Flexible Link Adaptation in Fully-Decoupled RAN: A Machine Learning Approach

Xianfeng Zou*, Jiwei Zhao*, Kai Yu*, Jiacheng Chen†, Haibo Zhou*, and Xuemin (Sherman) Shen‡
*School of Electronic Science and Engineering, Nanjing University, Nanjing, China, 210023
†Department of Mathematics and Theories, Peng Cheng Laboratory, Shenzhen, China, 518000
Email: {xianfengzou, jw_zhao, kaiyu}@smail.nju.edu.cn, chenjch02@pcl.ac.cn, haibozhou@nju.edu.cn
‡Department of Electrical and Computer Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada, N2L 3G1. Email: sshen@uwaterloo.ca

2023 IEEE/CIC International Conference on Communications in China (ICCC) | 979-8-3503-4538-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCC57788.2023.10233482

***Abstract*—Fully decoupled radio access network (FD-RAN), as an emerging radio access architecture through physical uplink-downlink decoupling and control-data decoupling, has potential in flexible spectrum utilization and network cooperation. However, real-time uplink feedback is challenging in FD-RAN due to the complete physical decoupling of control and data base stations. In this paper, we propose a flexible link adaptation mechanism that leverages outdated channel state information (CSI) to determine the appropriate Modulation and Coding Scheme (MCS) for the user in the FD-RAN downlink. Specifically, we first utilizes kernel recursive least squares to predict the CSI at the future moment. We then select the optimal modulation and coding scheme based on the predicted CSI and the frame error rate estimated by a neural network. Simulation results show that the proposed link adaptation mechanism has a significant throughput performance gain in various scenarios.**

***Index Terms*—FD-RAN, link adaptation, feedback delay, kernel recursive least squares, neural network**

I. Introduction

To meet the growing traffic demand from mobile users and Internet-of-things, fifth-generation (5G) cellular systems are currently being deployed worldwide. However, spectrum scarcity and high network operating costs still exist. To further improve the network performance, the fully decoupled radio access network (FD-RAN) [1] has been proposed to address these challenges. FD-RAN achieves efficient utilization of network resources through decoupling and collaboration, to further reduce operating costs [2]. However, the decoupling of control and data base stations (BSs) in FD-RAN makes it difficult to obtain real-time user feedback information for downlink transmission, which has a negative effect on the application of some advanced communication technologies. One of which is link adaptation (LA). LA is a crucial technology to improve link throughput. Based on the current channel state, the receiver dynamically adjusts the transmission parameters such as the modulation and coding scheme (MCS), and feedbacks it to the corresponding transmitter via the control channel. As shown in Fig. 1, in FD-RAN, the MCS is reported over the long feedback link due to the complete physical decoupling of control and data channels, which leads to a long feedback delay. The feedback delay results in outdated MCS received at the downlink BS and degrades the LA performance due to channel variation. Furthermore, LA needs to select the optimal MCS by considering the channel state. Traditional look-up table LA methods exploit static mappings between one-dimensional channel quality and each MCS's frame error rate (FER) performance to select the optimal MCS. However, the static mapping approach cannot cope with the complicated channel and multiple antennas scenario in FD-RAN. Consequently, it is of critical importance to develop a flexible link adaptation mechanism for FD-RAN.

Many channel prediction methods have been investigated to solve the MCS feedback delay problem in LA. A prediction method based on normalized least mean square(NLMS) filtering was proposed in [3]. [4] investigated several prediction algorithms to mitigate feedback delay effects in LA, including linear filter with stochastic approximation and Kalman filter. However, traditional linear prediction methods have significant limitations in predicting nonlinearly varying channels. These methods have no performance gain in long feedback scenarios like FD-RAN. In addition, in some previous studies on LA, the MCS selection is modeled as a multi-class classification problem, where each class represents a specific MCS [5] [6]. And a trainable model is trained to select the best MCS. Specifically, the authors in [5] proposed an LA method based on autoencoder and multi-class support vector machine (SVM), where autoencoder is used to extract channel features, and multi-class SVM is used to select MCS. [6] proposed a convolutional neural network (CNN)-based multi-class classifier for LA on MIMO-OFDM systems. Unfortunately, they have not considered the feedback delay.

In this paper, we propose a new flexible LA framework for FD-RAN. Considering the long feedback delay of FD-RAN, we first predict the future SINR-based channel features based on historical values. We then use a neural network to estimate the FER of the predicted channel with different MCS transmissions. Finally, we select the MCS with the maximum expected throughput to improve the LA performance of FD-RAN. The main contributions can be summarized as follows:

- We propose a channel prediction method based on the kernel recursive least squares (KRLS) algorithm, which performs least squares regression in high-dimensional space to predict nonlinear channel feature. This method solves the problem of outdated MCS in FD-RAN because it predicts the channel at the future transmission time.
- We exploit supervised learning neural networks to estimate the FER for a given channel feature and MCS. Then,

we select the MCS with the maximum expected throughput based on the estimated FER, which can significantly improve the throughput performance.
- We verify the proposed LA framework in various simulation scenarios. Results show that our method obtains considerable performance gains with the long feedback delay of FD-RAN.

The remainder of this paper is organized as follows. Section II describes the system model and problem statement. The proposed LA method is presented in Section III. The simulation results are presented in Section IV. Finally, the paper is concluded in Section V

II. System Model and Problem Statement

A. Signal Model

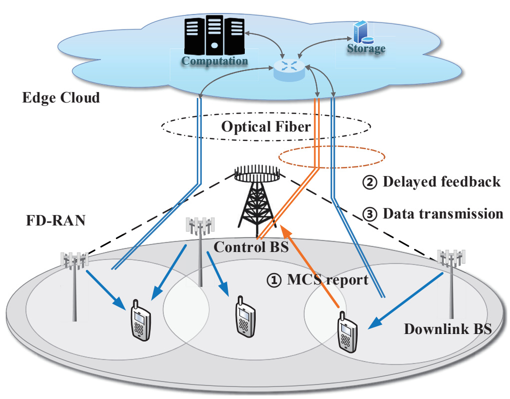

Fig. 1: Link adaptation in downlink FD-RAN

We consider a downlink single-user LA in FD-RAN architecture, which includes N_T transmit antennas at the downlink BS and N_R receive antennas at the UE. Given subcarrier $k \in 1, \ldots K$, $\mathbf{H}_k \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{T}}$ as the channel matrix, $\mathbf{W}_k \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{L}}$ as the precoding matrix, then the received symbol on the kth subcarrier $\mathbf{y}_k \in \mathbb{C}^{N_\mathrm{R} \times 1}$ is expressed as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{W}_k \mathbf{x}_k + \mathbf{n}_k, \tag{1}$$

where $\mathbf{x}_k \in \mathbb{C}^{N_L \times 1}$ is the transmit symbol, N_L is the number of spatial streams, $\mathbf{n}_k \sim \mathcal{CN}\left(0, \sigma^2 \cdot \mathbf{I}\right)$ is the complex zero-mean Gaussian noise with variance σ^2.

At the receiver, the post-equalization symbol $\mathbf{r}_k$ through linear equalization matrix G_k is given by

$$\mathbf{r}_k = \mathbf{G}_k \mathbf{y}_k. \tag{2}$$

After that, the post-processing SINR on kth subcarrier and lth stream is expressed as

$$\gamma_{k,l} = \frac{\left|\mathbf{K}_k(l,l)\right|^2}{\sum_{i \neq l} \left|\mathbf{K}_k(l,i)\right|^2 + \sigma^2 \sum_i^{N_L} \mathbf{G}_k(l,i)}, \tag{3}$$

where $\mathbf{K}_k = \mathbf{G}_k \mathbf{H}_k \mathbf{W}_k \in \mathbb{C}^{N_L \times N_L}$.

B. Link Adaptation in FD-RAN

Let $m \in \{0, 1, \cdots, M\}$ denotes the index of MCS. We consider the MCS set of 15 values defined by the LTE physical layer specifications [7]. The MCS selection problem can be expressed as

$$m^* = \arg\max_{m \in M} \left\{T_m * (1 - \mathrm{FER}(m, \gamma))\right\}, \tag{4}$$

where T_m is a fixed value representing the number of transmitted information bits when MCS m is used, $\mathrm{FER}(m, \gamma)$ denotes the FER with MCS m and channel state $\gamma = [\gamma_{k,l}]_{k=1,\ldots,\mathrm{K}; l=1,\ldots,N_\mathrm{L}}$.

Although the exact mathematical function of FER is not available, we can obtain the estimate of FER by the exponential effective SINR mapping (EESM) method [8]. EESM method maps the channel state vector γ to an equivalent SINR γ_{eff} of an AWGN channel. EESM is given by

$$\gamma_{eff} = -\beta \ln\left(\frac{1}{N_L K} \sum_{l=1}^{N_L} \sum_{k=1}^{K} e^{-\frac{\gamma_{k,l}}{\beta}}\right), \tag{5}$$

where β is the adjustment parameter. The FER can be estimated through γ_{eff} and the FER-SINR curve for the AWGN channel.

However, EESM is a loss-compression of channel quality and performs well for SISO channels, but it is difficult to obtain accurate FER estimation in a MIMO system [9]. Meanwhile, in FD-RAN, the MCS is reported over a long backhaul link, as shown in Fig. 1. Feedback delay and channel variation result in a mismatched MCS in downlink BSs, which can degrade the performance of LA. In the next section, we will discuss an accurate LA framework for FD-RAN.

III. Flexible Link Adaptation in FD-RAN

We propose a novel LA framework in this section to deal with the LA challenges in FD-RAN. We first construct a low-dimensional channel state feature based on post-processing SINRs. Subsequently, we present a channel feature prediction scheme with KRLS. Finally, an optimal MCS selection method based on multilayer perceptron (MLP) is proposed.

A. Low-dimensional Channel Feature

EESM uses a one-dimensional equivalent SINR as a channel quality metric. Still, for frequency-selective and spatial-selective channels, the one-dimensional feature does not contain enough channel information to perform LA. Multidimensional features perform better in this scenario.

It has been shown that the set of ordered post-processing SINRs over all subcarriers and spatial streams sufficiently parameterizes the approximate FER [10]. To balance the complexity and accuracy, we expect to extract a low-dimensional channel state feature. We first sort all post-processing SINRs in descending order and then select d of these sorted SINR values by equally spaced sampling. In this way, we obtain a $d-$dimensional channel feature $\mathbf{z}$, which is an approximation of the ordered SINR set. It can be defined as

$$\mathcal{S}(\gamma) \triangleq \mathbf{z} = (z_1, \ldots, z_d), \tag{6}$$

where $z_1, \ldots, z_d \in \gamma$, $\mathcal{S}$ represents the sorting and sampling operation.

B. Channel Feature Prediction with KRLS

One approach to tackle the feedback delay problem is to compensate for channel variation by channel prediction. In this work, we predict the channel feature for future moments from historical values. Then, the reported MCS is derived based on the predicted channel feature. Since the channel is time-varying, the predictor should adjust the relevant parameters online to keep track of the channel change. Traditional linear predictors work only with a small feedback delay. The neural network prediction scheme has high computational complexity and may not be suitable for LA with high real-time requirements. In this section, the KRLS algorithm [11] is employed to predict the channel. KRLS is a kernel adaptive filter widely used in nonlinear and non-stationary time series prediction. KRLS is also an online algorithm with relatively low computational complexity and can be deployed in real-time LA control.

Specifically, we formulate the channel feature $\mathbf{z}$ prediction problem as a time series prediction problem and consider the prediction of each dimension of $\mathbf{z}$ independently. At time t, the p-dimensional input of the predictor can be denoted as $x_l(t) = [z_l(t), z_l(t-1), \ldots, z_l(t-p+1)] \in \mathbb{R}^p$, the corresponding target output as $y_l(t) = z_l(t+\tau), l = 1, 2, \ldots, d$, τ is the feedback delay, and p is the window length of historical values. For simplicity, we drop the subscript l with the knowledge that prediction procedures have to be performed independently for each dimension of $\mathbf{z}$.

We first introduce the least squares (LS), a widely utilized method for linear data prediction. Assume a set of input-output pairs $\{x(i), y(i)\}|_{i=1}^{t}$, where t is the total number of samples. The input data matrix is $\mathrm{X} = [x(1), \ldots, x(t)]^T \in \mathbb{R}^{t\times p}$, the target output vector is $\mathbf{y} = (y(1), \ldots, y(t))^T \in \mathbb{R}^{t\times 1}$. The LS method aims to seek the optimal weight vector $\mathbf{w} \in \mathbb{R}^{p\times 1}$ that minimizes the cost function

$$\min_{\mathbf{w}} J = \|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2. \tag{7}$$

It is easy to obtain the optimal solution $\mathbf{w} = \mathrm{X}^{\dagger}\mathbf{y}$, where $(\cdot)^{\dagger}$ denotes the pseudo-inverse. Recursive least squares (RLS) is an online variant of LS that recursively calculates the optimal $\mathbf{w}$ when a new sample arrives [12].

The basic idea of KRLS is to execute RLS algorithms in high-dimensional feature spaces, which makes it possible to solve the problem of nonlinear time-varying channel prediction. First, the input $x(t)$ is mapped into a high-dimensional space. The nonlinear mapping procedure is $\varphi(\cdot) : \mathbf{X} \to \mathbf{F}$, where $\mathbf{F} = \{\varphi(x) \mid x \in \mathbf{X}\}$ represents the high-dimensional space. The predict value $\hat{y}(t)$ is obtained by $\varphi^T(x(t))w(t)$, where $w(t)$ represents the weight vector. Therefore, the cost function of the KRLS model is defined as

$$\min_{w(t)} J = \sum_{i=1}^{t} \left| y(i) - \varphi^T(x(i))w(t) \right|^2. \tag{8}$$

We define $\Phi(t) = [\varphi(x(1)), \ldots, \varphi(x(t))]$ as the map matrix. Then, $w(t)$ can be expressed as $\Phi(t)\alpha(t)$, where $\alpha(t) \in \mathbb{R}^t$ is the coefficient vector. The cost function can be redefined as

$$\min_{\alpha(t)} J = \|\mathbf{y}(t) - \mathrm{K}(t)\alpha(t)\|^2, \tag{9}$$

where $\mathrm{K}(t) = \Phi(t)^T\Phi(t) \in \mathbb{R}^{t\times t}$ is the kernel matrix, $\mathrm{K}(t)_{ij} = \langle \varphi(x(i)), \varphi(x(j)) \rangle = k(x(i), x(j))$ and k denotes the kernel function. The Gaussian kernel function $k(x(i), x(j)) = \exp\left(-\|x(i) - x(j)\|^2 / 2\sigma^2\right)$ is adopted in our work. The optimal $\alpha(t)$ can be express as $\mathrm{K}(t)^{\dagger}\mathbf{y}(t)$.

Through the above analysis, the dimension of $\mathrm{K}(t)$ is equal to the number of samples. As the sample size increases over time, computational complexity will also increase. We employ the approximate linear dependency(ALD) sparse rule in [11] to solve the problems. When a new sample comes, we consider the ALD condition

$$\delta(t) = \min_{a(t)} \left\| \sum_{i=1}^{m} a(t)_i \varphi(\tilde{x}(i)) - \varphi(x(t)) \right\|^2 \leq v, \tag{10}$$

where $a(t) \in \mathbb{R}^m$ is the optimal coefficient vector, v denotes the given threshold and $D_{t-1} = \{\tilde{x}(i)\}, i = 1, 2, \ldots, m$ is the dictionary composed of m independent input samples. The new sample can only be appended to the dictionary if (10) holds. Performing the minimization in (10), We can get

$$\begin{aligned} a(t) &= \tilde{\mathrm{K}}^{-1}(t-1)\tilde{k}_{t-1}(x(t)) \\ \delta(t) &= k_{tt} - \tilde{k}_{t-1}^T(x(t))a(t) \leq v, \end{aligned} \tag{11}$$

where $\tilde{\mathrm{K}}(t-1) \in \mathbb{R}^{m\times m}$ is a sparse kernel matrix, $\tilde{\mathrm{K}}(t-1)_{ij} = k(\tilde{x}(i), \tilde{x}(j))$,$\left(\tilde{k}_{t-1}(x(t))\right)_i = k(\tilde{x}(i), x(t)), i, j = 1, 2, \ldots, m$ and $k_{tt} = k(x(t), x(t))$.

We define $\tilde{\Phi}(t) = [\varphi(\tilde{x}(1)), \ldots, \varphi(\tilde{x}(m))]$, coefficient matrix $\mathrm{A}(t) \in \mathbb{R}^{t\times m}$ and the j-th row of A denotes the ALD coefficient $a(j)$. So we have $w(t) = \Phi(t)\alpha(t) \approx \tilde{\Phi}(t)\tilde{\alpha}(t)$, where $\tilde{\alpha}(t) = \mathrm{A}(t)^T\alpha(t) \in \mathbb{R}^m$. The corresponding approximation of kernel matrices is $\mathrm{K}(t) \approx \mathrm{A}(t)\tilde{\mathrm{K}}(t)\mathrm{A}(t)^T$, and the cost function is rewritten as

$$\min_{\tilde{\alpha}(t)} \|\mathbf{y}(t) - \mathrm{A}(t)\tilde{\mathrm{K}}(t)\tilde{\alpha}(t)\|^2. \tag{12}$$

By solving this LS problem, the optimal weight vector can be expressed as follows

$$\begin{aligned} \tilde{\alpha}(t) &= (\mathrm{A}(t)\tilde{\mathrm{K}}(t))^{\dagger}\mathbf{y}(t) \\ &= \tilde{\mathrm{K}}(t)^{-1}\left(\mathrm{A}(t)^T\mathrm{A}(t)\right)^{-1}\mathrm{A}(t)^T\mathbf{y}(t) \\ &= \tilde{\mathrm{K}}^{-1}(t)\mathrm{P}(t)\mathrm{A}(t)^T\mathbf{y}(t). \end{aligned} \tag{13}$$

In the next time step, the corresponding predicted value of the vector of historical values $x(t+1)$ is given by

$$\begin{aligned} \hat{y}(t+1) &\approx \varphi^T(x(t+1))\tilde{\Phi}(t)\tilde{\alpha}(t) \\ &\approx \sum_{i=1}^{m} \tilde{\alpha}(t)_i k(\tilde{x}(i), x(t+1)). \end{aligned} \tag{14}$$

The algorithm for channel prediction using ALD-KRLS is summarized in Algorithm 1, which consists of the recursive form of updating the weight vector $\tilde{\alpha}(t)$ in both cases where the ALD condition holds and not hold. We build different KRLS prediction models for each dimension of the channel feature $\mathbf{z} \in \mathbb{R}^d$. In each transmission time, we use Algorithm 1 to predict each dimension of the channel feature and combine all the predicted values sequentially to form the predicted channel feature $\hat{\mathbf{z}}$.

Algorithm 1: Prediction Based on ALD-KRLS

```
Initialization: $t_0 = p + \tau$, $\tilde{\mathrm{K}}^{-1}(t_0)$, $\tilde{\alpha}(t_0)$, $\mathrm{P}(t_0)$, $m = 1$, threshold $v$.
for transmission time $t = t_0 + 1, t_0 + 2, \cdots$ do
    Construct input vector $x(t)$ and new update sample $\{x^*(t), y^*(t)\} = \{x(t-\tau), y(t-\tau)\}$ based on the current and historical channel feature.
    Compute ALD condition (10) of $\{x^*(t), y^*(t)\}$ and update weight vector $\tilde{\alpha}(t)$:
    if $\delta > \nu$ then
        $D(t) = D(t-1) \cup x^*(t), m = m + 1,$
        $\tilde{\mathrm{K}}^{-1}(t) = \frac{1}{\delta(t)} \begin{bmatrix} \delta(t)\tilde{\mathrm{K}}^{-1}(t-1) + a(t)a(t)^T & -a(t) \\ -a(t)^T & \mathbf{1} \end{bmatrix},$
        $\mathrm{A}(t) = \begin{bmatrix} \mathrm{A}(t-1) & 0 \\ 0^T & 1 \end{bmatrix}, \quad \mathrm{P}(t) = \begin{bmatrix} \mathrm{P}(t-1) & 0 \\ 0^T & 1 \end{bmatrix}$
        $e(t) = y^*(t) - \tilde{k}_{t-1}^T(x^*(t))\tilde{\alpha}(t-1)$
        $\tilde{\alpha}(t) = \begin{bmatrix} \tilde{\alpha}(t-1) - a(t)e(t)/\delta(t) \\ e(t)/\delta(t) \end{bmatrix}$
    else
        $D(t) = D(t-1), \tilde{\mathrm{K}}(t) = \tilde{\mathrm{K}}(t-1), m = m$
        $\mathrm{A}(t) = \left[\mathrm{A}(t-1)^T, a(t)\right]^T, q(t) = \frac{\mathrm{P}(t-1)a(t)}{1 + a(t)^T \mathrm{P}(t-1)a(t)}$
        $\mathrm{P}(t) = \mathrm{P}(t-1) - q(t)a(t)^T \mathrm{P}(t-1)$
        $e(t) = y^*(t) - \tilde{k}_{t-1}^T(x^*(t))\tilde{\alpha}(t-1)$
        $\tilde{\alpha}(t) = \tilde{\alpha}(t-1) + \tilde{\mathrm{K}}^{-1}(t)q(t)e(t)$
    end if
    Calculate the predicted value:
        $\hat{y}(t) \approx \sum_{i=1}^{m} \tilde{\alpha}(t)_i k(\tilde{x}(i), x(t))$
end for
```

C. Optimal MCS Selection

In this section, we select the optimal MCS based on the predicted channel feature. As (4) shows, the selection of MCS requires an accurate FER estimate. In previous work [6], MCS selection is generally modeled as a multi-class classification problem. The classifier outputs the FER of multiple MCSs based on a given channel feature. However, a multiple-MCS label for the instantaneous channel feature is unavailable in a real-time communication system. In this section, we propose a supervised learning-based online optimal MCS selection method, where the training data can be obtained in real-time.

Since neural networks can approximate highly nonlinear functions, we apply a simple MLP network to fit the FER function. To reduce the complexity of the network, we replace all post-processing SINRs γ with the low-dimensional channel features $\mathbf{z}$. Thus, the FER function estimated by the MLP network can be expressed as $\rho(\mathbf{z}, m; \boldsymbol{\theta})$, where MCS m and channel feature $\mathbf{z}$ are the input to the network, θ denotes the trainable network parameters. We construct the training set $\{((\mathbf{z}_n, m_n), e_n)\}_{n=1}^{N}$ by observing N transmission frames, where $(\mathbf{z}_n, m_n)$ is the low-dimensional channel feature and the MCS used at the n-th transmission frame, and $e_n \in \{0, 1\}$ denotes the successful/unsuccessful transmission of the n-th frame. After that, we minimize the MLP network cross-entropy loss function over the training set and obtain the optimal MLP network parameters θ. The cross-entropy loss function is expressed as follows

$$J(\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^{N} \left(-e_n \log \rho - (1 - e_n) \log (1 - \rho)\right) \qquad (15)$$

Based on the trained network model, (4) can be rewritten as

$$m^* = \arg \min_{m \in \hat{M}} \left\{ T_m * \rho(m, \hat{\mathbf{z}}; \boldsymbol{\theta}) \right\}. \qquad (16)$$

To select the optimal MCS, we first extract the predicted channel feature $\hat{\mathbf{z}}$. Meanwhile, considering the channel time correlation, the optimal MCS values of two adjacent transmission frames are not too far apart. Based on this observation, the candidate MCS values for the current frame is limited to the neighborhood of the previous MCS, defined as

$$\hat{M}(t) = \{max(0, m^*(t-1) - \Delta m), \cdots, min(M, m^*(t-1) + \Delta m)\}. \qquad (17)$$

By traversing all the candidate MCS values, the trained model outputs the FER of different MCS. Finally, we select the MCS that provides the maximum expected throughput. A summary of the procedure of the proposed method is presented in Algorithm 2.

Algorithm 2: Optimal MCS Selection Based on MLP Network

```
for each frame $n$ of the first $N$ do
    Decode signal and save the training data:
        the successful/unsuccessful label: $e_n \in \{0, 1\}$,
        the feature:channel feature $\mathbf{z}_n$, used MCS $m_n$
    Select MCS through EESM and feedback
end for
Train the MLP network model with the saved train set
for each frame do
    if delayed feedback then
        Predict the channel feature $\hat{\mathbf{z}}$ by Algorithm 1
    end if
    Estimate FER with trained model and select MCS
    by Eq.(16)
end for
```

IV. Simulation Results

In this section, we evaluate the performance of the proposed LA framework through link-level simulations. The main parameters are summarized in Table I. The traditional prediction algorithms and EESM MCS selection method are taken as a benchmark to evaluate the performance of KRLS channel prediction and MLP MCS selection framework.

TABLE I: SIMULATION PARAMETERS

Parameters	Value	Parameters	Value
Carrier frequency	2GHz	Subcarrier spacing	15KHz
Bandwidth	10 MHz	MIMO configuration	2×2
Spatial stream	1	Channel knowledge	perfect
Equalizer	zero forcing	Channel model	3GPP EPA5
Feedback period	1ms	Feature dimension d	5

A. Performance of Ideal Feedback

We first set the feedback delay to 0 and evaluate the performance of the proposed MLP MCS selection method without channel prediction. The MLP network has 2 hidden layers. The activation functions in the hidden and output layers are ReLU and sigmoid, respectively. This lightweight model can be quickly retrained to address the changing communication environment. The simulated SINR values range from 0dB to 30dB with a 1 dB step. For every SINR value, we simulate 2×10^4 consecutive transmission frames. According to Algorithm 2, the first 10^4 transmission frames create the training set for MLP model training. In the subsequent 10^4 transmission frames, the trained model is applied to select the MCS in LA. We provide the throughput performance for the proposed MLP method and take the conventional EESM method as a benchmark. Fig. 2 shows the average throughput of different methods. The ideal throughput is obtained by simulating each MCS for each frame and selecting the largest MCS that is successful. We observe that the proposed method can achieve higher throughput than the EESM method, especially in the medium SINR region. In addition, we give the performance of the fixed MCS method without feedback. It can be seen that the optimal fixed MCS varies with the SINR value, but the performance is always worse than the LA methods. Therefore, LA is one of the crucial technologies to improve the FD-RAN link throughput.

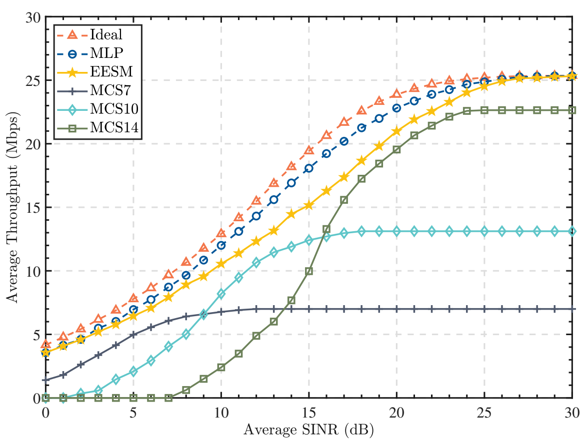

Fig. 2: Throughput versus average SINR

B. Performance of Delayed Feedback

In the case of delayed feedback, we compare the proposed KRLS-based prediction scheme with the Kalman filter [4] and NLMS prediction schemes [3]. For KRLS, the ALD threshold is set to 0.00001, and a Gaussian kernel with a bandwidth σ of 32 is used. The step size parameter of NLMS is set to 1. The optimal length of historical window p varies for different feedback delays. For simplicity, we fix $p = 10$ for all prediction schemes. Performance is evaluated on 10^4 transmission frames where the average SINR is set to 10 dB.

Figure. 3 shows the link throughput performance of different prediction schemes versus the length of feedback delay. From this figure, it can be seen that without prediction, the system performance degrades rapidly as the feedback delay increases. However, even in the case of delayed feedback, the proposed MLP MCS selection method outperforms the traditional EESM. Meanwhile, all prediction schemes can provide prediction gain when the feedback delay $\tau < 10ms$. However, when $\tau > 10ms$, the correlation between the input and output of the predictor is insufficient. Kalman filter and NLMS predictor cannot provide prediction gain, and the performance is even worse than no prediction. The KRLS scheme can always provide the highest prediction gain and improves the average throughput by 5% - 25% compared to delayed feedback. And the greater the feedback delay, the greater the performance gain. Similar results can be observed in the FER performance comparison in Figure. 4. A simple conclusion can be deduced, the proposed LA framework provides significant performance gains in the long feedback scenario of FD-RAN.

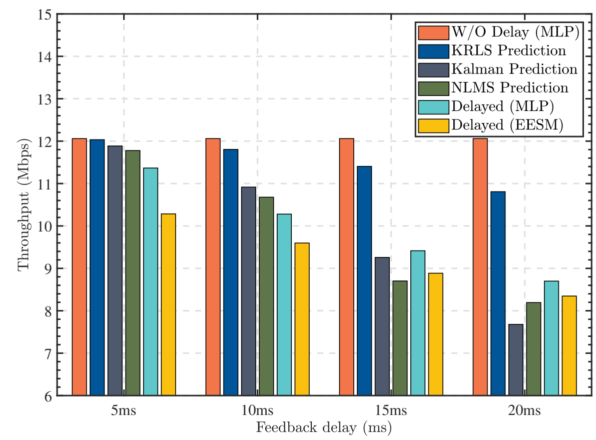

Fig. 3: Throughput of different prediction schemes versus feedback delay

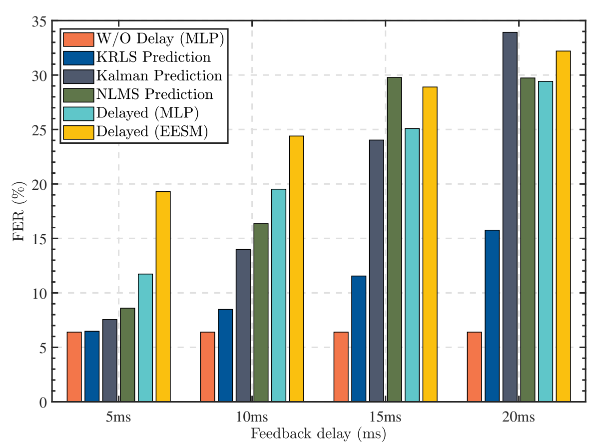

Fig. 4: FER of different prediction schemes versus feedback delay

C. Performance of Different Multi-path Fading Channel Model

To investigate the effect of user mobility, we simulate the performance of different schemes in three multi-path fading channels with different combinations of multi-path

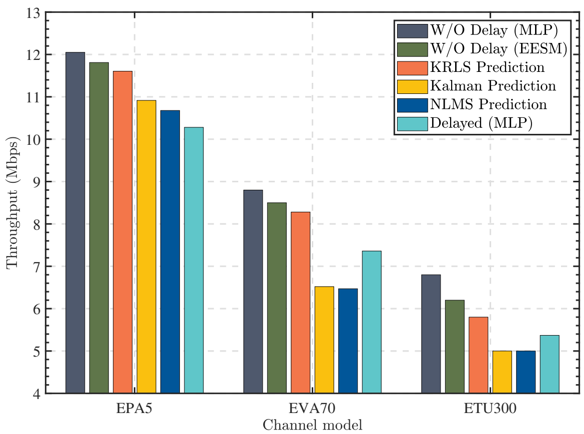

Fig. 5: Throughput of different schemes in different multi-path fading channel models

delay profile and maximum Doppler frequency. Specifically, the three delay profiles used in LTE [13], Extended Pedestrian A (EPA), Extended Vehicular A (EVA) and Extended typical urban (ETU) model are adopted in our simulations. These three delay profiles represent low, medium, and high delay spread, respectively, where higher delay spread means that the channel becomes more frequency selective. The maximum Doppler frequency of the three models are fixed at 5Hz, 70Hz and 300Hz respectively, which correspond to 3km/h, 38km/h and 162 km/h moving speed at 2 GHz carrier frequency. These three channel models are selected to be representative of the low, medium, and high user mobility environments. For the simulation of the three channels, we first evaluate the performance of different MCS selection methods (MLP & EESM) with ideal feedback. The average SINR is set to 10 dB. After that, the performance of different prediction schemes with MLP MCS selection is compared at a feedback delay of 10ms.

The performance is evaluated by the network throughput which is shown in Figure. 5. For the channels, EPA5 achieves the best performance. However, ETU300 has the most degraded performance due to the increase in Doppler frequency and delay spread which causes symbol interference. It can be seen that the performance of the proposed MLP MCS selection methods outperform the conventional EESM methods on all three channel models at ideal feedback.

In the simulation with 10ms feedback delay, Kalman and NLMS can only provide prediction gain in the low mobility channel EPA5. In higher mobility environments, the channels change rapidly, Kalman and NLMS cannot make effective channel predictions so that the performance is even worse than no prediction in EVA70 and ETU300. However, KRLS achieves the optimum prediction performance on all three channel models. Especially in a high mobility environment, KRLS can still capture the weak autocorrelation of the time-varying channel and provide prediction gain. Therefore, the proposed LA framework has stronger robustness than traditional methods in the complicated communication scenarios of FD-RAN.

V. Conclusions

In this paper, we have proposed a flexible LA framework for FD-RAN architecture. We have employed the KRLS algorithm for channel feature prediction to mitigate the effect of LA delay. The FER estimated with neural networks has been exploited for MCS selection. Simulation results have shown that the proposed framework can select the exact MCS in the long feedback delay scenario of FD-RAN, providing a significant throughput gain. For future work, we will consider the combination of LA and coordinated multiple base stations access mechanism to improve the FD-RAN performance.

Acknowledgement

This work was supported in part by the National Natural Science Foundation Original Exploration Project of China under Grant 62250004, the National Natural Science Foundation of China under Grant 62271244, the Natural Science Fund for Distinguished Young Scholars of Jiangsu Province under Grant BK20220067.

References

[1] Q. Yu, H. Zhou, J. Chen, Y. Li, J. Jing, J. J. Zhao, B. Qian, and J. Wang, "A fully-decoupled RAN architecture for 6G inspired by neurotransmission," *Journal of Communications and Information Networks*, vol. 4, no. 4, pp. 15–23, 2019.

[2] J. Zhao, Q. Yu, B. Qian, K. Yu, Y. Xu, H. Zhou, and X. Shen, "Fully-decoupled radio access networks: A resilient uplink base stations cooperative reception framework," *IEEE Transactions on Wireless Communications*, 2023.

[3] H. Touheed, A. U. Quddus, and R. Tafazolli, "Predictive cqi reporting for hsdpa," in *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, 2008, pp. 1–5.

[4] R. Abi Akl, S. Valentin, G. Wunder, and S. Stanczak, "Compensating for CQI aging by channel prediction: The LTE downlink," in *2012 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2012, pp. 4821–4827.

[5] Z. Dong, J. Shi, W. Wang, and X. Gao, "Machine learning based link adaptation method for MIMO system," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2018, pp. 1226–1231.

[6] M. Elwekeil, S. Jiang, T. Wang, and S. Zhang, "Deep convolutional neural networks for link adaptations in MIMO-OFDM wireless systems," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 665–668, 2018.

[7] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures,," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 36.213, 04 2017, version 14.2.0.

[8] J. Francis and N. B. Mehta, "EESM-based link adaptation in point-to-point and multi-cell OFDM systems: Modeling and analysis," *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 407–417, 2013.

[9] Z. Hanzaz and H. D. Schotten, "Analysis of effective SINR mapping models for MIMO OFDM in LTE system," in *2013 9th international wireless communications and mobile computing conference (IWCMC)*. IEEE, 2013, pp. 1509–1515.

[10] R. C. Daniels, C. M. Caramanis, and R. W. Heath, "Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and snr ordering," *IEEE Transactions on vehicular Technology*, vol. 59, no. 1, pp. 114–126, 2009.

[11] Y. Engel, S. Mannor, and R. Meir, "The kernel recursive least-squares algorithm," *IEEE Transactions on signal processing*, vol. 52, no. 8, pp. 2275–2285, 2004.

[12] A. H. Sayed, *Fundamentals of adaptive filtering*. John Wiley & Sons, 2003.

[13] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception,," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 36.104, 04 2017, version 14.3.0.