\kappa_t} e^{j\omega_t} \\ -\sqrt{\kappa_t} e^{j\omega_t} & 1 \end{bmatrix}, \tag{7}$$

where $\alpha_t \in [0, 2\pi]$ is a rotation angle. κ_t^{-1} and ω_t are the XPI factor and the phase offset of the transmitter. $\tilde{\mathbf{Q}}_i$ denotes the two-input two-output channel matrix between transmitter and DP-RIS, expressed as

$$\tilde{\mathbf{Q}}_i = \mathbf{\Sigma_1} \odot (\mathbf{C}_{s_p}^{\frac{1}{2}} \tilde{\mathbf{Z}}_i \mathbf{C}_{t_p}^{\frac{1}{2}}), \tag{8}$$

where $\tilde{\mathbf{Z}}_i = \begin{bmatrix} q_{1,i}^{LL} & q_{1,i}^{LR} \\ q_{1,i}^{RL} & q_{1,i}^{RR} \end{bmatrix}$ is the random channel matrix whose entries follow spatial uncorrelated i.i.d. Gaussian distribution with zero mean. $\mathbf{C}_{s_p}$ and $\mathbf{C}_{t_p}$ represent the RIS and transmitted polarization correlation matrices as

$$\mathbf{C}_{s_p} = \begin{bmatrix} 1 & \rho_s \\ \rho_s^* & 1 \end{bmatrix}, \quad \mathbf{C}_{t_p} = \begin{bmatrix} 1 & \rho_t \\ \rho_t^* & 1 \end{bmatrix}, \tag{9}$$

where ρ_s and ρ_t denote the polarization correlation coefficients at the RIS and transmitter, respectively.

Due to the DP imperfection which refers to polarization conversion, we define the cross polarization discrimination (XPD) to describe the channel's ability to discriminate orthogonal polarizations as

$$\beta_1 = \frac{\mathbb{E}[|h_n^{\mathrm{LL}}|^2]}{\mathbb{E}[|h_n^{\mathrm{RL}}|^2]} = \frac{\mathbb{E}[|h_n^{\mathrm{RR}}|^2]}{\mathbb{E}[|h_n^{\mathrm{LR}}|^2]} = (1-\lambda_1)/\lambda_1, \tag{10}$$

where $0 \le \lambda_1 \le 1$. XPD matrix can be defined as

$$\mathbf{\Delta}_1 = \begin{bmatrix} \sqrt{1-\lambda_1} & \sqrt{\lambda_1} \\ \sqrt{\lambda_1} & \sqrt{1-\lambda_1} \end{bmatrix}, \tag{11}$$

Furthermore, the final channel vector to include the spatial correlation can be modeled as

$$\bar{\mathbf{Q}}_1 = (\mathbf{\Delta_1} \otimes \mathbf{I}_N) \odot ((\mathbf{C}_{s_p} \otimes \mathbf{\Gamma}_{ris})^{\frac{1}{2}} \mathbf{Z}_1 (\mathbf{C}_{t_p} \otimes \mathbf{\Gamma}_t)^{\frac{1}{2}}), \tag{12}$$

where $\mathbf{\Gamma}_{ris} \in \mathbb{C}^{N \times N}$ and $\mathbf{\Gamma}_t \in \mathbb{C}^{N_t \times N_t}$ are the RIS spatial correlation and transmit spatial correlation matrices. The operator $\odot$ denotes the element-by-element Hadamard multiplication, $\otimes$ denotes the Kronecker product. $\mathbf{Z}_1 \in \mathbb{C}^{2N \times 2N_t}$ is the random matrix whose entries follow i.i.d. $\mathcal{CN}(0,1)$. We consider a RIS which consists of $\sqrt{N}$ elements along the y-axis and the z-axis. Thus, the spatial correlation matrices in [15] is used to describe the spatial correlations. $\mathbf{\Gamma}_{ris} = \mathbf{\Gamma}_{s_y} \otimes \mathbf{\Gamma}_{s_z}$, where $\mathbf{\Gamma}_{s_y/s_z} \in \mathbb{C}^{\sqrt{N} \times \sqrt{N}}$ denotes the spatial correlation with spatial correlation factor $|\xi_{s_y/s_z}| < 1$ along the y/z axis

between the m th row and n th column of the element as

$$\mathbf{\Gamma}_{s_y/s_z}^{(m,n)} = \begin{cases} \xi_{s_y/s_z}^{n-m}, m \leqslant n, \\ \left(\xi_{s_y/s_z}^{*}\right)^{m-n}, m > n. \end{cases} \tag{13}$$

Thus, the composite wireless channel between the DP receiver and RIS can be rewritten as

$$\mathbf{G}_1 = (\mathbf{I}_N \otimes \mathbf{F}_{ris})\bar{\mathbf{Q}}_1 \mathbf{m}_t(\alpha_t). \tag{14}$$

Incoming signals from transmitter are reflected towards specified directions at RIS by adjusting each individual RIS element. Then, the received SNR can be maximized by changing RIS phase shifts $\mathbf{\Theta} \in \mathbb{C}^{2N\times 2N}$ at the selected receive antenna as

$$\mathbf{\Theta} = \begin{bmatrix} \mathbf{\Phi}_0 & \\ & \mathbf{\Phi}_1 \end{bmatrix}, \tag{15}$$

where $\mathbf{\Phi}_0 = \mathrm{diag}\{e^{j\phi_{0,1}}, \cdots, e^{j\phi_{0,N}}\}$ stands for phase shifts of signals in LHCP and $\mathbf{\Phi}_1 = \mathrm{diag}\{e^{j\phi_{1,1}}, \cdots, e^{j\phi_{1,N}}\}$ stands for phase shifts of signals in RHCP.

Then, the wireless channel $\mathbf{G}_2$ between the RIS and the DP receive antennas can be formulated similarly as

$$\mathbf{G}_2 = (\mathbf{I}_N \otimes \mathbf{F}_r)\bar{\mathbf{Q}}_2 \mathbf{m}_s(\alpha_s), \tag{16}$$

where $\mathbf{F}_r$ is the XPI of the DP receive antennas in a similar form as $\mathbf{F}_t$ with χ_r^{-1} and φ_r. And $\mathbf{m}_s(\alpha_s)$ is expressed as

$$\mathbf{m}_s(\alpha_s) = \begin{bmatrix} \cos(\alpha_s) & -\sin(\alpha_s) \\ \sin(\alpha_s) & \cos(\alpha_s) \end{bmatrix} \mathbf{F}_{ris} \tag{17}$$

where α_s is the rotation angle.

We assume that $\bar{\mathbf{Q}}_2 \in \mathbb{C}^{2N_r\times 2N}$ is a random matrix under the DP setting of channel as

$$\bar{\mathbf{Q}}_2 = (\mathbf{\Delta}_2 \otimes \mathbf{I}_{N_r}) \odot ((\mathbf{C}_{r_p} \otimes \mathbf{\Gamma}_r)^{\frac{1}{2}} \mathbf{Z}_2 (\mathbf{C}_{s_p} \otimes \mathbf{\Gamma}_{ris}^{\frac{1}{2}})), \tag{18}$$

where $\mathbf{C}_{r_p}$ is the receive polarization correlation matrix as expressed by (7) with the receiver polarization correlation coefficient r_p. $\mathbf{Z}_2 \in \mathbb{C}^{2N_r\times 2N}$ is matrix whose entries are of i.i.d. $\sim \mathcal{CN}(0,1)$. The XPD of the channel between RIS and receiver is expressed as Δ_2 with λ_2 and β_2 in the same formulation as Δ_1. $\mathbf{\Gamma}_r \in \mathbb{C}^{N_r\times N_r}$ is spatial correlation matrix between RIS and receiver expressed as

$$\mathbf{\Gamma}_r^{(m,n)} = \begin{cases} \xi_r^{n-m}, m \leqslant n, \\ (\xi_r^{*})^{m-n}, m > n. \end{cases} \tag{19}$$

Besides, we assume that the RIS is located in the far-field at transceiver. Accordingly, the large-scale fading coefficients and the polarized rotation angle of RIS elements are constant. Then, the whole channel matrix $\mathbf{S}$ can be further restructured as

$$\mathbf{S} = D\frac{\zeta}{2}\sum_{n=1}^{N} e^{j(\omega_a+\phi_{n,0}))} \begin{bmatrix} 1 + e^{j(\phi_{n,0}-\phi_{n,1})} \\ 1 - e^{j(\phi_{n,0}-\phi_{n,1})} \end{bmatrix}, \tag{20}$$

where $\omega_a = \omega_t + \omega_s + \omega_r$ and D is given:

$$D = \begin{bmatrix} \cos(\alpha_a) & -\sin(\alpha_a) \\ \sin(\alpha_a) & \cos(\alpha_a) \end{bmatrix} \tag{21}$$

with $\alpha_a = \alpha_s + \alpha_t$. Then, we derive instantaneous SNR at the u-th chosen receive antennas of k-th elements as follows:

$$\gamma_{u,k} = \frac{\left|\sum_{n=1}^{N} \xi_{u,k} e^{j(\omega_t+\omega_s+\omega_r-\phi_{u,k})}\right|^2 E_s}{N_0} = \frac{N^2\xi^2 E_s}{N_0}. \tag{22}$$

III. Traditional Detector

A. ML Detector

The ML detector applied in the RIS-RPSSK scheme is given as

$$(\hat{u}, \hat{k}) = \arg \min_{u\in\mathbb{U}, k\in\mathbb{K}} \left\| y - N\sqrt{E_s}\zeta \mathbf{G}_2 \mathbf{\Theta}_{u,k} \mathbf{G}_1 \right\|^2, \tag{23}$$

where $\mathbb{U} = \{U_1, U_2, ..., U_{N_r}\}$ represents possible N_r receive antennas and $\mathbb{K} = \{K_1, K_2\}$ represents the polarization states of LHCP and RHCP.

B. Greedy Detector

The greedy detector based on the maximum instantaneous SNR at receiver is expressed as:

$$(\hat{u}, \hat{k}) = \arg \max_{u\in\mathbb{U} k\in\mathbb{K}} |y_{u,k}|^2. \tag{24}$$

where $\hat{u}, \hat{k}$ are the estimated receive antenna index.

IV. Proposed LSTM Detecter

An attention-LSTM based multi-label classifier is proposed for RIS-aided RSSK scehme, where LSTM plays a significant role in DL with the superiority of handling sequences and capturing long-term dependencies related to time varying.

1) Pre-processing: Before entering the LSTM neural network training, pre-processing is used to transform original data into a operative data set. A feature vector generator (FVG) needs to be introduced to preprocess the original data and convert multiple vectors/matrices of complex-valued IQ elements into real-valued vectors. The input will be extracted by using FVG, result can be shown as

$$\mathbf{E} = [|\mathbf{y}_{1\Re}|, \ldots, |\mathbf{y}_{2N_r\Re}|, |\mathbf{y}_{2N_r\Im}|, \ldots, |\mathbf{S}_{2N_r\Re}|, |\mathbf{S}_{2N_r\Im}|]^T, \tag{25}$$

2) LSTM model: In a traditional RNN, all w are identical. The network retains the input in memory and generates a new prediction that incorporates all previous inputs, typically employing a simple activation function such as tanh, which may result in gradient problem hindering the ability to effectively capture characteristics.

To address the issues, the LSTM network employs gates to control the addition or removal of information, which effectively processes sequential data and extracts temporal features, thereby overcoming the challenges associated with information transmission in long sequences and enabling long-term memory functionality. The LSTM unit structure is illustrated in Fig. 2. Specifically, an LSTM model consists of a cell state and three Sigmoid gate structures. The cell state serves as the core memory space of the LSTM model, evolving over time as information is recorded and updated through a gating mechanism implemented via the sigmoid function and dot multiplication. A sigmoid value of 0 indicates the discarding of information, while a value of 1 signifies complete

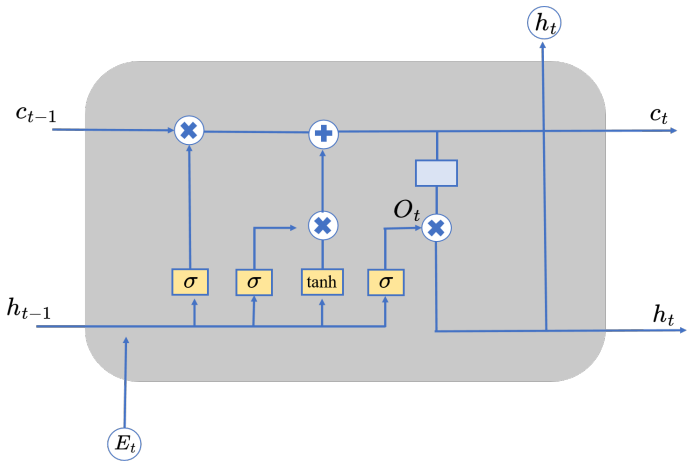

Fig. 2: LSTM unit structure

transmission. The gates are the forget gate f_t, input gatei_t, and output gate o_t at time t as

$$f_t = \sigma(W_{df}E_t + W_{hf}h_{t-1} + b_f) \tag{26}$$

$$i_t = \sigma(W_{di}E_t + W_{hi}h_{t-1} + b_i) \tag{27}$$

$$o_t = \sigma(W_{do}E_t + W_{ho}h_{t-1} + b_o), \tag{28}$$

where E_t is input at time t, σ represents the sigmoid activation function. $\mathbf{h}_{t-1}$ is the previous time hidden state. $W_{df}, W_{hf}, W_{di}, W_{hi}, W_{do}, W_{ho}$ are weight parameters and $\mathbf{b}_i, \mathbf{b}_f, \mathbf{b}_o$ are bias parameters.

The network executes memory cell and hidden state functions by leveraging the input gate at time t as

$$g_t = \tanh(W_{dg}E_t + W_{hg}h_{t-1} + b_g) \tag{29}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \tag{30}$$

$$h_t = o_t \odot tanh(c_t), \tag{31}$$

After bitwise multiplication operation, the output of the Sigmoid function is a vector with each bit value ranging from 0 to 1. Then f_t is used to multiply the vector c_{t-1}, effectively reducing the value of each position in the c_{t-1} vector to different degrees.

The input vector is transformed into a three-dimensional tensor to serve as input data for the LSTM layer. The output generated from three LSTM hidden layers is utilized as input for the fully connected (FC) layer. The sigmoid activation function is applied in this layer, defined as $f_{softmax}(x) = \frac{e^{x_i}}{\sum_{j=1}^{K} e^{x_j}}$, which effectively maps elements of the output vector to the range of (0, 1), facilitating a probabilistic interpretation. The output vector e_l from the FC layer can be mathematically expressed as:

$$e_l = f_{softmax}(W_s \times H_t + b_l), \tag{32}$$

where H_t is the hidden state from the LSTM layers, and b_l is the bias vector. Thus, the output bit b_g emerges as a probabilistic indicator, facilitating effective binary classification and subsequent processing in later network layers.

3) LSTM improvements: Following the construction of the basic LSTM model, further optimizations will be implemented. o_t refers to the state output of the decoder at time t and o_{t-1} denotes the state output at last time slot. By defining d_{t-1} is the label at time $t-1$, o_t can be expressed as

$$o_t = f(o_{t-1}, d_{t-1}, \sum_{j=1}^{Tx} a_{tj}q_j), \tag{33}$$

where q_j refers to the output of the j th input in the encoder, and $a_{tj} = \frac{e^{e_{tj}}}{\sum_{k=1}^{Tx} e^{e_{tk}}}$ is a weight as the degree to which the current decoder is aligned with the j-th input. This formula is to obtain the conditional probability P(a|e). e_{tj} can be calculated as

$$e_{tj} = g(o_{t-1}, q_j) = V \cdot \tanh(W \cdot q_j + U \cdot o_{t-1} + b), \tag{34}$$

where g can be approximated by a neural network to calculate the relationship score between o_{t-1} and q_j. The large score means that the degree of attention is high and the attention distribution can be more concentrated on the input.

This improved model incorporates an attention mechanism. The overall model structure is shown as Fig. 3.

4) Training: The selection of specific parameters is detailed in the Table.I. By employing the ReLU activation function, deep supervised networks can be trained. The final layer uses the softmax activation function to output a probability distribution. The loss function selected is categorical cross-entropy. Nadam can be employed to minimize this loss. In addition to features, training the network requires input labels. To train our network, we employ a randomly generated dataset consisting of 10^6 samples. The att-LSTM detector demonstrates significant accuracy after 50 epochs with sufficient iterations. We allocated 70% of the dataset for training, while 30% was reserved for validation and testing, enabling a comprehensive evaluation of the model's performance.

TABLE I: CONFIGURATION AND PARAMETERS FOR LSTM

Parameters	Value	Parameters	Value
Input nodes	$4N_r$	Learning rate	0.001
Hidden nodes	[256 128]	Optimizer	Nadam
Output nodes	$2N_r$	Epoch	50
Output layer activation	softmax	Loss function	Cross-entropy

5) Prediction: After completing pre-processing and model training, we use the model for prediction of actual transmission. In the j-th time slot, the transmitted signal can be predicted by F_j.

The minimum distance index u and k is obtained by the Euclidean distance between the received signal and the predicted symbol vector. The index corresponding to the minimum distance represents the estimated symbol:

$$\hat{l} = \arg \min_{l \in \{1,\ldots L_{all}\}} \left\| \mathbf{y}^{(k)} - (\mathbf{S})^{(k)}_{(J,k)_l} \right\|_F^2. \tag{35}$$

Then $\hat{l}$-th spatial and polarization indices are the final solution.

V. SIMULATION RESULTS

This section presents a comparative analysis of the BER performance of the proposed RIS-RPSSK scheme and the RIS-

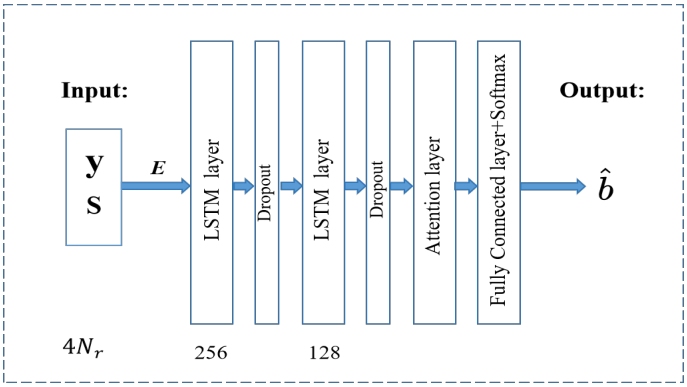

Fig. 3: Structure of the LSTM detector

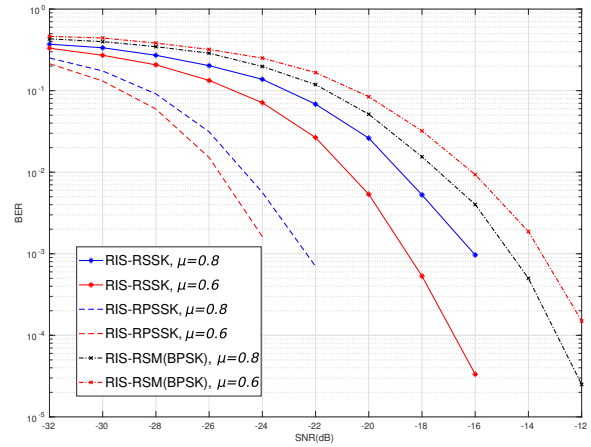

Fig. 4: BER Performances comparison between RIS-RSSK, RIS-RSM and RIS-RPSSK schemes versus different spatial correlations $\mu = 0.6, 0.8$.

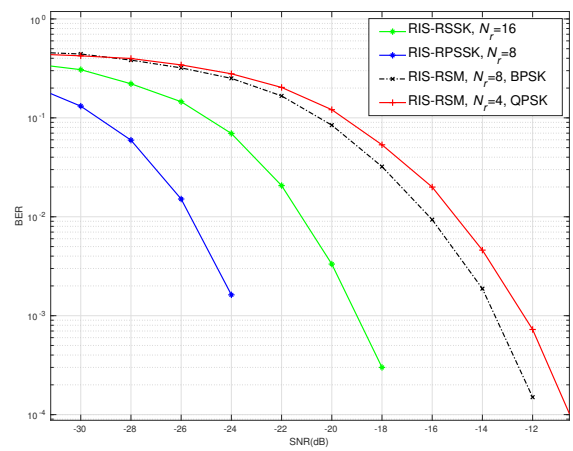

Fig. 5: BER Performances comparison between RIS-RPSSK scheme, RIS-RSSK and two RIS-RSM schemes with 4 bpcu and $\mu = 0.6$.

RSSK scheme. Moreover, the performance of the LSTM detector are evaluated by comparing it with the ML detector and the greedy detector, which is called GD in figures for convenience, utilizing TensorFlow in Python. The simulations are conducted on an 13-th Gen Intel Core i7-13700KF (24CPUs), NVIDIA GeForce RTX 3080 Ti GPU and 32GB RAM.

The polarization correlation factors are chosen as $s_p = t_p = r_p = 0.1$. The XPI of the system is $\chi_t^{-1} = \chi_{ris}^{-1} = \chi_r^{-1} = 25$ dB and the all phase offsets are fixed as $\varphi_t = \varphi_{ris} = \varphi_r = 0$. We generate the angles of $\alpha_t, \alpha_s \in \mathcal{U}[0, 2\pi]$ randomly with a uniform distribution.

For the comparison with RIS-RSSK scheme and RIS-RSM (BPSK) scheme, we first compare the BER under different spatial correlation level as $\mu = 0.6, 0.8$ under $N_r = 8$, $N = 32$. Fig. 4 illustrates that the proposed RIS-RPSSK significantly outperforms the RIS-RSSK scheme and RIS-RSM (BPSK) scheme in terms of BER performance under conditions of high spatial correlation while achieving a higher data rate. For a BER of 10^{-3}, the RIS-RPSSK offers approximately 4 dB of SNR advantage over the RIS-RSSK scheme and 9 dB over the RIS-RSM (BPSK) scheme. Additionally, it is noted that the proposed RIS-RPSSK scheme experiences a minor degradation of about 1 dB, attributed to cross-talk in polarized channels. The proposed RIS-RPSSK scheme employs two polarization states, allowing the receiver to concurrently capture signals from both polarization components. This dual-polarization approach enhances signal distinction and consequently mitigates multipath propagation effects.

Furthermore, we utilize a configuration with an equivalent data rate of 4 bits per channel use (bpcu). As depicted in Fig. 5, the Bit Error Rate (BER) of the proposed RIS-RPSSK scheme under the initial conditions of $N_r = 8$ surpasses that of the RIS-RSSK scheme with $N_r = 16$ and RIS-RSM scheme with $N_r = 4$, QPSK or $N_r = 8$, BPSK. At the BER of 10^{-3}, the proposed RIS-RPSSK scheme obtain about 4 dB, 9dB and 10 dB gains over the setting of RIS-RPSSK and RIS-RSM with two configurations. This enhancement enables more efficient utilization of spectral resources and facilitates a more compact antenna array design.

Additionally, Fig. 6 illustrates the BER performance of the proposed Att-LSTM detector in comparison to the ML detector and greedy detector for varying N. The results indicate that the Att-LSTM detector offers a notable improvement in BER performance, achieving an enhancement of approximately 3 dB. Additionally, similar to the other detectors, an increasing number of N also brings a beneficial effect on the BER performance in the Att-LSTM detector. It is evident that the BER performance of Att-LSTM detector closely approach that of ML detector across N cases. In Fig. 7, the impact of the number of N_r is analysed on the BER performance of the RIS/RSKK for each detector. The proposed Att-LSTM detector is also capable of providing better BER performance than the greedy detector.

The order complexities of Att-LSTM detector can be caluculated as $\mathcal{O}(f_1 f_2)$ (f_1, f_2 are neurons of layers) which mainly depend on the fixed parameters thus having lower computation.

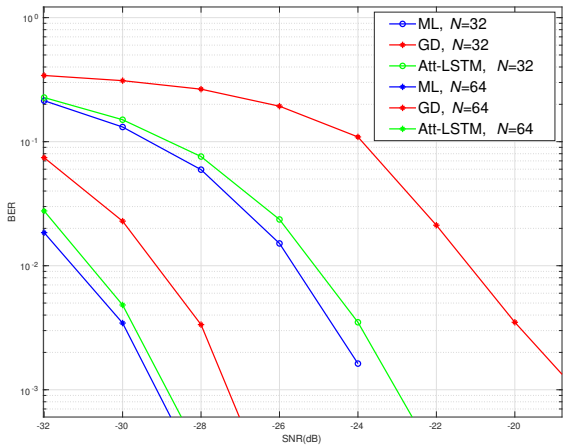

Fig. 6: BER performance comparison of the ML detector, greedy detector, Att-LSTM detector for RIS-RPSSK with $N_r = 8$ for different N.

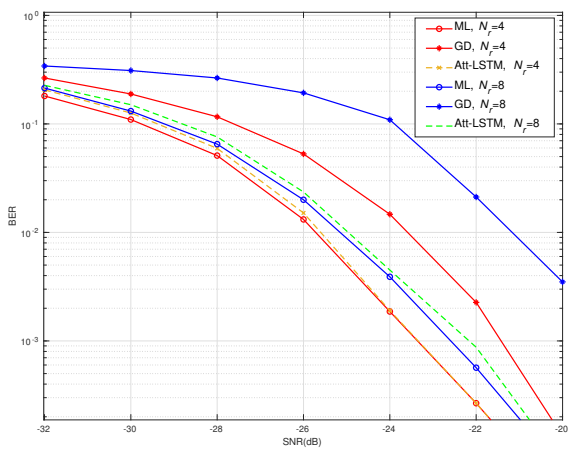

Fig. 7: BER performance comparison of the ML detector, greedy detector, Att-LSTM detector for RIS-RPSSK with $N = 32$ for different N_r.

This makes the Att-LSTM detector a more suitable option for future scenarios where computational efficiency is crucial due to the exponential growth in complexity associated with ML detector $\mathcal{O}(UKN)$.

VI. Conclusion

In this paper, we propose the RIS-aided RPSSK scheme utilizing dual-polarization antennas and dual-polarization elements to enhance SE. We establish channel models that account for all spatial and polarization impairments. Additionally, we introduce the Att-LSTM detector to address the limitations of traditional detectors. Results demonstrate that the Att-LSTM techniques outperform others while maintaining lower complexity. Overall, the RIS-RPSSK framework presents a promising candidate applied in future wireless communication systems.

References

[1] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.

[2] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.

[3] C. You, B. Zheng, and R. Zhang, "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2604–2620, Nov. 2020.

[4] X. G. Zhang et al., "Polarization-controlled dual-programmable metasurfaces," *Adv. Sci.*, vol. 7, no. 11, pp. 1–10, Apr. 2020.

[5] Y. Han et al., "Dual-Polarized RIS-Assisted Mobile Communications," *IEEE Trans. Wireless Communications*, vol. 21, no. 1, pp. 591-606, Jan. 2022.

[6] A. E. Canbilen, E. Basar, and S. S. Ikki, "Reconfigurable intelligent surface-assisted space shift keying," in *IEEE wireless communications letters*, vol. 9, no. 9, pp. 1495–1499, 2020.

[7] J. Zhang, Y. Wang, J. Zhang, and L. Ding, "Polarization shift keying (polarSK): System scheme and performance analysis," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10139–10155, Nov. 2017.

[8] S. Dhanasekaran, "Space-polarization shift keying modulation for MIMO channels," *Wireless Pers. Commun.*, vol. 86, no. 3, pp. 1509–1539, 2016.

[9] J. Liu and M. D. Renzo, "Data-driven and Model-driven Deep Learning Detection for RIS-aided Spatial Modulation," *Proc. IEEE 5G World Forum (5GWF)*, Montreal, QC, Canada, Oct. 2021, pp. 88-92.

[10] Salan, F. Bayar, H. Ilhan and E. Aydin, "Deep Learning Based Detection on RIS Assisted RSM and RSSK Techniques," *2023 IEEE Globecom Workshops (GC Wkshps)*, Kuala Lumpur, Malaysia, 2023, pp. 1153-1157.

[11] H. Zhang, L. Zhang, Y. Jiang and Z. Wu, "LSTM and ResNets Deep Learning Aided End-to-End Intelligent Communication Systems," *2021 2nd Information Communication Technologies Conference (ICTC)*, Nanjing, China, 2021, pp. 156-160.

[12] 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.

[13] Emil Bjornson, Jakob Hoydis and Luca Sanguinetti (2017), "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency", Foundations and Trends® in *Signal Processing*: Vol. 11, No. 3-4, pp 154–655. DOI: 10.1561/2000000093.

[14] E. Ibrahim, R. Nilsson and J. van de Beek, "Joint polarization and spatial modulation using reconfigurable intelligent surface," in *Proc. IEEE WCNC*, Dubai, United Arab Emirates, Apr. 2024, pp. 1-6.

[15] N. Bouhlel, M. Saad, F. Bader and J. Palicot, "Dual-polarized generalized spatial modulation for wireless terabit system", *Proc. IEEE 31st Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, pp. 1-7, Aug. 2020.

[16] C. Zhang, Y. Liu, B. K. Ng, and C.-T. Lam, "RIS-assisted differential transmitted spatial modulation design," *Signal Process.*, vol. 2024, p. 109767, Nov. 2024.