Improved ADMM Signal Detection Algorithm for Multi-layer RIS Transmitter

Zhiwen Bai[1], Xinbo Xu[2], Shuyi Chen[1(✉)], Weixiao Meng[1], Sebastián E. Godoy[3], and Gabriel Saavedra[3]

[1] School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China
chenshuyitina@163.com
[2] Beijing Institute of Spacecraft System Engineering, Beijing, China
[3] Universidad de Concepción, Concepció, Chile

Abstract. Among numerous emerging technologies targeting future wireless communication scenarios and demands, Reconfigurable Intelligent Surface stand out as a research hotspot due to their advantages in controlling the wireless communication environment with lower cost and energy consumption. This article uses double-layer RIS to replace traditional components as multiple-in multiple-out transmitters, and utilizes the diffraction between RIS layers to achieve signal level error control encoding. Under this model, this article simplifies the Alternating Direction Method of Multipliers algorithm using the proximal gradient method, and optimizes the algorithm by introducing parameters and training with deep learning. This reduces the complexity of the algorithm to a certain extent and improves its performance.

Keywords: Reconfigurable Intelligence Surface · MIMO Detection · Deep Learning · Neural Network · Alternating Direction Method of Multipliers

1 Introduction

RIS is composed of passive components, which can adjust the phase shift of a large number of quasi passive and low-cost reflective components appropriately by changing the electrical parameters of the structural units, thereby reconfiguring the wireless channel environment [1] and obtaining the ability to manipulate the wireless channel. In addition to its role in auxiliary communication, RIS can also serve as an alternative to specific traditional modules in transmitters and receivers [2]. For example, RIS can

This paper is supported by the National Science Fund for Young Scholars No. 62201176, National Natural Science Foundation of China under Grand No.62431009 and No. 62271168, the Key R&D Plan of Heilongjiang Province of China under Grand No. JD22A001, the Young Elite Scientist Sponsorship Program by CAST No. YESS20210339, the Natural Science Foundation of Heilongjiang Province of China No. YQ2023F005, and the Heilongjiang Province Outstanding Doctoral Dissertation Funding Project No. LJYXL2022-049.

H.-H. Chen and W. Meng (Eds.): WiSATS 2024, LNICST 606, pp. 277–291, 2025.
https://doi.org/10.1007/978-3-031-86203-8_22

achieve modulation function by inputting a data source into a controller and mapping the data source into signals of different amplitudes and phases [3]. In addition, RIS supports multi antenna transmission by using each element atom as the transmitting antenna, where the precoding matrix is input into the controller to achieve simulated beamforming [4].

At present, some research has shown that RIS can directly achieve traditional digital modulation, including frequency shift keying, phase shift keying, and quadrature amplitude modulation [5]. The author of [6] proposes a simplified BFSK implementation architecture based on RIS. According to the methods designed in [7–9], multiple modulation schemes such as QPSK, 8PSK, and 16QAM can be directly implemented using RIS. The author of [10] proposed a method to extend RIS based wireless transmission from Single Input Single Output (SISO) to Multiple Input Multiple Output (MIMO).

MIMO, as one of the most fundamental technologies in wireless communication systems, can significantly improve spectrum efficiency and energy efficiency by deploying a large number of antennas on users and base stations, endowing the system with the ability to achieve high-speed and low latency transmission. MIMO technology effectively utilizes spatial resources and improves spectral efficiency through the use of multiple antennas, not only improving data transmission speed and quality, but also enhancing signal coverage and system reliability. Currently, research on MIMO signal detection algorithms is mainly divided into two categories: optimal and suboptimal [11]. The Maximum Likelihood Detection [12] algorithm is known as the optimal algorithm for MIMO signal detection, which selects the detection signal with the smallest candidate signal norm by traversing all sets of transmitted signals. However, in this case, the detection algorithm complexity is high. The suboptimal algorithm includes two detection algorithms: linear and nonlinear. Linear detection algorithms include detection algorithms such as Zero Forcing and Minimum Mean Square Error. Nonlinear detection algorithms, such as Sphere Decoding and Serial Interference Cancellation detection.

With the development of deep learning and artificial intelligence [13], deep learning has also been applied to large-scale MIMO signal detection in pursuit of performance improvement. The main implementation methods of deep learning for signal detection in MIMO are mainly divided into two types: data-driven methods and model-driven methods [14]. In order to combine the advantages of the two implementation methods, deep unfolding is proposed, and its core idea is to unfold each iteration of an iterative algorithm into a layer of neural network. Therefore, by utilizing deep neural network-based methods, many MIMO detection signal processing algorithms have been designed. A algorithm called DetNet in reference [15] combines fully connected neural networks and MIMO detectors. DetNet is formed by unfolding a projection gradient descent algorithm for ML optimization and shows better performance than AMP and SDR detectors, but it takes tens of hours of offline training process. Compared to traditional methods, DetNet can have better performance. However, its drawbacks are also obvious, which are higher time complexity and lack of deployability compared to traditional methods. Reference [16] proposes a deep learning based OAMP Net2 detection algorithm based on the OAMP algorithm, which introduces training parameters and improves the accuracy of nonlinear estimation in the OAMP algorithm. Reference [17] constructed an ADMM Net signal detection network model based on deep learning using the ADMM.

However, these detection algorithms are all based on traditional transmission and reception systems. With the implementation of new transmitter hardware architectures and prototypes based on RIS, there has not been much research on whether these detection algorithms can adapt to RIS based transmitters, and what impact RIS transmission and reception systems will have on them.

This article is based on the study of signal level error control encoding using diffraction between multiple RIS layers [18]. The use of dual layer RIS as a transmitter and traditional detection algorithms for signal detection has been proven to be compatible with existing MIMO traditional detection algorithms. On this basis, this article verifies its compatibility with deep learning based detection algorithms, referencing the ideas of ADMM Net. Based on the use of RIS transmitters, this article further introduces parameters to improve the relevant performance of the algorithm to a certain extent.

The organizational structure of this article is as follows: In the second section, we briefly introduces the MIMO system model based on RIS transmitters. In the third section, a new MIMO detection scheme based on ADMM is proposed. The fourth section conducts numerical comparisons and summarizes this article in the fifth section.

2 System Model

2.1 Diffraction Channel Coding Based on RIS

The RIS layer can serve as the transmitting antenna for MIMO systems, and by placing the RIS layer in front of the MIMO receiving antenna, certain types of signal processing can be achieved. Therefore, a basic transceiver structure based on RIS can be established, where the transmitter consists of three parts: modulation based on RIS, inter layer transmission based on RIS, and beamforming based on RIS. The first part consists of an RF signal generator and a RIS layer containing M electromagnetic units to achieve signal modulation; The second part is the propagation of signals between two layers of RIS; The third part also consists of a RIS layer containing N electromagnetic units, achieving beamforming based on RIS.

Due to the fact that the distance between adjacent metasurfaces is on the same wavelength scale, it is not possible to model them using large-scale wireless channel modeling methods. However, this distance satisfies the conditions for near-field propagation. In this case, the propagation between adjacent layers of RIS can be modeled using the Rayleigh Summerfeld diffraction equation to construct an equivalent propagation matrix.

The propagation coefficient between two layers is

$$\mathbf{G} = \begin{pmatrix} g_{1,1} & \cdots & g_{1,N} \\ \cdots & g_{m,n} & \cdots \\ g_{M,1} & \cdots & g_{M,N} \end{pmatrix} \tag{1}$$

$g_{\mathrm{m,n}}$ is the propagation factor from the first layer RIS unit to the second layer RIS unit established through the Rayleigh Sammerfeld diffraction equation

$$g_{m,n} = \frac{z_{1,k,l} - z_{2,u,v}}{r_{k,l,u,v}^2} \left(\frac{1}{2\pi r_{k,l,u,v}} + \frac{1}{j\lambda} \right) exp\left(\frac{j2\pi r_{k,l,u,v}}{\lambda} \right) \tag{2}$$

The position of unit m in the first layer is $x_{1,k,l}$, $y_{1,k,l}$, $z_{1,k,l}$, and the position of unit n in the second layer is $x_{2,u,v}$, $y_{2,u,v}$, $z_{2,u,v}$, $r_{k,l,u,v} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$ is the distance between the two units.

If the number of electromagnetic units in the first layer RIS is smaller than that in the second layer RIS, the output signal dimension, i.e. The received signal in the second layer, increases. In fact, this is equivalent to adding an additional additional signal, and the added additional signal is a linear combination of the original signal.

Firstly, represent the input data as

$$\boldsymbol{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_M \end{pmatrix} \tag{3}$$

By inputting data information into FPGA, signal modulation can be achieved by adjusting the behavior of RIS units. The modulated signal is represented as:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix} \tag{4}$$

The signal after double-layer RIS encoding can be represented as:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \tag{5}$$

Among $\mathbf{v} = \mathbf{G}\mathbf{x}$.

We can achieve optimal code performance by maximizing the minimum code distance, which is represented as:

$$D\left(\vec{d}_i, \vec{d}_j\right) = \left\| v_i - \mathbf{v}_j \right\| \tag{6}$$

Therefore, constructing an optimization problem model is

$$\begin{aligned} &\max_{\mathbb{G}} \min_{i,j \in [0,2^M], i \neq j} \left\{ D(\overrightarrow{d_i}, \overrightarrow{d_j}) \right\} \\ &\text{s.t. } d_c \in [\frac{\lambda}{6}, \frac{\lambda}{4}] \\ &d_l > \lambda \end{aligned} \tag{7}$$

In Eq. (7), d_c represents the distance between two electromagnetic units, and d_l represents the distance between two RIS layers. By solving the optimization problem, the generation matrix can be obtained.

The minimum code distance determines the encoding ability, and the larger the minimum code distance, the better the encoding performance. Based on the minimum code

distance, error patterns can be detected and the number of errors that can be corrected can be determined. Traditional channel coding is based on the Galois domain and calculates code distance through binary operations. In contrast, the diffraction channel coding based on RIS directly operates on the signal, and the code distance is calculated as the Euclidean distance in the complex domain, rather than the Galois domain. In addition, since the signal is continuous in amplitude and phase, detectable error modes cannot be determined based on the minimum code distance. In traditional channel coding, if a bit changes from 0 to 1 or from 1 to 0, an error will occur. However, in RIS based transmitters, channel effects cause both amplitude and phase changes in the signal, allowing for any type of change, thus improving the performance of communication systems and reducing channel effects to a certain extent.

If the number of electromagnetic units in the first layer RIS is different from that in the second layer RIS, the dimension of the transformation matrix between the two layers of RIS obtained will change. Therefore, this article assumes that the number of electromagnetic units in the first layer is smaller than the number of electromagnetic units in the second layer, and the output signal dimension increases, that is, the received signal in the second layer increases. In fact, this is equivalent to adding an additional additional signal, and the added additional signal is a linear combination of the original signal.

The transmission effect of signals passing through RIS can be represented in matrix form, so signals passing through multiple layers of metasurfaces can be mathematically equivalent to the product of multiple matrices. On this basis, a transmitter based on multi-layer metasurfaces can be built.

2.2 MIMO System Based on RIS Transmitter

A MIMO system based on RIS transmitters considers a traditional receiver consisting of a dual layer RIS transmitter and an N antenna. The dual layer RIS transmitter can be equivalent to a traditional transmitter with M transmitting antennas, which sends signal vectors to the receiver through wireless channels, as shown in Fig. 1:

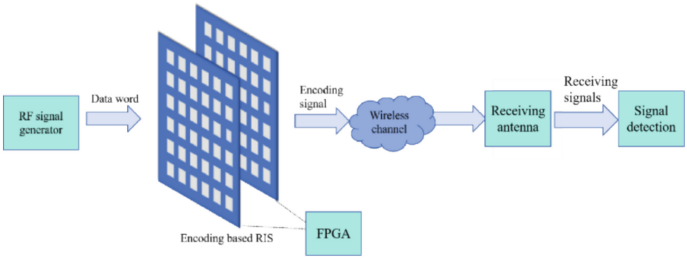

Fig. 1. MIMO system based on RIS transmitter

Therefore, MIMO systems based on RIS transmitters can also be written as:

$$\overline{\mathbf{y}}=\overline{\mathbf{H}}_1\mathbf{G}\overline{\mathbf{x}}+\overline{\mathbf{n}} \tag{8}$$

Among them $\overline{\mathbf{y}}$ is the received signal vector, $\overline{\mathbf{x}}$ is the transmitted signal vector, $\overline{\mathbf{n}}$ is the additive Gaussian white noise, which mean is 0 and the variance is σ^2. $\overline{\mathbf{H}}_1$ representing the Rayleigh fading channel matrix, $\mathbf{G}$ representing the generation matrix between two layers of RIS. This article assumes that the channel matrix $\overline{\mathbf{H}}_1$ is a known quantity, and the generation matrix between the two layers of RIS is also known. Therefore, $\overline{\mathbf{H}}_1\mathbf{G}$ in Eq. (8) can be rewritten to represent the entire channel, represented by $\overline{\mathbf{H}}$.

$$\overline{\mathbf{y}}=\overline{\mathbf{H}}\overline{\mathbf{x}}+\overline{\mathbf{n}} \tag{9}$$

To obtain an equivalent real channel model, the receiver can rearrange Eq. (9) and transform its complex form into real form, represented as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{10}$$

$$\mathbf{y} = \begin{bmatrix} \Re(\overline{\mathbf{y}}) \\ \Im(\overline{\mathbf{y}}) \end{bmatrix}, \mathrm{x} = \begin{bmatrix} \Re(\overline{\mathrm{x}}) \\ \Im(\overline{\mathrm{x}}) \end{bmatrix}, \mathrm{n} = \begin{bmatrix} \Re(\overline{\mathrm{n}}) \\ \Im(\overline{\mathrm{n}}) \end{bmatrix}, \mathrm{H} = \begin{bmatrix} \Re(\overline{\mathrm{H}}) & -\Im(\overline{\mathrm{H}}) \\ \Im(\overline{\mathrm{H}}) & \Re(\overline{\mathrm{H}}) \end{bmatrix} \tag{11}$$

wherein $\mathbf{y} \in \mathbb{R}^{\mathrm{N}}$, $\mathbf{x} \in \mathrm{S}^{\mathrm{M}}$, $\mathbf{n} \in \mathbb{R}^{\mathrm{N}}$, $\mathbf{G} \in \mathbb{R}^{\mathrm{N}\times\mathrm{M}}$, $\mathbf{H} \in \mathbb{R}^{\mathrm{N}\times\mathrm{N}}$, $\mathrm{N} = 2\overline{\mathrm{N}}$, $\mathrm{M} = 2\overline{\mathrm{M}}$。

3 Signal Detection Algorithm

In this section, we propose a MIMO detection algorithm based on deep unfolding. We first simplify the existing ADMM algorithm using the proximal gradient method, and then introduce parameters for training.

3.1 MIMO Detection Algorithm Based on ADMM

In the ADMM detection algorithm, the objective function for reconstructing the MIMO signal detection problem is:

$$\begin{gathered} \min_{\mathbf{x}\in\mathbb{R}^M, \mathbf{z}\in\mathbb{R}^N} I_{S^M}(\mathbf{x}) + \frac{1}{2\lambda}\|\mathbf{z}\|_2^2 \\ \text{s.t.}\mathbf{z} = \mathbf{H}\mathbf{x} - \mathbf{y} \end{gathered} \tag{12}$$

In Eq. (12), $\lambda > 0$, $\mathbf{x}$ is the received signal to be restored. The received signal can be divided into two groups of variables, $\mathbf{x}$ variables and $\mathbf{z}$ variables, $I_{\mathrm{S}^{\mathrm{M}}}(\mathbf{x})$ is the indicator functions of convex set modulation constellation maps. When $\mathrm{x} \in \mathrm{S}^{\mathrm{M}}$, the value of is 0; At time $\mathrm{x} \notin \mathrm{S}^{\mathrm{M}}$, the value of is infinite.

The expression for its augmented Lagrangian function is:

$$\mathrm{L}(\mathbf{x}, \mathbf{z}, \mathbf{u}) = I_{\mathrm{S}^{\mathrm{M}}}(\mathbf{x}) + \frac{1}{2\lambda}\|\mathbf{z}\|_2^2 + \rho\mathbf{u}^{\mathrm{T}}(\mathbf{H}\mathbf{x} - \mathbf{y} - \mathbf{z}) + \frac{\rho}{2}\|\mathbf{H}\mathbf{x} - \mathbf{y} - \mathbf{z}\|_2^2 \tag{13}$$

Among them $\mathbf{u}$ are dual variables.

By applying the dual ascent method, the L(**x**, **z**, **u**) can minimize the impact on the original variable (**x**, **z**); Regarding the maximization of dual variables. The resulting iterative process is:

$$\begin{aligned}
\mathbf{x}_{k+1} &= \arg\min_{\mathbf{x}\in\mathbb{R}^M} I_{S^M}(\mathbf{x}) + \frac{\rho}{2}\|\mathbf{H}\mathbf{x} - \mathbf{y} - \mathbf{z}_K + \mathbf{u}_k\|_2^2 \\
\mathbf{z}_{k+1} &= \arg\min_{\mathbf{z}\in\mathbb{R}^M} \frac{1}{2\lambda}\|\mathbf{z}\|_2^2 + \frac{\rho}{2}\|\mathbf{H}\mathbf{x}_{k+1} - \mathbf{y} - \mathbf{z}_K + \mathbf{u}_k\|_2^2 \\
&= \frac{\lambda\rho}{1+\lambda\rho}(\mathbf{H}\mathbf{x}_{k+1} - \mathbf{y} - \mathbf{z}_K + \mathbf{u}_k) \\
\mathbf{u}_{k+1} &= \mathbf{H}\mathbf{x}_{k+1} - \mathbf{y} - \mathbf{z}_{K+1} + \mathbf{u}_k
\end{aligned} \tag{14}$$

G. Franca proposed an optimized variant of the ADMM algorithm, which improves convergence speed and stability by introducing heavy ball acceleration and relaxation parameters. From this, the ADMM (R-HB-ADMM) for relaxation and heavy sphere acceleration was derived [19]:

$$\begin{aligned}
\mathbf{x}_{k+1} &= \arg\min_{\mathbf{x}\in\mathbb{R}^M} I_{S^M}(\mathbf{x}) + \frac{\rho}{2}\left\|\mathbf{H}\mathbf{x} - \mathbf{y} - \hat{\mathbf{z}}_K + \hat{\mathbf{u}}_k\right\|_2^2 \\
\mathbf{z}_{k+1} &= \arg\min_{\mathbf{z}\in\mathbb{R}^M} \frac{1}{2\lambda}\|\mathbf{z}\|_2^2 + \frac{\rho}{2}\left\|\alpha(\mathbf{H}\mathbf{x}_{k+1} - \mathbf{y}) + (\mathbf{1}-\alpha)\hat{\mathbf{z}}_K + \hat{\mathbf{u}}_k\right\|_2^2 \\
&= \frac{\lambda\rho}{1+\lambda\rho}(\alpha(\mathbf{H}\mathbf{x}_{k+1} - \mathbf{y}) + (\mathbf{1}-\alpha)\hat{\mathbf{z}}_K + \hat{\mathbf{u}}_k) \\
\mathbf{u}_{k+1} &= \hat{\mathbf{u}}_k + \alpha(\mathbf{H}\mathbf{x}_{k+1} - \mathbf{y}) + (\mathbf{1}-\alpha)\hat{\mathbf{z}}_K - \mathbf{z}_{k+1} \\
\hat{\mathbf{z}}_{k+1} &= \mathbf{z}_{k+1} + \gamma(\mathbf{z}_{k+1} - \mathbf{z}_K) \\
\hat{\mathbf{u}}_{k+1} &= \mathbf{u}_{k+1} + \gamma(\mathbf{u}_{k+1} - \mathbf{u}_K)
\end{aligned} \tag{15}$$

wherein $\gamma = 1 - \mu$, $\alpha \in (0, 2)$.

Due to α, the convergence speed is accelerated, and at $\alpha = 1, \gamma_k = 0$ that time, Eq. (14) is equal to Eq. (15), which is a traditional ADMM detection. When the objective function is convex, the convergence of the R-HB-ADMM algorithm shows a linear trend, so it can quickly approximate the optimal solution.

From the **x** update formula in Eq. (15), it can be seen that as the scale of MIMO systems expands, the computational complexity of the R-HB-ADMM algorithm is significant. Therefore, in order to further improve signal detection performance and adapt to future technological developments, it is necessary to further adjust the algorithm to reduce its computational complexity.

We use the proximal gradient method for simplification, and the function expression of the proximal gradient method is:

$$\mathrm{f}(x) \le \hat{\mathrm{f}}(x; x_k) \equiv \mathrm{f}(x_k) + \nabla\mathrm{f}(x_k)^{\mathrm{T}}(x - x_k) + \frac{1}{2\delta_k}\|x - x_k\|_2^2 \tag{16}$$

If $f(\mathbf{x}) = \frac{1}{2}\left\|\mathbf{H}\mathbf{x} - \mathbf{y} - \hat{\mathbf{z}}_K + \hat{\mathbf{u}}_k\right\|_2^2$, the iterative formula for the simplified x can be obtained as:

$$\mathbf{x}_{k+1} = \arg\min_{\mathbf{x}\in\mathbb{R}^M} I_{S^M}(\mathbf{x}) + \rho\nabla h(\mathbf{x}_k)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}_k) + \frac{\rho}{2\delta_k}\|\mathbf{x} - \mathbf{x}_k\|_2^2$$

$$
\begin{aligned}
&= \arg\min_{\mathbf{x}\in\mathbb{R}^{M}} I_{S^{M}}(\mathbf{x}) + \frac{\rho}{2\delta_k}\|\mathbf{x}-\mathbf{x}_k+\delta_k\nabla h(\mathbf{x}_k)\|_2^2 \\
&= \prod\nolimits_{S^{M}}(\mathbf{x}_k-\delta_k\nabla h(\mathbf{x}_k)) \\
&= \prod\nolimits_{S^{M}}(\mathbf{x}_k-\delta_k\mathbf{H}^{T}(\mathbf{H}\mathbf{x}_k-\mathbf{y}-\hat{\mathbf{z}}_K+\hat{\mathbf{u}}_k))
\end{aligned} \tag{17}
$$

In the formula,$\delta_k \le \frac{1}{L_f}$, L_f is the maximum eigenvalue of $\mathbf{H}^{T}\mathbf{H}$, and it is hoped that it can take the smallest possible value.

3.2 Improved ADMM Detection Algorithm Based on Deep Unfolding

On the basis of using the proximal gradient descent method for simplification, in order to further reduce computational complexity, parameter terms can be added and neural networks can be used to further simplify the proposed ADMM detection algorithm, thereby making the ADMM algorithm have faster convergence speed and better performance. Considering that different μ values in γ in the R-HB-ADMM algorithm will have different effects on the results, it will also be used γ as a training parameter to obtain its optimal value. In addition, introduce residuals $\mathbf{e}_{k+1}$, accelerated residuals $\hat{\mathbf{e}}_{k+1}$, cumulative errors $\mathbf{v}_{k+1}$, and accelerated cumulative errors $\hat{\mathbf{v}}_{k+1}$. Its expression is:

$$
\begin{aligned}
\mathbf{e}_{k+1} &= \mathbf{H}\mathbf{x}_{k+1}-\mathbf{y} \\
\hat{\mathbf{e}}_{k+1} &= \mathbf{e}_{k+1}+\psi_k(\mathbf{e}_{k+1}-\mathbf{e}_k) \\
v_{k+1} &= \alpha\mathbf{e}_{k+1}+(1-\alpha)\hat{z}_k+\hat{u}_k \\
\hat{\mathbf{v}}_{k+1} &= \mathbf{v}_{k+1}+\gamma_k(\mathbf{v}_{k+1}-\mathbf{v}_k)
\end{aligned} \tag{18}
$$

By using the variables introduced in (17), the following equation can be obtained:

$$
\begin{aligned}
\mathbf{z}_{k+1} &= \beta\mathbf{v}_{k+1} \\
\hat{\mathbf{z}}_{k+1} &= \beta\hat{\mathbf{v}}_{k+1} \\
\mathbf{u}_{k+1} &= (1-\beta)\mathbf{v}_{k+1} \\
\hat{\mathbf{u}}_{k+1} &= (1-\beta)\hat{\mathbf{v}}_{k+1} \\
\mathbf{v}_{k+1} &= \alpha\mathbf{e}_{k+1}+(1-\alpha\beta)\hat{\mathbf{v}}_k
\end{aligned} \tag{19}
$$

Among them, $\beta = \frac{\lambda\rho}{1+\lambda\rho}$, $\mathbf{z}_k$ represents the k-th estimated value of noise in the channel received by the receiver, and $\mathbf{u}_k$ represents the k-th estimated value of the dual variable in the ADMM algorithm. Among them, $\mathbf{z}_k$ and $\mathbf{u}_k$ are only scalar multiples of the same variable $\mathbf{v}_k$. In summary, according to Eqs. (17), (18), and (19), the iterative function of ADMM can be obtained as follows:

$$
\begin{aligned}
\mathbf{x}_{k+1} &= \eta_{\mathcal{S}}(\mathbf{x}_k-\delta_k\mathbf{H}^{T}(\hat{\mathbf{e}}_k+(1-2\beta)\hat{\mathbf{v}}_k);\theta) \\
\mathbf{e}_{k+1} &= \mathbf{H}\mathbf{x}_{k+1}-\mathbf{y} \\
\hat{\mathbf{e}}_{k+1} &= \mathbf{e}_{k+1}+\psi_k(\mathbf{e}_{k+1}-\mathbf{e}_k) \\
\mathbf{v}_{k+1} &= \alpha\hat{\mathbf{e}}_{k+1}+(1-\alpha\beta)\hat{\mathbf{v}}_k \\
\hat{\mathbf{v}}_{k+1} &= \mathbf{v}_{k+1}+\gamma_k(\mathbf{v}_{k+1}-\mathbf{v}_k)
\end{aligned} \tag{20}
$$

Among $\eta_S(x;\theta) = \sum_{i=1}^{|s|-1} \tanh(\theta(x-\tau_i))$, $\tau_i = \frac{1}{2}(c_i + c_{i+1})$, $i = 1, \ldots, |S|-1$.

Due to the close correlation between the performance of the algorithm and the values of (α, β, δ, γ, θ, ψ) these six parameters introduced, it is possible to choose these six parameters as training parameters and utilize the powerful training and learning capabilities of the neural network to find the most suitable parameters. The algorithm pseudocode is as follows:

Algorithm: ADMM net

1: Input: y, H

2: Output: $\mathbf{x}_L$

3: Initialization: $\mathbf{x}_0 = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}, \mathbf{e}_0 = 0, \hat{\mathbf{e}}_0 = 0, \mathbf{v}_0 = 0, \mathbf{v}_0 = 0$

4: for k=0:L-1

5: $\mathbf{x}_{k+1} = \eta_{\mathcal{S}}(\mathbf{x}_k - \delta_k \mathbf{H}^T(\hat{\mathbf{e}}_k + (1-2\beta)\mathbf{v}_k); \theta)$

6: $\mathbf{e}_{k+1} = \mathbf{H}\mathbf{x}_{k+1} - \mathbf{y}$

7: $\hat{\mathbf{e}}_{k+1} = \mathbf{e}_{k+1} + \psi_k(\mathbf{e}_{k+1} - \mathbf{e}_k)$

8: $\mathbf{v}_{k+1} = \alpha\hat{\mathbf{e}}_{k+1} + (1-\alpha\beta)\mathbf{v}_k$

9: $\mathbf{v}_{k+1} = \mathbf{v}_{k+1} + \gamma_k(\mathbf{v}_{k+1} - \mathbf{v}_k)$

10: end

In each iteration process, each iteration step is learned using a learnable layer, and the output of each layer is the input of the next layer. After multiple iterations, that is, passing through multiple learnable layers, the output of the last layer is the final estimated value of the transmitted signal. As shown in Fig. 2:

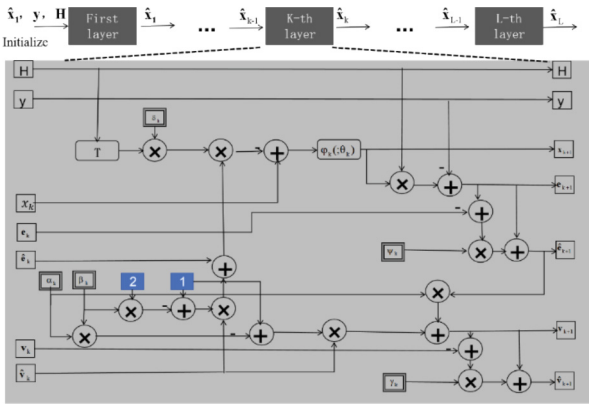

Fig. 2. ADMM network model

4 Numerical Results

In this section, we first compare the complexity with other deep learning based networks to demonstrate that the proposed ADMM RIS detection network has a lower overall complexity. Then, through simulation, it was found that it has a good convergence speed. Finally, a performance comparison was made between the ADMM RIS detection algorithm and traditional detection algorithms, as well as deep learning based detection algorithms.

4.1 Complexity

The deep learning based ADMM detection algorithm designed in this article has a computational complexity lower than OAMP Net2. And it is worth noting that, unlike DetNet, the training parameters are also directly proportional to the product of the number of transmitting and receiving antennas. The parameters that need to be trained in the proposed network are only related to the number of layers L, because each layer in the network has only 6 learnable parameters. So in a network with a total of L layers, the total number of trainable parameters is 6L, far less than DeTNeT. As shown in Table 1:

Table 1. Comparison of Complexity of Different Algorithms

Algorithm	Computational complexity	Training parameter quantity
DetNet	$O(M^2)$	$L \times (6M \times N + 2N + M)$
OAMP-Net2	$O(M^3)$	4L
ADMM RIS	$O(M^2)$	6L

According to Table 1, it can be seen that the computational complexity of ADMM Net is relatively low, and the number of parameters that need to be trained is not too large. Due to the need to balance the performance and computational cost of the model, most networks that use deep unfolding for signal detection currently have layers generally within 30. Therefore, although ADMM RIS has two more training parameters in each layer than OAMP Net2, its slightly increased training time can also be accepted.

4.2 Analysis of Simulation Results

The simulation parameter settings are shown in Table 2.

Uniformly use Signal to Noise Ratio (SNR) as the horizontal axis and consider Symbol Error Rate (SER) as the performance metric.

In order to fully compare the performance of the proposed ADMM detection algorithm, this paper compares it with various algorithms under different antenna numbers, modulation methods, and whether RIS is used as the transmitter. The comparison algorithms include zero forcing (ZF) algorithm, minimum mean square error (MMSE) algorithm, zero forcing serial interference cancellation (ZF-SIC) detection algorithm, spherical decoding (SD) detection algorithm, and DetNet, OAMP Net2 detection algorithms using deep learning.

Table 2. Simulation parameters

Modulation mode	QPSK, 16QAM, 64QAM
Number of antennas	7×7, 16×16, 32×32
Number of network layers	20
Channel conditions	rayleigh channel
Learning rate	0.001
Training dataset	10000

To explore the convergence speed of the algorithm proposed in this article, the relationship between the iteration levels of SER and the proposed ADMM detection algorithm was simulated, and the results are shown in Fig. 3:

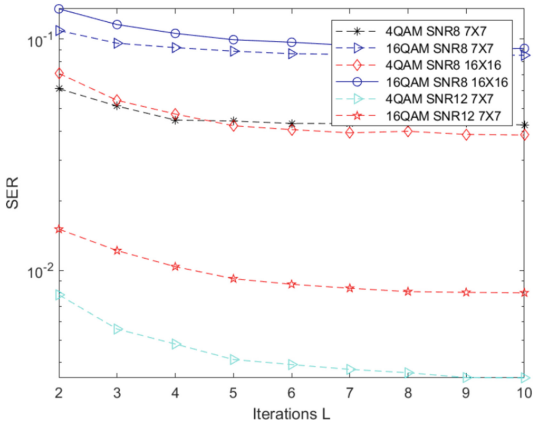

Fig. 3. Comparison of SER and ADMM network layers in different situations

It can be observed that in different situations, the proposed ADMM detection algorithm can achieve convergence around layer 5, with a faster convergence speed. Due to the trade-off between runtime and error performance, as the neural network deepens and the number of iteration layers further increases, it is expected that the probability of error will be lower.

After using RIS as the transmitter, when the number of antennas is 7 and the modulation method changes, the results are shown in Fig. 4:

It can be observed that after using RIS as the transmitter, both the traditional detection algorithm and the proposed ADMM detection algorithm have significantly reduced SER. At the same signal-to-noise ratio, the proposed ADMM detection method will decrease SER by about an order of magnitude, indicating that the intelligent metasurface as the transmitter is suitable for the proposed MIMO detection algorithm, can be compatible with it, and can improve its performance. As the modulation order changes,

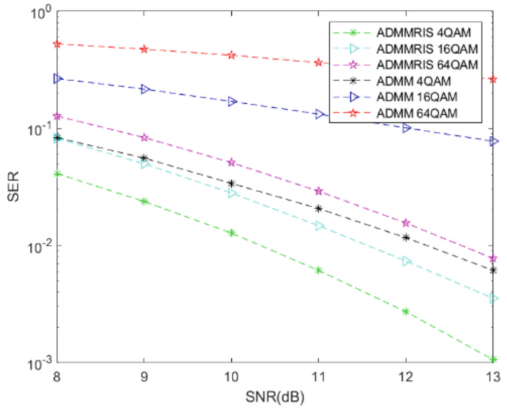

Fig. 4. Comparison of Different Modulation Methods of RIS Transmitters Used in ADMM Detection Algorithm

the performance still improves after diffraction propagation between the two RIS layers, indicating that RIS transmitters are suitable for different modulation methods.

After using RIS as the transmitter, the results of changing the number of antennas under 4QAM modulation are shown in Fig. 5:

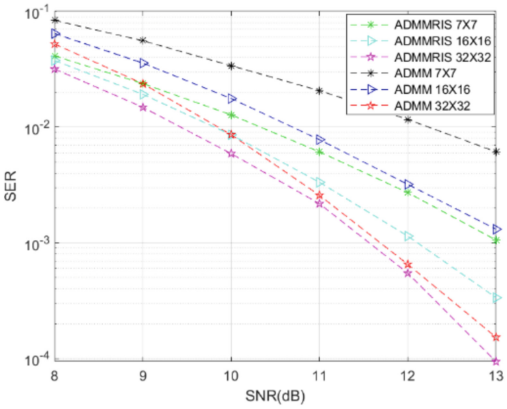

Fig. 5. Does the ADMM detection algorithm use RIS transmitters and compare different antennas

It can be observed that with the use of RIS transmitters, as the number of receiving antennas changes, the performance still improves after diffraction propagation between two RIS layers, indicating that RIS transmitters are suitable for different numbers of antennas. However, as the number of antennas increases, the power of RIS to improve performance gradually weakens. When SER is, the number of antennas is 7, and the performance improves by about 2 dB. When the number of antennas is 16, the performance improves by about 1dB, and when the number of antennas is 32, the performance improves by about 0.5 dB.

When the modulation method is 4QAM and the number of antennas is 7, the impact of using RIS as the transmitter on the deep learning based detection algorithm is shown in Fig. 6:

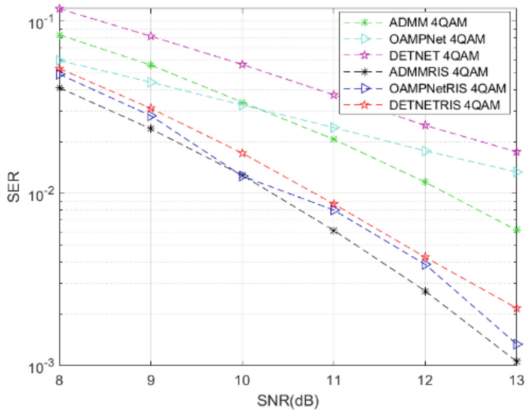

Fig. 6. Comparison of using RIS transmitters based on deep learning detection algorithms

From Fig. 6, it can be observed that when the number of antennas is 7 and the modulation method is 4QAM, the proposed ADMM detection algorithm performs similarly to OAMP Net2 without using RIS transmitters. The former performs better when the signal-to-noise ratio is below 10dB, and the latter performs better when the signal-to-noise ratio is above 10dB. But when using RIS as a transmitter, its performance is better than deep learning based DetNet and OAMP Net2 detection algorithms. When SER is, the proposed algorithm is about 0.7dB better than DetNet and about 0.3dB better than OAMP Net2.

When both the transmitting and receiving antennas are 7 and 4QAM is used for modulation, the simulation results of each control algorithm are shown in Fig. 7:

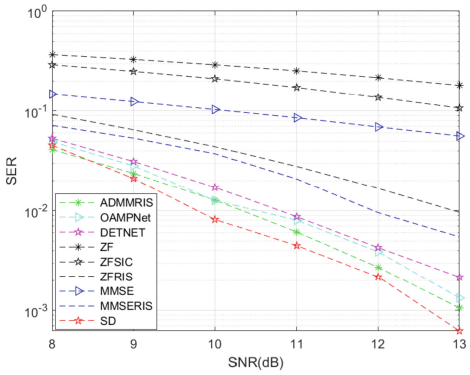

Fig. 7. Comparison of Different Detection Algorithms

As shown in the Fig. 7, with the increase of SNR, the bit error rates of several methods will decrease. Moreover, under the same SNR conditions, the ADMM detection algorithm proposed in this paper has a significantly lower bit error rate after using RIS transmitters than the traditional ZF, MMSE, and ZF-SIC algorithms without RIS transmitters, and is significantly lower than the ZF and MMSE detection algorithms using RIS transmitters. It is slightly better than the DETNET and OAMP-Net2 algorithms based on deep learning. Performance comparable to SD detection algorithm.

5 Conclusion

In this article, we propose a MIMO system model based on double-layer RIS, explore its compatibility with existing MIMO communication systems, and improve the ADMM algorithm. By introducing parameters and using deep learning for training, we optimize the algorithm, reduce its complexity to a certain extent, and enhance its performance. The algorithm will be further optimized in the future to make it more suitable for large-scale antenna scenarios.

References

1. Pan, C., Zhou, G., et al.: An overview of signalrocessing techniques for RIS/IRS-aided wireless systems. IEEE J. Sel. Top. Signal Process. **16**(5), 883–917 (2022)
2. Li, Z., Chen, W., Wu, Q., et al.: Toward transmissive RIS transceiver enabled uplink communication systems: design and optimization. IEEE Internet Things Things J. **11**(4), 6788–6801 (2024)
3. Li, Q., Wen, M., Direnzo, M., et al.: Single-RF MIMO: from spatial modulation to metasurface-based modulation. IEEE Wireless Commun. **28**(4), 88–95 (2021)
4. Lu, Y., Hao, M., Mackenzie, R.: Reconfigurable intelligent surface based hybrid precoding for THz communications. Intell. Converged Networks **3**(1), 103–118 (2022)
5. Jia, L., Tan, W., Chen, X., et al.: Realization of reconfigurable intelligent surface-based index modulation transmission. IEEE Global Commun. Conf. 633–638 (2022)
6. Zhao, J., Yang, X., Dai, J., et al.: Programmable time-domain digital-coding metasurface for non-linear harmonic manipulation and new wireless communication systems. Natl. Sci. Rev. **6**(2), 231–238 (2019)
7. Tang, W., Li, X., Dai, J., et al.: Wireless communications with programmable metasurface: transceiver design and experimental results. China Commun. **16**(5), 46–61 (2019)
8. Tang, W., Dai, J., Chen, M., et al.: Programmable metasurface-based RF chain-free 8PSK wireless transmitter. Electron. Lett. **55**(7), 417–420 (2019)
9. Dai, J., Tang, W., Wankai, L., et al.: Realization of multi-modulation schemes for wireless communication by time-domain digital coding metasurface. IEEE Trans. Antennas Propag. **68**(3), 618–1627 (2020)
10. Tang, W., Aai, J., Chen, M., et al.: MIMO transmission through reconfigurable intelligent surface: system design, analysis, and implementation. IEEE J. Sel. Areas Commun. **38**(11), 2683–2699 (2020)
11. Yang, S., Hanzo, L.: Fifty years of MIMO detection: the road to large-scale MIMOs. IEEE Commun. Surv. Tutor. **17**(4), 1941–1988 (2015)
12. Zhu, X., Murch, R.D.: Performance analysis of maximum likelihood detection in a MIMO antenna system. IEEE Trans. Commun. **50**(2), 187–191 (2002)

13. He, C., Huang, S., Chen, S., Wang, Z.: A low bitrates image semantic coding method based on semantic communication. J. Signal Process. **39**(3), 410–418 (2023)
14. He, H., Jin, S., Wen, C., et al.: Model driven deep learning for physical layer communications. IEEE Wirel. Commun. **26**(5), 77–83 (2019)
15. Samuel, N., Diskin, T., Wiesel, A.: Deep MIMO detection. In: 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications, pp. 1–5 (2017)
16. He, H., Wen, C., Jin, S., et al.: Model-driven deep learning for MIMO detection. IEEE Trans. Signal Process. **68**, 1702–1715 (2020)
17. Kim, M., Park, D.: Learnable MIMO detection networks based on inexact ADMM. IEEE Trans. Wirel. Commun. **20**(1), 565–576 (2021)
18. Chen, S., Hui, Y., Qin, Y., et al.: RIS-based diffractional channel cooding: error control coding on-the-air. submitted to IEEE Wireless Communications Magazine
19. Franca, G., Robinson, D.P., Vidal, R.: A nonsmooth dynamical systems perspective on accelerated extensions of ADMM. IEEE Trans. Automat. Cont. **68**(5), 2966–2978 (2023)