

Recurrent Neural Network-based Frequency-Domain Channel Prediction for Wideband Communications

Wei Jiang^{*†} and Hans D. Schotten^{†*}

^{*}German Research Center for Artificial Intelligence (DFKI)
Trippstadter Street 122, Kaiserslautern, 67663 Germany
Emails: {wei.jiang, hans.schotten}@dfki.de

[†]Institute for Wireless Communication and Navigation, University of Kaiserslautern
Building 11, Paul-Ehrlich Street, Kaiserslautern, 67663 Germany
Emails: {wei.jiang@dfki.uni-kl.de, schotten@eit.uni-kl.de}

Abstract—Outdated channel state information (CSI) severely degrades the performance of adaptive transmission systems that adapt their transmissions to channel fading. In contrast with mitigation methods that sacrifice scarce wireless resources to compensate for such a performance loss, channel prediction provides an efficient solution. A few predictors for frequency-flat channels were by far proposed, whereas those suited to frequency-selective channels are seldom explored. In this paper, therefore, we propose to apply a recurrent neural network to build a frequency-domain channel predictor for wideband communications. As an application example, integrating a predictor into a multi-input multi-output orthogonal frequency-division multiplexing system to improve the correctness of antenna selection is provided. Performance assessment is carried out in multi-path fading channels defined by 3GPP Extended Vehicular A and Extended Typical Urban models. Results reveals that this predictor is effective to combat the outdated CSI with reasonable computational complexity. It outperforms the Kalman filter-based predictor notably and has an intrinsic flexibility to enable multi-step prediction.

I. INTRODUCTION

Outdated channel state information (CSI) due to the feedback delay between a receiver and a transmitter severely deteriorates the performance of a wide variety of wireless systems, e.g., multiple-input multiple-output (MIMO) [1], massive MIMO [2], cooperative relaying [3], interference alignment [4], orthogonal frequency-division multiplexing (OFDM) [5], and physical layer security [6], [7]. In the fifth generation (5G) system, on the one hand, new applications such as Tactile Internet [8] and autonomous driving impose a huge demand for ultra-reliable, secure and high-available wireless links. On the other hand, acquiring accurate CSI gets harder in some 5G deployment scenarios, e.g., millimeter wave and high-speed trains.

To cope with the outdated CSI, a large number of mitigation algorithms and protocols that compensate for the performance loss passively with a cost of scarce wireless resources have been proposed in the literature [9]. In contrast, channel prediction provides an efficient approach by improving the quality of CSI directly without spending extra wireless resources, and therefore attracts much attention from researchers [10]. Besides

the statistical methods [11]–[14] that model a fading channel as an autoregressive (AR) process and apply a Kalman filter (KF) to realize a linear predictor, Artificial Intelligence (AI) technology is also being discussed recently. Making use of its capability of time-series prediction [15], a recurrent neural network (RNN) was firstly proposed in [16] to build a narrow-band single-antenna predictor and was further extended to MIMO channels by [17], [18]. The feasibility of applying a deep neural network to predict fading channels was also studied in [19]. In [20], the authors of this paper proposed to employ a real-valued RNN to implement a multi-step predictor and further verified its effectness in a MIMO system [21].

However, the aforementioned predictors focused mainly on frequency-flat channels, whereas those suited to frequency-selective channels are seldom explored. In this paper, therefore, we propose to utilize a RNN to build a frequency-domain channel predictor for wideband communications. As a concrete application, integrating this predictor into a MIMO-OFDM system so as to improve the correctness of selecting transmit antennas is illustrated. Performance assessment is carried out in multi-path fading environment specified by 3GPP Extended Vehicular A (EVA) and Extended Typical Urban (ETU) channel models [22]. A number of influential factors, namely, estimation errors due to additive noise, inter-antenna correlation, the Doppler shift, and signal interpolation errors, are taken into account during the assessment. Moreover, a comparison of performance and complexity between the RNN and KF predictor is conducted.

The rest of this paper is organized as follows: Section II gives the system model. Section III proposes the RNN-based frequency-domain predictor. Section IV describes a prediction-assisted MIMO-OFDM system. Section V evaluates its performance and complexity. Finally, Section VI remarks this paper.

II. SYSTEM MODEL

To begin with, the discrete-time baseband model for a single-antenna system in a frequency-selective channel is provided:

$$y[t] = \sum_{l=0}^{L-1} h_l[t]x[t-l] + z[t], \quad (1)$$

^{*}This work was supported by German Federal Ministry of Education and Research (BMBF) under the TACNET4.0 project with grant no. *KIS15GT1007*.

where $h_l[t]$ denotes the l^{th} tap for a time-varying channel filter, $x[t]$ and $y[t]$ represent the transmitted and received signals at time t , respectively, and $z[t]$ is additive noise. Dropped time index for brevity, a frequency-selective channel is modeled as a linear channel filter $\mathbf{h} = [h_0, h_1, \dots, h_{L-1}]^T$, where L is the filter length. This channel can be converted into N independent frequency-flat sub-carriers by means of the OFDM modulation [23]. The signal transmission over the n^{th} sub-carrier at time t can be modeled as

$$\tilde{y}_n[t] = \tilde{h}_n[t]\tilde{x}_n[t] + \tilde{z}_n[t], \quad n = 0, 1, \dots, N-1, \quad (2)$$

where $\tilde{x}_n[t]$, $\tilde{y}_n[t]$, and $\tilde{z}_n[t]$ stand for the transmitted signal, received signal, and noise, respectively, in the frequency domain. According to the picket fence effect in discrete Fourier transform (DFT) [24], the frequency response of the channel filter denoted by $\tilde{\mathbf{h}} = [\tilde{h}_0, \tilde{h}_1, \dots, \tilde{h}_{N-1}]^T$ is the DFT of $\mathbf{h}' = [h_0, h_1, \dots, h_{L-1}, 0, \dots, 0]^T$ that is the filter \mathbf{h} padding with $N-L$ zeros at the tail.

The extension of (2) to a multi-antenna system with N_t transmit and N_r receive antennas is straightforward by applying the same OFDM modulation into MIMO channels. Thus, on the n^{th} sub-carrier, the signal transmission is represented by

$$\tilde{\mathbf{y}}_n[t] = \tilde{\mathbf{H}}_n[t]\tilde{\mathbf{x}}_n[t] + \tilde{\mathbf{z}}_n[t], \quad n = 0, 1, \dots, N-1, \quad (3)$$

where $\tilde{\mathbf{y}}_n[t]$ represents N_r received symbols for sub-carrier n at time t , $\tilde{\mathbf{x}}_n[t]$ corresponds to N_t transmit symbols, $\tilde{\mathbf{z}}[t]$ is a vector of additive noise. $\tilde{\mathbf{H}}_n[t] = [\tilde{h}_{n_r n_t}^{n_r n_t}[t]]_{N_r \times N_t}$ denotes the frequency-domain channel matrix at time t , where $1 \leq n_r \leq N_r$, $1 \leq n_t \leq N_t$, and $\tilde{h}_{n_r n_t}^{n_r n_t} \in \mathbb{C}^{1 \times 1}$ stands for the channel frequency response on sub-carrier n between transmit antenna n_t and receive antenna n_r , which can be derived by the DFT of its channel filter denoted by $\mathbf{h}^{n_r n_t} = [h_0^{n_r n_t}, h_1^{n_r n_t}, \dots, h_{L-1}^{n_r n_t}]^T$.

Owing to the feedback delay τ , the available CSI at the transmitter is in some extent outdated, which may substantially differ from the actual CSI at the instant of signal transmission, i.e., $\tilde{\mathbf{H}}_n[t] \neq \tilde{\mathbf{H}}_n[t+\tau]$. It was widely recognized that the outdated CSI severely degrades the performance of wireless systems adapting their transmissions to the channel fading [1]–[6]. The task of channel prediction is to obtain a predicted channel matrix $\hat{\mathbf{H}}_n[t+\tau]$ that approximates its actual value $\tilde{\mathbf{H}}_n[t+\tau]$ as close as possible on the basis of the available CSI $\tilde{\mathbf{H}}_n[t]$.

III. FREQUENCY-DOMAIN CHANNEL PREDICTION

According to [25], a fading channel can be modeled as an autoregressive process of order p , and a Kalman filter is employed to build a time-domain channel predictor [11]–[14]. In this paper, we further transform it into a frequency-domain predictor that is capable of predicting channel's frequency response over each sub-carrier, as follows:

$$\hat{\mathbf{H}}_n[t+1] = \sum_{k=1}^p a_k \tilde{\mathbf{H}}_n[t-k+1], \quad (4)$$

where a_1, a_2, \dots, a_p denote the AR coefficients that can be figured out from its discrete-time autocorrelation function $R[m] = J_0(2\pi f_d T_s N |m|)$, where f_d indicates the maximal

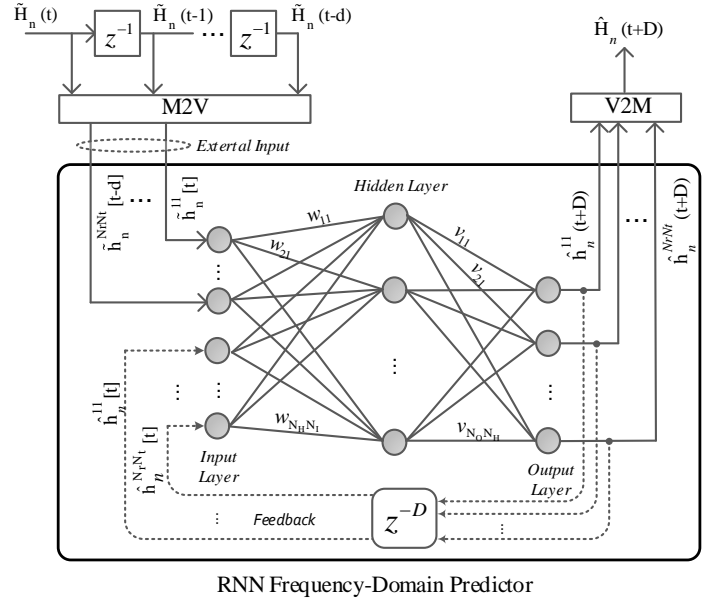


Fig. 1. Schematics of the RNN-based frequency-domain multi-step predictor for frequency-selective multi-antenna channels.

Doppler shift, the symbol period T_s equals the inverse of the sampling rate $T_s = 1/f_s$, N represents the total number of sub-carriers or the size of DFT, and $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. Despite it is simple to implement, the KF predictor has two limitations: 1) only providing one-step prediction $\hat{\mathbf{H}}[t+1]$ rather than multi-step ones $\hat{\mathbf{H}}[t+D]$; 2) relying on the knowledge of the Doppler shift f_d that is difficult to obtain in practice.

Recently, recurrent neural networks [15] are also being considered to implement channel prediction thanks to its strong capability exhibited in time-series prediction. However, the focus of the RNN predictors is by far on frequency-flat fading channels, whereas those for frequency-selective channels are still missing. To fill this gap, we propose in this paper a RNN-based frequency-domain predictor for frequency-selective multi-antenna channels. The main idea is to convert a frequency-selective channel into a set of parallel frequency-flat sub-carriers, and then utilize a RNN predictor to forecast the frequency response of sub-carriers. Fig.1 illustrates the block diagram of the proposed predictor. The applied network is composed by three layers: an input layer having N_I neurons that include external input and feedback, a hidden layer with N_H neurons, and an output layer with N_O neurons. Each connection between the output of a neuron in the predecessor layer and the input of a neuron in the successor layer is assigned with a weight. As shown in Fig.1, w_{ij} denotes the weight connecting the i^{th} input and the j^{th} hidden neuron, while v_{ml} is the weight for hidden neuron l and output m , where $1 \leq i \leq N_I$, $1 \leq l \leq N_H$, and $1 \leq m \leq N_O$.

The operation of a RNN predictor is divided into two phases: training and predicting. As long as a network's parameters such as the number of layers and neurons have been determined, a

training procedure starts from an initial state where all weights are randomly selected. Providing a training data set, i.e., a series of channel response samples, the RNN processes each sample and compares its resulting prediction against the desired value. The prediction error is propagated back through the network in order to update the weights iteratively until a certain convergence condition reaches. The details of RNN and its training procedure can refer to the literature such as [16], [26].

If a RNN is trained by a series of channel frequency response over a certain sub-carrier $\{\tilde{\mathbf{H}}_n[t] | t=1, 2, \dots\}$, it can be applied to predict unknown frequency-domain CSI per sub-carrier. At time t over sub-carrier n , as shown in Fig.1, the channel matrix $\tilde{\mathbf{H}}_n[t]$, as well as its d -step delays $\tilde{\mathbf{H}}_n[t-1], \dots, \tilde{\mathbf{H}}_n[t-d]$, are fed into the RNN as the external input. To adapt the input layer, channel matrices with the dimension of $N_r \times N_t$ need to be vectorized as:

$$\tilde{\mathbf{h}}_n = \text{vec}(\tilde{\mathbf{H}}_n) = [\tilde{h}_n^{11}, \tilde{h}_n^{12}, \dots, \tilde{h}_n^{N_r N_t}]^T, \quad (5)$$

which is implemented through a Matrix-to-Vector (M2V) module as shown in this figure. Meanwhile, feeding the D -step delay of the output, namely $\hat{\mathbf{h}}_n[t] = [\hat{h}_n^{11}[t], \dots, \hat{h}_n^{N_r N_t}[t]]^T$, back to the input layer, together with the external input, the whole input at time t is thus $\tilde{\mathbf{h}}_n[t], \tilde{\mathbf{h}}_n[t-1], \dots, \tilde{\mathbf{h}}_n[t-d], \hat{\mathbf{h}}_n[t]$. The RNN output is then the prediction for D steps ahead, i.e., $\hat{\mathbf{h}}_n[t+D]$, which can be recovered to a predicted channel matrix $\hat{\mathbf{H}}_n[t+D]$ by a Vector-to-Matrix (V2M) module.

From the perspective of a pilot-assisted system, only a subset of sub-carriers instead of all N sub-carriers needs to be predicted if channel interpolation is utilized. Suppose one pilot is inserted every N_P sub-carriers, there are a total of $P = \lfloor \frac{N}{N_P} \rfloor$ pilot sub-carriers. Given their predicted CSI $\hat{\mathbf{H}}_p[t+D]$, where for example $p = (i-1)N_P$ and $i=1, \dots, P$, the prediction for all sub-carriers $\hat{\mathbf{H}}_n[t+D]$, $n=0, \dots, N-1$ can be obtained by interpolating $\hat{\mathbf{H}}_0[t+D], \hat{\mathbf{H}}_{N_P}[t+D], \dots, \hat{\mathbf{H}}_{(P-1)N_P}[t+D]$.

IV. PREDICTION-ASSISTED TAS IN MIMO-OFDM

To further shed light on the frequency-domain channel prediction, transmit antenna selection (TAS) in a MIMO system with N_t transmit and N_r receive antennas in a frequency-selective fading channel is depicted here as one of its applications. The frequency-selective channel is converted into N parallel frequency-flat sub-carriers by means of the fast Fourier transform (FFT) modulator and the inverse FFT (IFFT) demodulator, together with the cyclic prefix (CP), as shown in Fig.2. There exist two selection schemes: bulk or per-tone selection, as mentioned in [27]. For a better observation on the effect of channel prediction, we adopt the latter, i.e., each sub-carrier decides its best antenna individually instead of the same selection for all sub-carriers.

An OFDM symbol carries a payload of M data symbols denoted by $\mathbf{d} = [d_1, d_2, \dots, d_M]^T$, while the remaining $P = N - M$ sub-carriers are reserved for comb-type pilot symbols that are uniformly inserted in sub-carriers $p = (i-1)N_P$, where $i=1, \dots, P$ and N_P is the interval of pilots. Thus, the frequency

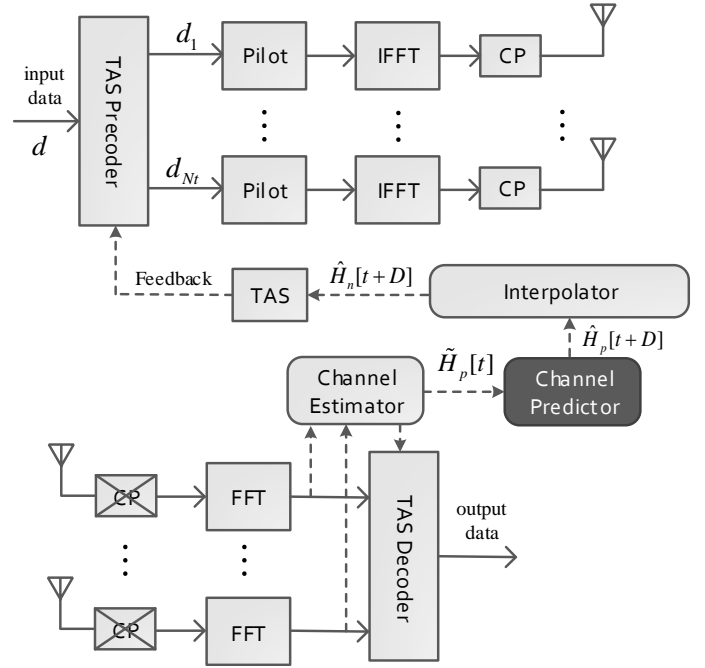


Fig. 2. Illustration of prediction-assisted TAS in a MIMO-OFDM system, where the CSI on pilot sub-carriers $\tilde{\mathbf{H}}_p[t]$ is fed into a channel predictor to get their D -step-ahead prediction $\tilde{\mathbf{H}}_p[t+D]$. Then, a channel interpolator recovers the CSI on all sub-carriers $\tilde{\mathbf{H}}_n[t+D]$, which is used for antenna selection, unlike the outdated CSI $\tilde{\mathbf{H}}_n[t]$ in a traditional TAS system.

response of pilot sub-carrier p denoted by $\tilde{\mathbf{H}}_p[t]$ can be estimated at the receiver. Taking advantage of channel's frequency correlation, a frequency-domain interpolation is conducted to recover the CSI on all sub-carriers including data and pilot ones, i.e., $\tilde{\mathbf{H}}_n[t]$, $n=0, 1, \dots, N-1$. Following the per-tone selection scheme [27], each data sub-carrier chooses its own transmit antenna that has the largest channel gain. The key difference is that the traditional TAS system directly applies the outdated CSI $\tilde{\mathbf{H}}_n[t]$ to select in sub-carrier n , following

$$\eta_n[t] = \arg \max_{1 \leq n_t \leq N_t} \left\| \tilde{\mathbf{h}}_n^{n_t}[t] \right\|, \quad (6)$$

where $\eta_n[t]$ represents the index of the best antenna at time t for sub-carrier n , $\tilde{\mathbf{h}}_n^{n_t}[t]$ is the n_t^{th} column vector of the channel matrix $\tilde{\mathbf{H}}_n[t]$, and $\|\cdot\|$ stands for the Euclidean norm of a vector. The receiver feeds a set of selected antenna indices for all data sub-carriers $\{\eta_n[t] | 0 \leq n \leq N-1, n \neq p\}$ back to the transmitter through a feedback channel. In the n^{th} sub-carrier of OFDM symbol $t+D$, the TAS precoder allocates a data symbol to the best antenna $\eta_n[t]$, while other antennas are null by padding a symbol of 0. Using a simplest system configuration with $N_t=2$, $N=8$, $N_P=4$, and $M=6$ as an example, the data vectors for two transmit chains are probably

$$\begin{aligned} \mathbf{d}_1 &= [0, d_1, d_2, 0, 0, 0, 0, d_6]^T \\ \mathbf{d}_2 &= [0, 0, 0, d_3, 0, d_4, d_5, 0]^T, \end{aligned} \quad (7)$$

where sub-carriers $n=0$ and 4 are reserved for pilot symbols, and $n=1, 2$, and 7 select the first antenna to carry data symbols d_1, d_2 , and d_6 , respectively.

Due to the channel fading, the outdated CSI $\tilde{\mathbf{H}}_n[t]$ may differ substantially from the actual CSI $\tilde{\mathbf{H}}_n[t+D]$, leading to a remarkable performance degradation [28]. With the aid of channel prediction, a selection decision can be made in terms of the predicted CSI that is possible to closely approximate the actual CSI. At time t , as depicted in Fig.2, estimating the pilots at the t^{th} OFDM symbol can get the CSI of pilot sub-carriers $\tilde{\mathbf{H}}_p[t]$, which is fed into the channel predictor to get the predicted CSI $\hat{\mathbf{H}}_p[t+D]$. A frequency-domain channel interpolator is applied to obtain the CSI on all sub-carriers denoted by $\hat{\mathbf{H}}_n[t+D]$, $n=0, 1, \dots, N-1$ so as to replace the outdated CSI $\tilde{\mathbf{H}}_n[t]$, $n=0, 1, \dots, N-1$ in a traditional TAS system. Thus, the best transmit antenna over sub-carrier n can be selected as

$$\hat{\eta}_n[t] = \arg \max_{1 \leq n_t \leq N_t} \left\| \hat{\mathbf{h}}_n^{n_t}[t+D] \right\|. \quad (8)$$

V. PERFORMANCE AND COMPLEXITY ASSESSMENTS

In this section, the performance and complexity of the proposed predictor are assessed and compared with the KF predictor. Outage probability achieved by the prediction-assisted TAS in a MIMO-OFDM system is evaluated by Monte-Carlo simulations. The applied frequency-selective channels follow the 3GPP EVA and ETU models with a maximal Doppler shift of $f_d=70\text{Hz}$ and 300Hz , respectively. The effects of additive noise, given signal-to-noise ratios (SNRs) equaling to 20dB and 30dB, and inter-antenna correlation with coefficients of $\alpha=0.3$ and 0.9, are also taken into account. The signal bandwidth (or the sampling rate) is 1MHz, which is converted into $N=64$ parallel sub-carriers by the OFDM modulation, resulting in a sub-carrier spacing around $\Delta f=15\text{KHz}$. Through the observation in the simulation, the optimal number of hidden neurons is set to $N_H=10$ and the length of tapped delay line is $d=3$. The simulation parameters are summarized in Table I.

A. Computational Complexity

In general, the number of complex multiplications is used as a measure for the computational complexity. As can be derived from Fig.1, the hidden and output layer need to conduct $N_I N_H$ and $N_O N_H$ times multiplication per prediction, respectively, amounting to a total number of $\Omega_{rnn}=N_H(N_I+N_O)$. The

number of required input neurons is proportional to the number of MIMO subchannels $N_r N_t$, we have $N_I=(d+2)N_r N_t$, and similarly the number of output neurons is $N_O=N_r N_t$. Then, the complexity of the RNN predictor can be indicated by $\Omega_{rnn}=(d+3)N_H N_r N_t$. In contrast, derived from (4), the KF predictor requires $\Omega_{kf}=p N_r N_t$ times multiplication per prediction. Since a small filter order such as $p=4$ is generally optimal and thus $(d+3)N_H > p$, it is concluded that the KF predictor is computationally simpler than the RNN predictor.

Further, it is meaningful to make clear how many computing resources are required. The number of pilot sub-carriers per OFDM symbol is N/N_P , and there are f_s/N OFDM symbols per second, from which the number of predictions per second can be figured out, i.e., $\psi=f_s/N_P$. The required multiplications per second by the KF predictor is exactly the product of ψ and Ω_{kf} , i.e., $\Omega_{kf}^{(s)}=f_s \Omega_{kf}/N_P$, and $\Omega_{rnn}^{(s)}=f_s \Omega_{rnn}/N_P$ in the case of the RNN predictor. In terms of the parameters given in Table I and assume $N_P=4$, we have $\Omega_{kf}^{(s)}=4 \times 10^6$ and $\Omega_{rnn}^{(s)}=60 \times 10^6$. Compared with off-the-shelf digital signal processors (DSPs), e.g., TI 66AK2x that provides a capability of nearly 2×10^4 Million Instructions executed Per Second (MIPS), the required resource of the RNN predictor is 0.3%. Even if in a massive MIMO system with a dimension of 32×4 , it consumes 10% of the resource of a single DSP. In summary, the required computing resource of channel prediction is affordable, which is promising from the practical perspective.

B. Performance

To train the RNN, we build a training data set that contains a series of consecutive CSI $\{\tilde{\mathbf{H}}_{tr}[t] | t=1, 2, \dots\}$ extracted from an arbitrary sub-carrier during 10 periods of fluctuation (i.e., channel's coherence time). A training process starts from an initial state where all weights are randomly selected. At iteration t , feeding the channel matrix $\tilde{\mathbf{H}}_{tr}[t]$ into the RNN, the resultant output is compared with the desired value and the prediction error $\hat{\mathbf{H}}_{tr}[t+D]-\tilde{\mathbf{H}}_{tr}[t+D]$ is propagated back through the network so as to update the weights by means of training algorithms such as Levenberg-Marquardt [26]. This process is iteratively carried out until the RNN reaches a certain convergence condition. In contrast, the KF predictor does not need a training process. Its filter coefficients required in (4) can be figured out if f_d and f_s are known.

Suppose the applied multi-antenna system is a uniform linear array having $N_r=1$ receive and $N_t=4$ transmit antennas, a single transmit antenna with the largest instantaneous channel gain is selected. The outage probability, an important performance metric over fading channels, is defined as $P(R)=\Pr\{\log_2(1+\text{SNR}) < R\}$, where \Pr is the notation of mathematical probability and R means a target end-to-end data rate that is set to 1bps/Hz in our simulation. Three different CSI modes are compared:

- The perfect mode where the transmit antenna in subcarrier n at OFDM symbol $t+D$ is chosen in terms of the actual CSI $\tilde{\mathbf{H}}_n[t+D]$, despite it never exists in practice owing to delay and noise.

TABLE I
SIMULATION PARAMETERS

Parameters	Values
Sampling rate	$f_s = 1\text{MHz}$
Maximum Doppler shifts	$f_d = 70$ or 300Hz
MIMO	4×1
OFDM DFT size	$N = 64$
Channel models	3GPP ETU and EVA [22]
Neural Network	3-layer RNN
Training algorithm	Levenberg-Marquardt [26]
Length of tapped delay	$d = 3$
Number of hidden neuron	$N_H = 10$

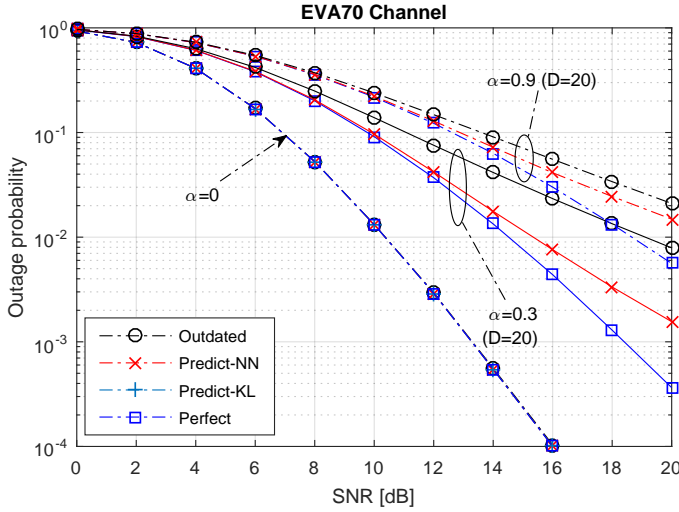


Fig. 3. Performance comparison of using the outdated, predicted, and perfect CSI to select antennas in a MIMO-OFDM system over the EVA channel.

- As in a traditional TAS system, only the outdated CSI $\hat{\mathbf{H}}_n[t]$ is available.
- With the aid of channel prediction, the predicted CSI $\hat{\mathbf{H}}_n[t+D]$ that is possible to closely approximates the actual CSI is used.

Fig.3 illustrates the performance achieved by the MIMO-OFDM system in the EVA channel with a maximal Doppler frequency of $f_d=70\text{Hz}$. First, the RNN predictor is tuned to one-step prediction mode ($D=1$) in order to compare with the KF predictor directly. The multi-antenna channels are independent and identically distributed (*i.i.d.*) as marked in the figure by the correlation coefficient of $\alpha=0$. There is no performance loss due to the outdated CSI and the curves of the perfect, predicted, and outdated modes are identical. That is because the prediction step $D=1$ is equivalent to a time length of $64\mu\text{s}$ that is negligible relative to the coherence time $T_c \approx 1/f_d = 14.28\text{ms}$. In order to make the CSI really outdated, the RNN predictor is then reset to multi-step mode and the prediction step is increased to $D=20$, corresponding to 1.28ms . To check whether the predictor is effective in correlated channels, we impose the correlation matrix recommended in 3GPP LTE standards [22] on the EVA channels. Under the medium correlation indicated by $\alpha=0.3$, the outdated CSI has a performance loss of 4.5dB given $P(R)=10^{-2}$ in comparison with the perfect mode, while the channel prediction can take nearly 4dB back. As mentioned previously, the KF predictor can merely predict one step ahead and therefore is not applicable to this case. Further increased the channel correlation to $\alpha=0.9$, as shown in Fig.3, the performance of multi-step prediction is also clearly better than that of the outdated CSI. With the increase of channel correlation, the system performance degrades because the available spatial diversity gain vanishes gradually, independently of the application of channel prediction.

The performance evaluation is conducted also in *i.i.d.* ETU channels with a maximal Doppler frequency of $f_d=300\text{Hz}$.

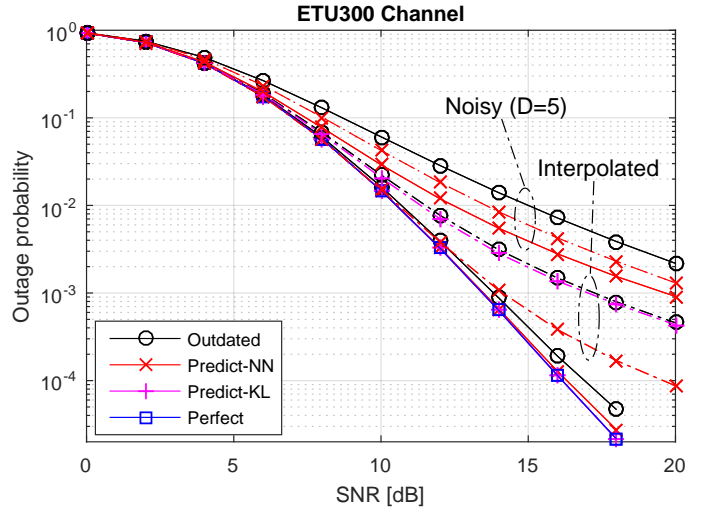


Fig. 4. Performance comparison of using the outdated, predicted, and perfect CSI to select antennas in a MIMO-OFDM system over the ETU channel.

First, the RNN predictor is set to one-step mode corresponding to the same time length of $64\mu\text{s}$, whereas the coherence time drops to $T_c \approx 1/f_d = 3.3\text{ms}$. The effect of the outdated CSI is not negligible this time, as shown in Fig.4, where a performance loss of about 0.7dB is observed at $P(R)=10^{-4}$. Using the curve of the perfect mode as a benchmark, the KF predictor achieves the optimal performance. Although the RNN predictor is slightly inferior to the KF predictor, it is still quite close to the optimal performance and outperforms the outdated mode clearly. As mentioned in Section III, channel interpolation is involved in the process of channel estimation, so the effect of interpolation errors that is defined as the difference between the perfect CSI and the interpolated CSI is taken into account. For the purpose of a better illustration, the results for pilot's insertion interval $N_P=3$ are selected to show in the figure. Even in the mode of one-step prediction $D=1$, the outdated CSI has a remarkable loss of 3.6dB in comparison with the perfect mode at the outage probability of 10^{-3} . The KF predictor is vulnerable to interpolation errors, as shown in the figure, which achieves a worse result that is very comparable to the outdated mode. In contrast, the RNN predictor outperforms the KF predictor with an SNR gain of 3.1dB . In addition to interpolation errors, the available CSI is also impaired by estimation errors since additive noise cannot be avoided in the process of channel estimation. Under the assumption that the SNR of pilots is $\text{SNR}_p=20\text{dB}$, a performance evaluation with a prediction step of $D=5$ is conducted. The results reveal that additive noise has a notable impact on the performance, but the RNN predictor still receives an SNR gain of around 2dB at $P(R)=10^{-2}$ compared with that of the outdated CSI. If the SNR of pilots is further increased to $\text{SNR}_p=30\text{dB}$, a performance gain of nearly 3dB can be expected. In conclusion, the RNN predictor is effective to against the outdated CSI no matter whether the frequency-selective channels are noiseless or noisy, independent or correlated.

VI. CONCLUSIONS

This paper proposed a recurrent neural network-based frequency-domain predictor for frequency-selective multi-antenna channels. The application of this channel predictor into a MIMO-OFDM system so as to improve the correctness of selecting transmit antennas at the transmitter was illustrated. Performance assessment was carried out in multi-path fading environment specified by 3GPP EVA and ETU channel models taking into account the influential factors including spatial correlation, the Doppler shift, as well as interpolation and estimation errors. Numerical results verified the effectiveness of the RNN predictor to combat the outdated CSI no matter whether channels are noiseless or noisy, correlated or independent. Although its computational complexity is higher than the Kalman filter, the required computing resource is still affordable relative to off-the-shelf hardware. More importantly, the proposed predictor has a flexibility of conducting multi-step prediction and is more robust against interpolation errors.

REFERENCES

- [1] T. R. Ramya and S. Bhashyam, "Using delayed feedback for antenna selection in MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6059–6067, Dec. 2009.
- [2] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338–351, Sep. 2013.
- [3] J. L. Vicario *et al.*, "Opportunistic relay selection with outdated CSI: outage probability and diversity analysis," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2872–2876, Jun. 2009.
- [4] P. Aquilina and T. Ratnarajah, "Performance analysis of IA techniques in the MIMO IBC with imperfect CSI," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1259–1270, Feb. 2015.
- [5] Z. Wang *et al.*, "Resource allocation in OFDMA networks with imperfect channel state information," *IEEE Commun. Lett.*, vol. 18, no. 9, pp. 1611–1614, Sep. 2014.
- [6] A. Hyadi *et al.*, "An overview of physical layer security in wireless communication systems with CSIT uncertainty," *IEEE Access*, vol. 4, pp. 6121–6132, Sep. 2016.
- [7] A. Weinand *et al.*, "Providing physical layer security for mission critical machine type communication," in *Proc. of IEEE ETFA*, Berlin, Germany, Sep. 2016.
- [8] TACNET 4.0 project. [Online]. Available: <http://www.tacnet40.de>
- [9] W. Jiang *et al.*, "A robust opportunistic relaying strategy for co-operative wireless communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2642–2655, Apr. 2016.
- [10] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec. 2007.
- [11] T. Eyceoz *et al.*, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *IEEE Commun. Lett.*, vol. 2, no. 9, pp. 254–256, Sep. 1998.
- [12] A. Duel-Hallen *et al.*, "Long-range prediction of fading signals," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.
- [13] W. Peng *et al.*, "Channel prediction in time-varying Massive MIMO environments," *IEEE Access*, vol. 5, pp. 23 938–23 946, Nov. 2017.
- [14] J.-Y. Wu and W.-M. Lee, "Optimal linear channel prediction for LTE-A uplink under channel estimation errors," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 4135–4142, Oct. 2013.
- [15] J. Connor *et al.*, "Recurrent neural networks and robust time series prediction," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 240–254, Mar. 1994.
- [16] W. Liu *et al.*, "Recurrent neural network based narrowband channel prediction," in *Proc. IEEE Vehicular Tech. Conf. (VTC)*, Melbourne, Australia, May 2006.
- [17] C. Potter *et al.*, "MIMO beam-forming with neural network channel prediction trained by a novel PSO-EA-DEPSO algorithm," in *Proc. IEEE Intl. Joint Conf. on Neural Networks (IJCNN)*, Hong Kong, China, Jun. 2008.
- [18] K. T. Truong and R. W. Heath, "Fading channel prediction based on combination of complex-valued neural networks and chirp Z-transform," *IEEE Trans. Neural Netw.*, vol. 25, no. 9, pp. 1686–1695, Sep. 2014.
- [19] R.-F. Liao *et al.*, "The Rayleigh fading channel prediction via deep learning," *Wireless Commu. and Mobile Compu.*, Jul. 2018.
- [20] W. Jiang and H. D. Schotten, "Multi-antenna fading channel prediction empowered by artificial intelligence," in *Proc. IEEE Vehicular Tech. Conf. (VTC)*, Chicago, USA, Aug. 2018.
- [21] —, "Neural network-based channel prediction and its performance in multi-antenna systems," in *Proc. IEEE Vehicular Tech. Conf. (VTC)*, Chicago, USA, Aug. 2018.
- [22] *3GPP TS36.104: Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception*, Sep. 2018, v15.4.0.
- [23] W. Jiang and T. Kaiser, *Signal Processing for 5G: Algorithms and Implementations*. United Kindom: John Wiley&Sons, 2016, ch. 3.
- [24] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*, 1st ed. Prentice-Hall, 1975.
- [25] K. Baddour and N. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650–1662, Jul. 2005.
- [26] X. Fu *et al.*, "Training recurrent neural networks with the Levenberg Marquardt algorithm for optimal control of a grid-connected converter," *IEEE Trans. Neural Netw.*, vol. 26, no. 9, pp. 1900–1912, Sep. 2015.
- [27] H. Zhang and R. U. Nabar, "Transmit antenna selection in MIMO-OFDM systems: Bulk versus per-tone selection," in *Proc. IEEE Intl. Conf. on Commu. (ICC)*, Beijing, China, May 2008.
- [28] X. Yu *et al.*, "Unified performance analysis of transmit antenna selection with OSTBC and imperfect CSI over Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 494–508, Jan. 2018.