



# Robust Beamforming Under Channel Prediction Errors for Time-Varying MIMO System

ZHU Yuting<sup>1</sup>, LI Zeng<sup>2,3</sup>, ZHANG Hongtao<sup>1</sup>

(1. Key Lab of Universal Wireless Communications, Ministry of Education of China, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. ZTE Corporation, Shenzhen 518057, China;

3. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China)

DOI: 10.12142/ZTECOM.202303011

<https://link.cnki.net/urlid/34.1294.TN.20230830.1912.007>, published online August 31, 2023

Manuscript received: 2022-12-23

**Abstract:** The accuracy of acquired channel state information (CSI) for beamforming design is essential for achievable performance in multiple-input multiple-output (MIMO) systems. However, in a high-speed moving scene with time-division duplex (TDD) mode, the acquired CSI depending on the channel reciprocity is inevitably outdated, leading to outdated beamforming design and then performance degradation. In this paper, a robust beamforming design under channel prediction errors is proposed for a time-varying MIMO system to combat the degradation further, based on the channel prediction technique. Specifically, the statistical characteristics of historical channel prediction errors are exploited and modeled. Moreover, to deal with random error terms, deterministic equivalents are adopted to further explore potential beamforming gain through the statistical information and ultimately derive the robust design aiming at maximizing weighted sum-rate performance. Simulation results show that the proposed beamforming design can maintain outperformance during the downlink transmission time even when channels vary fast, compared with the traditional beamforming design.

**Keywords:** time-varying channels; time-division duplex; robust beamforming; channel prediction errors; weighted sum-rate maximization

**Citation** (Format 1): ZHU Y T, LI Z, ZHANG H T. Robust beamforming under channel prediction errors for time-varying MIMO system [J]. *ZTE Communications*, 2023, 21(3): 77 – 85. DOI: 10.12142/ZTECOM.202303011

**Citation** (Format 2): Y. T. Zhu, Z. Li, and H. T. Zhang, “Robust beamforming under channel prediction errors for time-varying MIMO system,” *ZTE Communications*, vol. 21, no. 3, pp. 77 – 85, Sept. 2023. doi: 10.12142/ZTECOM.202303011.

## 1 Introduction

The rapid growth of intelligent devices and applications generates large demands for high data rate transmission. Massive multiple-input multiple-output (MIMO) technique, as a key technology in the fifth generation of wireless networks, provides huge potential capacity gain by employing multiple antennas and exploiting the extra degree of freedom without extending the extra bandwidth<sup>[1]</sup>. The achievable performance heavily relies on the exploitation of spatial multiplexing gain, as well as enough channel knowledge of the base station<sup>[2]</sup>. The former is tightly related to multi-user MIMO (MU-MIMO) communications and appropriate beamforming design to enhance the intended signal and suppress unintended interference. As to the latter, in a time-division duplex (TDD) mode, the strategy used to obtain chan-

nel state information (CSI) is reciprocity, where the transmitter uses uplink (UL) channel information to speculate the downlink (DL) channel state in the next transmission interval<sup>[3]</sup>. However, this strategy cannot face the channel aging problem, especially in a high-mobility environment where the channel characteristic is varying fast during the transmission period, which leads to a serious impact on the performance of beamforming in the MIMO system.

Traditional beamforming algorithms, such as classic zero-forcing (ZF) and weighted minimum mean-squared-error (WMMSE), actually work well only when accurate and instantaneous CSI is obtained. Note that the WMMSE precoder<sup>[4-5]</sup> is designed according to the sum-rate maximization criterion so it performs better than ZF. In the time-varying TDD system, the channel estimation for multiple users is based on their sounding reference signals (SRS), which means we can only obtain the accurate CSI of the current SRS transmission time slot while cannot measure the channel state of the intermediate DL time slots in the case of a high-speed moving scene.

This work was supported by the ZTE Industry-University-Institute Cooperation Funds under Grant No. 2021ZTE01-03.

Thus, the acquired CSI is inevitably outdated, leading to outdated beamforming design and then performance degradation. Considering the CSI used for DL transmission is not always perfect in practice, several studies of robust beamforming have been carried out<sup>[6-7]</sup>. Ref. [6] introduced an iterative algorithm to maximize the linear assignment weighted sum rate with statistical CSI under channel fading conditions. In Ref. [7], a robust coordinated beamforming method is proposed to maintain the weighted sum-rate performance with inaccurate CSI. The channel aging effect still weakens the robust benefit dramatically<sup>[8]</sup>, so more accurate CSI is needed from the viewpoint of CSI, which can be achieved by channel prediction technology.

Channel prediction technology is an effective way to combat the detrimental effects of channel aging and keep the track of the time-varying channel<sup>[9]</sup>, which predicts the future channels by exploiting the potential temporal correlation from those in the past<sup>[10]</sup>. There are various prediction methods from autoregressive (AR) model-based methods to deep learning<sup>[11]</sup>. The AR model is commonly used, which treats the time-varying channel as a wide-sense stationary stochastic process<sup>[9, 12]</sup>, and deep learning requires a mass of samples for training<sup>[13]</sup>. Besides, by utilizing the time and spatial properties of channels, the authors in Ref. [14] proposed an angle-domain channel tracking scheme for high-speed railway communications, and spatial-temporal sparse structures are further exhibited to propose a novel estimation and prediction scheme for time-varying massive MIMO channels<sup>[15]</sup>. Furthermore, to solve the problem of tracking the nonlinear channels, the kernel recursive least squares (KRLS)<sup>[16-17]</sup> algorithm is proposed, mapping the channel sample space to the high-dimensional space, and it shows great adaptiveness.

Based on the predicted CSI, linear or nonlinear interpolation can be employed for predicting the DL channel between the current channel and the predicted one<sup>[18]</sup>. Ref. [19] evaluated the benefit of channel prediction in adaptive beamforming systems and showed the tolerable Doppler spread increase for a given outage performance by using prediction. However, there still exist errors in the channel prediction, which becomes more obvious as the mobility increases<sup>[20]</sup>. The existence of errors between the acquired CSI and the real-time CSI will result in the performance error cost since the designed beamforming does not match the real channel. However, most of the existing works focus on the impact that the channel prediction error brings<sup>[21]</sup>, while no robust beamforming for improving the performance under channel prediction errors is considered. Moreover, the channel prediction errors are often constructed as complex Gaussian random variables with independent and identically distributed (i.i.d.), zero mean and unit variance entries, which does not consider the statistical characteristics.

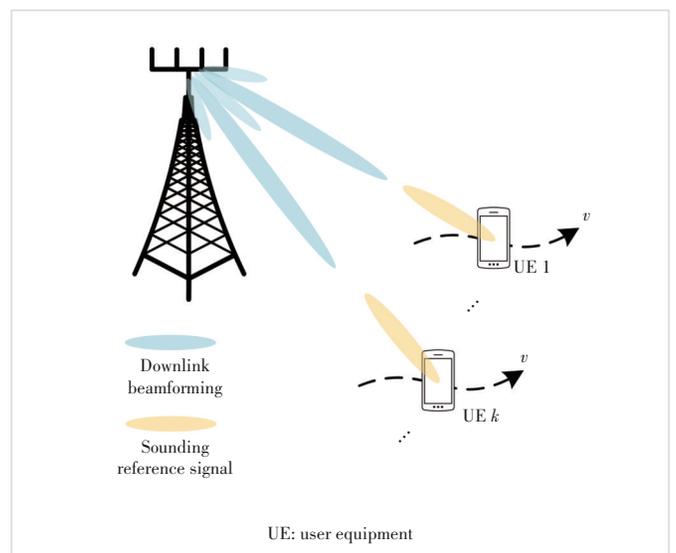
Motivated by the above, we propose a robust beamforming design under channel prediction errors based on channel pre-

dition for the time-varying MIMO system in this paper. The main contributions of this work are summarized as follows. 1) A framework of TDD system for beamforming based on channel prediction is introduced, with a novel prediction error model that combines channel statistical information and a Gaussian random matrix, including the joint correlation properties of realistic massive MIMO channels. 2) To maximize the sum-rate performance, we propose a robust beamforming design based on channel prediction, utilizing the deterministic equivalents method to explore further potential beamforming gain brought by the static statistical information, breaking through the performance limitation brought by the channel prediction errors. 3) The beamforming gains compared with traditional benchmarks under different mobile scenes are validated by simulation, and we investigate the performance during the half-frame period to reflect the loss caused by the time-varying effect.

The rest of this paper is organized as follows. Section 2 presents the proposed framework of the system including the system model and the formulation of the discussed problem. Section 3 introduces the proposed robust beamforming design. Section 4 provides the simulation results and performance discussion. Finally, we conclude this paper in Section 5.

## 2 System Model and Problem Formulation

In this paper, we focus on the TDD system, and the CSI used for DL transmission is obtained depending on the channel reciprocity. As shown in Fig. 1, we consider a MU-MIMO system where one base station (BS) equipped with  $N_t$  transmit antennas serves  $K$  pieces of mobile user equipment (UE), and each is equipped with  $N_r$  receive antennas. Denote  $\mathbf{H}_{i,k}^S \in \mathbb{C}^{N_r \times N_t}$  as the CSI matrix spanning from the BS to the  $k$ -th piece of UE



▲ Figure 1. Illustration of a time-division duplex (TDD) multi-user multiple-input multiple-output (MU-MIMO) system where  $v$  denotes the speed of mobile UE

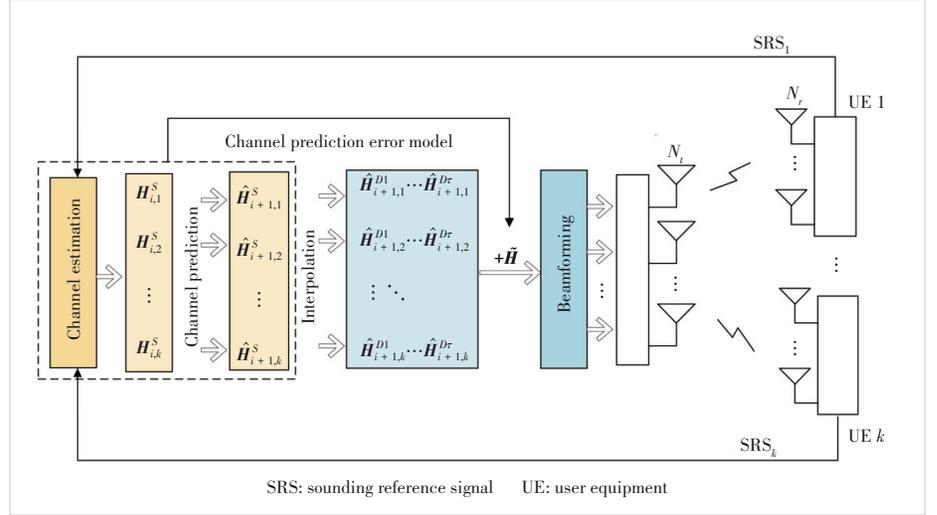
obtained through SRS in the  $i$ -th half-frame and  $\mathbf{H}_{i,k}^{Dr} \in \mathbb{C}^{N_r \times N_t}$  as the CSI matrix for DL transmission in the  $t$ -th DL time slot of the  $i$ -th half-frame period. As illustrated in Fig. 2, the whole procedure of channel acquisition can be divided into three steps, which are channel estimation, channel prediction and interpolation. Based on the channel prediction algorithm and interpolation for the DL channel between the currently estimated channel and the predicted one,  $\mathbf{H}_{i,k}^{Dt}$  used for DL beamforming depends on the predicted DL channel matrix  $\hat{\mathbf{H}}_{i,k}^{Dt}$ , and the channel prediction error is considered further.

### 2.1 Channel Prediction

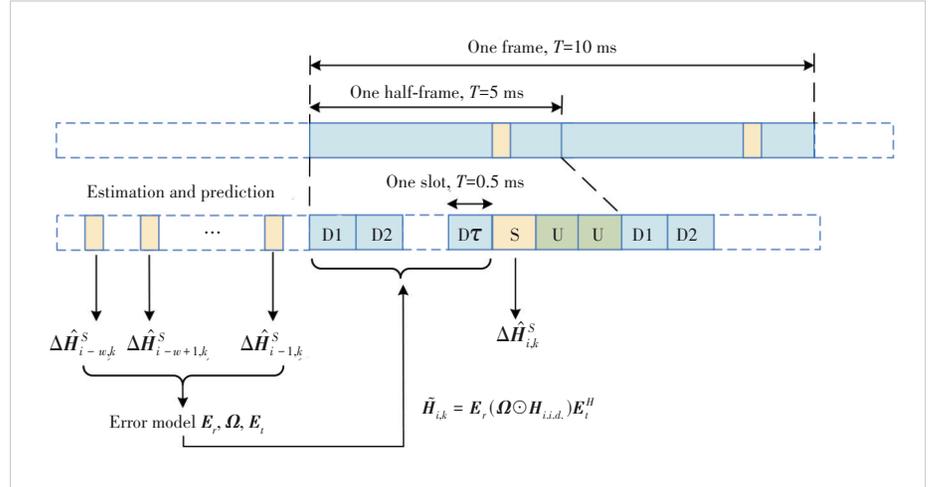
In the TDD system, firstly, by exploiting the reciprocity between UL and DL, we can obtain the uplink channel information by channel estimation such as minimum mean square error (MMSE) and least square (LS), and use them to calculate parameters including beamforming for DL transmission. The channel estimation is to diminish the effect of noise and extract the channel model from received data. In this paper, we assume the channel estimation is perfect and we can obtain the estimated CSI for UE  $\mathbf{H}_{i,1}^S, \mathbf{H}_{i,2}^S, \dots, \mathbf{H}_{i,K}^S$  through the processing of SRS.

Then, based on a certain number of past CSI samples, we can predict the future state of channels for UE  $\hat{\mathbf{H}}_{i+1,1}^S, \hat{\mathbf{H}}_{i+1,2}^S, \dots, \hat{\mathbf{H}}_{i+1,K}^S$  by exploiting the correlation characteristics of the channels. In this paper, to track nonlinear channels and predict the time-varying channel, we adopt a sparse sliding-window KRLS algorithm, where the sample set is updated dynamically based on the correlation analysis among samples. And it is to be mentioned that the proposed beamforming scheme in the following is not limited to a certain prediction method.

During the special time slot for SRS, with the currently estimated channel and predicted one  $\hat{\mathbf{H}}_{i+1,k}^S$ , the next step is to interpolate the DL channel between them and here we take the linear interpolation method. Define  $\tau$  and  $N_{\text{slot}}$  as the number of DL time slots and that of the total slots in one half-frame, respectively. Assuming that besides the DL time slots there is one SRS slot and two UL slots, with the detail of the time slot structure shown in Fig. 3, the deployed linear interpolation is expressed as



▲ Figure 2. System model for proposed beamforming scheme



▲ Figure 3. Channel prediction error model and time slot structure

$$\hat{\mathbf{H}}_{i+1,k}^{Dt} = \mathbf{H}_{i,k}^S + (t+2) \cdot \frac{\hat{\mathbf{H}}_{i+1,k}^S - \mathbf{H}_{i,k}^S}{N_{\text{slot}}}, \quad 1 \leq t \leq \tau. \quad (1)$$

### 2.2 Channel Prediction Error Model

Considering that the predicted channels obtained by channel prediction techniques still exist errors inevitably, compared with the real channels, we take channel prediction errors into consideration and introduce an error model. With  $\mathbf{H}_{i,k}^S$  denoting the perfect estimated CSI for the  $k$ -th piece of UE in the  $i$ -th SRS period, the former channel prediction errors collected can be described as

$$\Delta \mathbf{H}_{i,k}^S = \mathbf{H}_{i,k}^S - \hat{\mathbf{H}}_{i,k}^S, \quad (2)$$

and a sliding-window size  $w$  is set for collection, as shown in Fig. 3. By exploiting the statistical characteristics of the historical channel prediction error sample set  $[\Delta \mathbf{H}_{i-w,k}^S, \Delta \mathbf{H}_{i-w+1,k}^S, \dots, \Delta \mathbf{H}_{i-1,k}^S]$ , we use the jointly corre-

lated channel model to describe the potential correlations of each channel prediction error in the past slots and express the error model for the  $k$ -th piece of UE in the DL time slots of the  $i$ -th half-frame period,

$$\tilde{\mathbf{H}}_{i,k} = \mathbf{E}_r (\mathbf{\Omega} \odot \mathbf{H}_{i,i.d.}) \mathbf{E}_l^H. \quad (3)$$

The statistical information  $\mathbf{E}_r$ ,  $\mathbf{E}_l$  and  $\mathbf{\Omega}$  are gathered from practical error samples for the UE, where  $\mathbf{E}_r$  and  $\mathbf{E}_l$  are deterministic unitary matrices that reveal the correlation of the transmit and receive antennas respectively, while each element of  $\mathbf{\Omega}$  is the square root of the element of the eigenmode channel coupling matrix.  $\mathbf{H}_{i,i.d.}$  is a random matrix with zero mean and unit variance entries, following i.i.d. For illustration purposes, we depict the error model with the time slot structure in Fig. 3.

For simplicity, we use  $\mathbf{H}_k$  as a substitute for  $\mathbf{H}_{i,k}^{Di}$  in the following part to denote the CSI matrix for the  $k$ -th UE considering the beamforming design for only one typical DL transmission time slot. Based on the analysis mentioned above, the channel matrix for UE can be divided into two parts: the determined prediction channel matrix  $\hat{\mathbf{H}}_k$  and the random error matrix  $\tilde{\mathbf{H}}_k$ .

### 2.3 Problem Formulation

Denote the linear precoding matrix for the  $k$ -th piece of UE as  $\mathbf{V}_k \in \mathbb{C}^{N_t \times 1}$ . The receive signal of the  $k$ -th piece of UE  $\mathbf{y}_k$  at this DL time slot is

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{V}_k \mathbf{s}_k + \sum_{m=1, m \neq k}^K \mathbf{H}_k \mathbf{V}_m \mathbf{s}_m + \mathbf{n}_k, \quad (4)$$

where the transmit signal to the  $k$ -th piece of UE  $\mathbf{s}_k$  is independent random variables with zero mean and unit variance;  $\mathbf{n}_k$  represents the additive white Gaussian noise with the distribution  $\text{CN}(0, \sigma^2 \mathbf{I}_{N_r})$ . Then, the achievable data rate for UE at a DL time slot can be expressed as

$$R_k = \mathbb{E} \left\{ \log \det \left( \mathbf{I} + \mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_k^H \left( \sigma^2 \mathbf{I} + \sum_{m=1, m \neq k}^K \mathbf{H}_k \mathbf{V}_m \mathbf{V}_m^H \mathbf{H}_k^H \right)^{-1} \right) \right\}. \quad (5)$$

To simplify the objective, we treat the aggregate interference-plus-noise as the Gaussian noise for UE and denote its covariance matrix as

$$\mathbf{Q}_k = \sigma^2 \mathbf{I} + \mathbb{E} \left\{ \sum_{m=1, m \neq k}^K \mathbf{H}_k \mathbf{V}_m \mathbf{V}_m^H \mathbf{H}_k^H \right\}. \quad (6)$$

Then the achievable rate for UE at a DL time slot can be re-

written as

$$R_k = \mathbb{E} \left\{ \log \det \left( \mathbf{I} + \mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{Q}_k^{-1} \right) \right\}. \quad (7)$$

Furthermore, the achievable weighted sum rate of the system can be expressed as

$$R = \frac{\tau}{N_{\text{slot}}} \sum_{k=1}^K \alpha_k R_k. \quad (8)$$

Aiming at optimizing the beamforming matrix to maximize the expected weighted sum rate at a typical DL slot, we can express the optimization problem as

$$\begin{aligned} \max_{\{\mathbf{V}_k\}} & \sum_{k=1}^K \alpha_k R_k, \\ \text{s.t.} & \sum_{k=1}^K \text{tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq P, \end{aligned} \quad (9)$$

where  $\{\alpha_k\}$  is the weighting coefficient to ensure fairness among the UE and  $P$  denotes the power budget of the BS. Based on the channel model where prediction errors are taken into consideration, we investigate the optimal robust beamforming design for the proposed problem in the next section.

## 3 Proposed Robust Beamforming Design

Due to the randomness introduced by the channel prediction error, it is difficult to obtain the close-form expressions of the objective. To deal with the randomness, a robust beamforming design is proposed by considering deterministic equivalents. The NP-hard Problem (9) can be resolved by exploiting the relationship between the rate and the mean-square error (MSE) matrix, inspired by Ref. [5]. Denote the estimated signal at the receiver as  $\hat{\mathbf{s}}_k = \mathbf{U}_k^H \mathbf{y}_k$ , where  $\mathbf{U}_k$  denotes the receive beamformer of the  $k$ -th piece of UE. Assuming that signal  $\mathbf{s}_k$  and noise  $\mathbf{n}_k$  are independent, the MSE matrix can be written as

$$\begin{aligned} \mathbf{MSE}_k &= \mathbb{E}_{\mathbf{s}_k, \mathbf{n}_k} \{ (\hat{\mathbf{s}}_k - \mathbf{s}_k)(\hat{\mathbf{s}}_k - \mathbf{s}_k)^H \} = \\ & (\mathbf{I} - \mathbf{U}_k^H \mathbf{H}_k \mathbf{V}_k)(\mathbf{I} - \mathbf{U}_k^H \mathbf{H}_k \mathbf{V}_k)^H + \sum_{m \neq k}^K \mathbf{U}_k^H \mathbf{H}_k \mathbf{V}_m \mathbf{V}_m^H \mathbf{H}_k^H \mathbf{U}_k + \sigma^2 \mathbf{U}_k^H \mathbf{U}_k. \end{aligned} \quad (10)$$

Fixing all of the transmitting beamforming matrices and adopting the well-known MMSE receiver, we can rewrite the MSE matrix as

$$\begin{aligned} \mathbf{E}_k &= \mathbf{MSE}_k(\mathbf{U}_k^{\text{mmse}}) = \\ & \mathbf{I} - \mathbf{V}_k^H \mathbf{H}_k^H (\mathbf{Q}_k + \mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_k^H)^{-1} \mathbf{H}_k \mathbf{V}_k. \end{aligned} \quad (11)$$

Note that  $R_k = \mathbb{E} \{ \log \det(\mathbf{E}_k^{-1}) \}$  can be expressed as a convex function of  $\mathbf{E}_k$ . Using the first-order condition of the convex function, we can obtain

$$\begin{aligned} & \mathbb{E}\{\log \det(\mathbf{E}_k)^{-1}\} \geq \mathbb{E}\{\log \det(\mathbf{E}_k^{(d)})^{-1}\} - \\ & \mathbb{E}\{\text{tr}(\mathbf{E}_k^{(d)})^{-1}(\mathbf{E}_k - \mathbf{E}_k^{(d)})\} \geq \\ & \mathbb{E}\{\log \det(\mathbf{E}_k^{(d)})^{-1}\} + \text{tr}(\mathbf{I}) - \mathbb{E}\{\text{tr}(\mathbf{E}_k^{(d)})^{-1}\mathbf{MSE}_k\}, \end{aligned} \quad (12)$$

where  $\mathbf{E}_k^{(d)}$  represents  $\mathbf{E}_k$  with a fixed set of beamforming matrix  $\{\mathbf{V}_k^{(d)}\}$  at the  $d$ -th iteration,  $k = 1, \dots, K$ . By ignoring the first two constant terms of the right side of Inequality (12), the last part can be expressed as a function of  $\mathbf{V}_k$ . For brevity, define functions  $\eta_k^{\text{pri}}[\mathbf{f}] = \mathbb{E}\{\mathbf{H}_k \mathbf{f} \mathbf{H}_k^H\}$ ,  $\tilde{\eta}_k^{\text{pri}}[\mathbf{f}] = \mathbb{E}\{\mathbf{H}_k^H \mathbf{f} \mathbf{H}_k\}$ ,  $\eta_k[\mathbf{f}] = \mathbb{E}\{\tilde{\mathbf{H}}_k \mathbf{f} \tilde{\mathbf{H}}_k^H\}$ , and  $\tilde{\eta}_k[\mathbf{f}] = \mathbb{E}\{\tilde{\mathbf{H}}_k^H \mathbf{f} \tilde{\mathbf{H}}_k\}$ . Denote the whole right side of Inequality (12) as  $g(\mathbf{V}_k | \{\mathbf{V}_k^{(d)}\})$ , which is described in detail as

$$\begin{aligned} g(\mathbf{V}_k | \{\mathbf{V}_k^{(d)}\}) &= c_k^{(d)} + \text{tr}((\mathbf{A}_k^{(d)})^H \mathbf{V}_k) + \text{tr}(\mathbf{A}_k^{(d)} \mathbf{V}_k^H) - \\ & \text{tr}(\mathbf{B}_k^{(d)} \mathbf{V}_k \mathbf{V}_k^H) - \text{tr}\left(\mathbf{C}_k^{(d)} \sum_{m \neq k} \mathbf{V}_m \mathbf{V}_m^H\right), \end{aligned} \quad (13)$$

where

$$\begin{aligned} c_k^{(d)} &= \mathbb{E}\{\log \det(\mathbf{E}_k^{(d)})^{-1}\} + \text{tr}(\mathbf{I}) - \mathbb{E}\{\text{tr}((\mathbf{E}_k^{(d)})^{-1})\} - \\ & \sigma^2 \mathbb{E}\left\{\text{tr}\left((\mathbf{E}_k^{(d)})^{-1}(\mathbf{U}_k^{(d)})^H \mathbf{U}_k^{(d)}\right)\right\}, \end{aligned} \quad (14)$$

$$\mathbf{A}_k^{(d)} = \hat{\mathbf{H}}_k^H (\mathbf{Q}_k^{(d)})^{-1} \hat{\mathbf{H}}_k \mathbf{V}_k^{(d)} + \tilde{\eta}_k[\mathbf{Q}_k^{(d)}]^{-1} \mathbf{V}_k^{(d)}, \quad (15)$$

$$\begin{aligned} \mathbf{B}_k^{(d)} &= \hat{\mathbf{H}}_k^H (\mathbf{Q}_k^{(d)})^{-1} \hat{\mathbf{H}}_k + \tilde{\eta}_k[\mathbf{Q}_k^{(d)}]^{-1} - \mathbb{E}\left\{\mathbf{H}_k^H (\mathbf{Q}_k^{(d)} + \right. \\ & \left. \mathbf{H}_k \mathbf{V}_k^{(d)} (\mathbf{V}_k^{(d)})^H \mathbf{H}_k^H)^{-1} \mathbf{H}_k\right\}, \end{aligned} \quad (16)$$

$$\begin{aligned} \mathbf{C}_k^{(d)} &= \mathbb{E}\left\{\mathbf{H}_k^H \left((\mathbf{Q}_k^{(d)})^{-1} - \mathbb{E}\left\{(\mathbf{Q}_k^{(d)} + \right. \right. \right. \\ & \left. \left. \mathbf{H}_k \mathbf{V}_k^{(d)} (\mathbf{V}_k^{(d)})^H \mathbf{H}_k^H\right\}^{-1}\right) \mathbf{H}_k\right\}. \end{aligned} \quad (17)$$

Denote  $f(\{\mathbf{V}_k^{(d)}\})$  as the left side of Inequality (12) and  $g(\mathbf{V}_k | \{\mathbf{V}_k^{(d)}\})$  can be regarded as the lower-bounding function which minorizes the objective function  $f(\{\mathbf{V}_k^{(d)}\})$ , since it is a convex function of  $\mathbf{V}_k$  that satisfies  $g(\mathbf{V}_k | \{\mathbf{V}_k^{(d)}\}) \leq f(\{\mathbf{V}_k^{(d)}\})$  and  $g(\mathbf{V}_k | \{\mathbf{V}_k^{(d)}\}) = f(\{\mathbf{V}_k^{(d)}\})$ . It can be iteratively optimized and converge to a stationary point, which has been proved in Ref. [22].

The update of  $\mathbf{V}_k$  for all users can be decoupled across transmitters. Then, we update beamforming matrices as  $\mathbf{V}_k^{(d+1)} = \arg \max_{\mathbf{V}_k} g(\mathbf{V}_k | \{\mathbf{V}_k^{(d)}\})$ , and the optimal solution to the

equation is obtained by the Lagrange multipliers method with Lagrange multiplier  $\mu$ . According to the first-order optimal conditions, the iterative equation of the beamformer can be obtained as

$$\mathbf{V}_k^{(d+1)} = (\mathbf{D}_k^{(d)} + \mu \mathbf{I})^{-1} (\alpha_k \mathbf{A}_k^{(d)}), \quad (18)$$

where

$$\mathbf{D}_k^{(d)} = \alpha_k \mathbf{B}_k^{(d)} + \sum_{m \neq k} \alpha_m \mathbf{C}_m^{(d)}, \quad (19)$$

and  $\mu$  can be obtained by using the bisection method.

For the last part of  $\mathbf{B}_k^{(d)}$  and  $\mathbf{C}_m^{(d)}$ , which have not been transformed into the deterministic matrix, there are complicated calculations of the random variables, leading to difficulty in obtaining closed-form expressions. Drawing on the essence of Ref. [23], we will utilize deterministic equivalents to obtain the approximate closed-form expression of MIMO capacity, and the derivation process is presented as follow.

Note that  $R_k = \mathbb{E}\{\log \det(\mathbf{I} + \mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{Q}_k^{-1})\}$  can be rewritten as  $\mathbb{E}\left\{\log \det\left(\mathbf{I} + \mathbf{Q}_k^{-\frac{1}{2}} \mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{Q}_k^{-\frac{1}{2}}\right)\right\}$ . Take  $\mathbf{Q}_k^{-\frac{1}{2}} \mathbf{H}_k \mathbf{V}_k$  as a whole part  $\mathcal{H}_k$  and denote the term  $\mathbf{Q}_k^{-\frac{1}{2}} \mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{Q}_k^{-\frac{1}{2}}$  as a Hermitian matrix  $\mathbf{Z}_{N_r, k}$ . Let  $F_{\mathbf{Z}_{N_r, k}}(\lambda)$  denote the expected cumulative distribution of the eigenvalues  $\lambda_1, \dots, \lambda_{N_r}$  of  $\mathbf{Z}_{N_r, k}$ , and the Shannon transform  $\mathcal{V}_{\mathbf{Z}_{N_r, k}}$  is described as  $\mathcal{V}_{\mathbf{Z}_{N_r, k}}(x) = \int_0^\infty \log\left(1 + \frac{1}{x} \lambda\right) dF_{\mathbf{Z}_{N_r, k}}(\lambda)$ . Then  $R_k$  can be simplified as  $\sum \mathbb{E}_{\lambda_i}[\log(1 + \lambda_i)]$  and translated into  $N_r \int_0^\infty \log(1 + \lambda) dF_{\mathbf{Z}_{N_r, k}}(\lambda)$ , finally equivalent to  $N_r \mathcal{V}_{\mathbf{Z}_{N_r, k}}(1)$ . The Stieltjes transform for  $F_{\mathbf{Z}_{N_r, k}}(\lambda)$  is

$$\mathcal{S}_{\mathbf{Z}_{N_r, k}}(y) = \int \frac{1}{y - \lambda} dF_{\mathbf{Z}_{N_r, k}}(\lambda) = \frac{1}{N_r} \mathbb{E}\{\text{tr}[(y \mathbf{I}_{N_r} - \mathbf{Z}_{N_r, k})^{-1}]\}. \quad (20)$$

Then the relation between the Stieltjes transform  $\mathcal{S}_{\mathbf{Z}_{N_r, k}}(y)$  and the Shannon transform  $\mathcal{V}_{\mathbf{Z}_{N_r, k}}$  can be established as

$$\mathcal{V}_{\mathbf{Z}_{N_r, k}}(x) = \int_x^\infty \left(\frac{1}{y} + \mathcal{S}_{\mathbf{Z}_{N_r, k}}(-y)\right) dy. \quad (21)$$

Thus, we can obtain the closed-form expression of the ergodic user rate  $R_k$  by establishing the closed-form expression of the Stieltjes transform and Shannon transform. The approximate closed-form Stieltjes transform expressions can be calculated by applying the free probability theory<sup>[24]</sup>, denoted as Eqs. (22) and (23) with Notations (24) - (27).

$$\begin{aligned} \mathcal{S}_{Z_{N_s}}(y\mathbf{I}_{N_r}) &= (y\tilde{\Phi}_k - \\ & (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \hat{\mathbf{H}}_k \mathbf{V}_k^{(d)} \Phi_k^{-1} \mathbf{V}_k^{(d)H} \hat{\mathbf{H}}_k^H (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}})^{-1}, \end{aligned} \quad (22)$$

$$\begin{aligned} \mathcal{S}_{\mathcal{H}_k^u \mathcal{H}_k}(y\mathbf{I}_{N_r}) &= (y\tilde{\Phi}_k - \\ & (\mathbf{V}_k^{(d)})^H \hat{\mathbf{H}}_k^H (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \tilde{\Phi}_k^{-1} (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \hat{\mathbf{H}}_k \mathbf{V}_k^{(d)})^{-1}, \end{aligned} \quad (23)$$

$$\begin{aligned} \tilde{\Phi}_k &= \mathbf{I}_{N_r} - (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \eta_k \left[ (\mathbf{V}_k^{(d)})^H \mathcal{S}_{\mathcal{H}_k^u \mathcal{H}_k}(y\mathbf{I}_{N_r}) \mathbf{V}_k^{(d)} \right] (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}}, \end{aligned} \quad (24)$$

$$\begin{aligned} \Phi_k &= \mathbf{I}_{d_k} - (\mathbf{V}_k^{(d)})^H \tilde{\eta}_k \left[ (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \mathcal{S}_{Z_{N_s}}(y\mathbf{I}_{N_r}) (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \right] \mathbf{V}_k^{(d)}, \end{aligned} \quad (25)$$

$$\begin{aligned} \Gamma_k &= -\tilde{\eta}_k \left[ (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \mathcal{S}_{Z_{N_s}}(-\mathbf{I}_{N_r}) (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \right] + \\ & \hat{\mathbf{H}}_k^H (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \tilde{\Phi}_k^{-1} (\mathbf{Q}_k^{(d)})^{-\frac{1}{2}} \hat{\mathbf{H}}_k, \end{aligned} \quad (26)$$

$$\begin{aligned} \tilde{\Gamma}_k &= -\eta_k \left[ \mathbf{V}_k^{(d)} \mathcal{S}_{\mathcal{H}_k^u \mathcal{H}_k}(-\mathbf{I}_{N_r}) (\mathbf{V}_k^{(d)})^H \right] + \\ & \hat{\mathbf{H}}_k \mathbf{V}_k^{(d)} \Phi_k^{-1} (\mathbf{V}_k^{(d)})^H \hat{\mathbf{H}}_k^H. \end{aligned} \quad (27)$$

Then the deterministic equivalent form of rate  $\bar{R}_k$  can be expressed as:

$$\begin{aligned} \bar{R}_k &= \log \det(\mathbf{I}_{N_r} + \tilde{\Gamma}_k \mathbf{Q}_k^{-1}) + \log \det(\Phi_k) - \\ & \text{tr} \left( \eta_k \left[ \mathbf{V}_k \mathcal{S}_{\mathcal{H}_k^u \mathcal{H}_k}(-\mathbf{I}_{N_r}) \mathbf{V}_k^H \right] \mathbf{Q}_k^{-\frac{1}{2}} \mathcal{S}_{Z_{N_s}}(-\mathbf{I}_{N_r}) \mathbf{Q}_k^{-\frac{1}{2}} \right), \end{aligned} \quad (28)$$

$$\begin{aligned} \bar{R}_k &= \log \det(\mathbf{I}_{N_r} + \Gamma_k \mathbf{V}_k \mathbf{V}_k^H) + \log \det(\tilde{\Phi}_k) - \\ & \text{tr} \left( \tilde{\eta}_k \left[ \mathbf{Q}_k^{-\frac{1}{2}} \mathcal{S}_{Z_{N_s}}(-\mathbf{I}_{N_r}) \mathbf{Q}_k^{-\frac{1}{2}} \right] \mathbf{V}_k \mathcal{S}_{\mathcal{H}_k^u \mathcal{H}_k}(-\mathbf{I}_{N_r}) \mathbf{V}_k^H \right). \end{aligned} \quad (29)$$

Based on the aforementioned analysis, the approximate closed-form expressions of the second part of  $\mathbf{B}_k^{(d)}$  can be derived as Eq. (30) by exploiting the relationship between it and the original expression  $R_k$ . The closed-form expressions of  $\mathbf{C}_k^{(d)}$  can be derived similarly as Eq. (31).

$$\begin{aligned} \mathbb{E} \left\{ \mathbf{H}_k^H (\mathbf{Q}_k^{(d)} + \mathbf{H}_k \mathbf{V}_k^{(d)} (\mathbf{V}_k^{(d)})^H \mathbf{H}_k^H)^{-1} \mathbf{H}_k \right\} &= \\ \frac{\partial R_k}{\partial (\mathbf{V}_k \mathbf{V}_k^H)} &= \frac{\partial \bar{R}_k}{\partial (\mathbf{V}_k \mathbf{V}_k^H)} = (\mathbf{I}_{N_r} + \Gamma_k \mathbf{V}_k \mathbf{V}_k^H)^{-1} \Gamma_k, \end{aligned} \quad (30)$$

$$\begin{aligned} \mathbb{E} \left\{ \mathbf{H}_k^H \left( (\mathbf{Q}_k^{(d)})^{-1} - \mathbb{E} \left\{ (\mathbf{Q}_k^{(d)} + \mathbf{H}_k \mathbf{V}_k^{(d)} (\mathbf{V}_k^{(d)})^H \mathbf{H}_k^H)^{-1} \right\} \right) \mathbf{H}_k \right\} &= \\ \frac{\partial R_m}{\partial (\mathbf{V}_m \mathbf{V}_m^H)} &= \frac{\partial \bar{R}_k}{\partial (\mathbf{V}_m \mathbf{V}_m^H)} = \tilde{\eta}_k^{\text{pri}} (\mathbf{Q}_k^{-1}) - \tilde{\eta}_k^{\text{pri}} \left( (\mathbf{Q}_k + \tilde{\Gamma}_k)^{-1} \right). \end{aligned} \quad (31)$$

We then obtain

$$\begin{aligned} \mathbf{B}_k^{(d)} &= \hat{\mathbf{H}}_k^H (\mathbf{Q}_k^{(d)})^{-1} \hat{\mathbf{H}}_k + \tilde{\eta}_k \left[ (\mathbf{Q}_k^{(d)})^{-1} \right] - \\ & (\mathbf{I}_{N_r} + \Gamma_k \mathbf{V}_k \mathbf{V}_k^H)^{-1} \Gamma_k, \end{aligned} \quad (32)$$

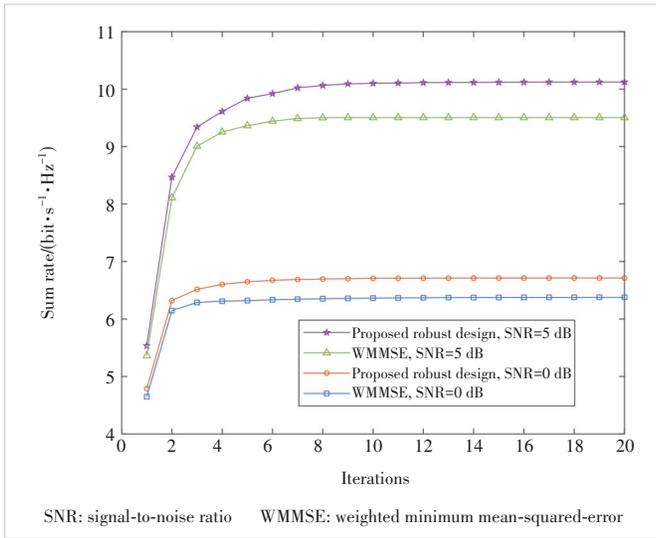
$$\mathbf{C}_k^{(d)} = \tilde{\eta}_k^{\text{pri}} (\mathbf{Q}_k^{-1}) - \tilde{\eta}_k^{\text{pri}} \left( (\mathbf{Q}_k + \tilde{\Gamma}_k)^{-1} \right). \quad (33)$$

By substituting Eqs. (15), (19), (32) and (33) into Eq. (18), we can derive the close-form of  $\mathbf{V}_k^{(d+1)}$  and the sum-rate maximization problem (9) is finally able to be solved by alternately optimizing  $\mathbf{V}_k$  until the convergence or the maximum iteration times are reached.

## 4 Simulation Results

In this section, our simulation results show the system performance of the proposed beamforming design under channel prediction errors. We consider a MU-MIMO system with the number of UE  $K$  set to 4 and each piece of UE is equipped with  $N_r = 4$  receive antennas. The transmit antenna number  $N_t = 32$  and the transmit power budget  $P$  is set to 1. The weight  $\alpha_k$  for each piece of UE is set equally. In addition,  $d_k = N_r$ ,  $\tau = 7$ ,  $N_{\text{slot}} = 10$ , and  $w = 10$ . The cluster delay line (CDL) channel model is used in the simulation to generate a time-varying channel for each piece of UE, where the implementations exactly follow the 3GPP 5G new radio standard protocol TR 38.901<sup>[25]</sup>. Specifically, we adopt an urban macro scenario and consider a CDL-A delay profile where the delay spread is set as 100 ns. The Doppler shift is computed by  $f_d = v f_c / c$ , which is a combination of user speed  $v$ , carrier frequency  $f_c$  and the speed of light  $c$ , with  $f_c = 2$  GHz. Different UE velocities are set to indicate different time-varying channels for performance comparisons. Several combinations of acquired CSI and beamforming methods are discussed here. The simplest case with no prediction and ZF algorithm is considered, where the CSI used for DL beamforming is the estimated CSI obtained in the previous SRS time slot. With the predicted CSI for DL time slots, ZF and WMMSE are considered for comparisons. Moreover, the case with perfect CSI known in the transmitter and WMMSE beamforming is also discussed as an ideal case.

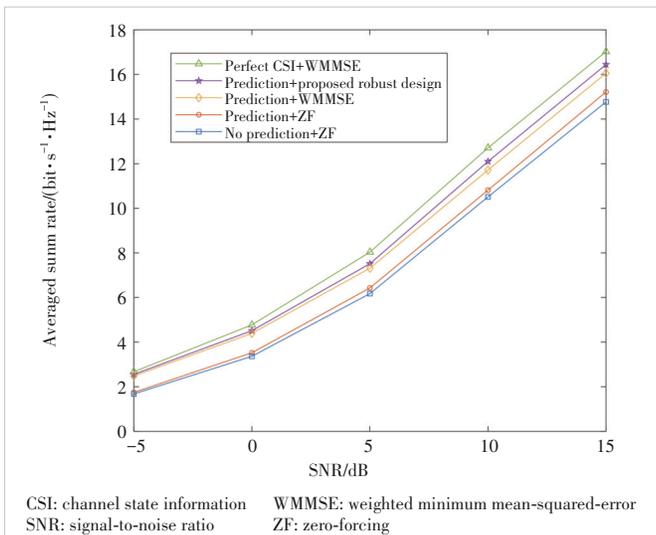
Fig. 4 illustrates the convergence behavior of the proposed beamforming design and WMMSE for the cases of SNR = 0 dB and SNR = 5 dB. Since the extra prediction errors are taken into consideration in the proposed design and then the



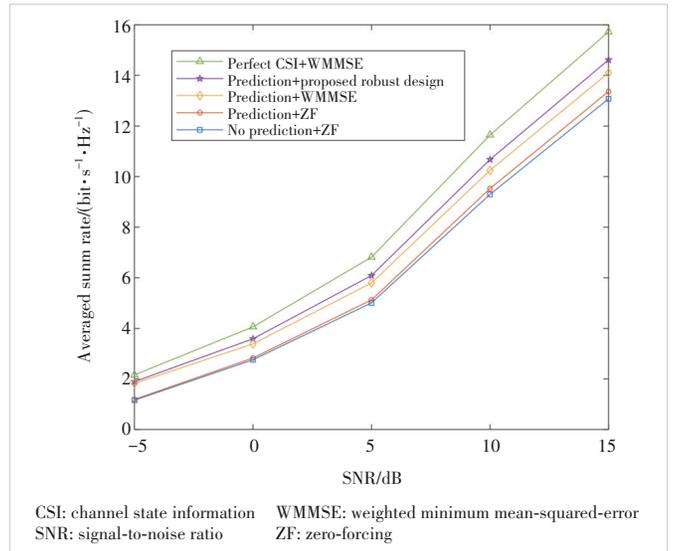
▲ Figure 4. Convergence behavior of the proposed beamforming design and WMMSE

randomness of the objective is to be solved, it is supposed to be more complex to conduct. As shown in this figure, the proposed one converges a little slower than WMMSE but both behave almost the same. And the proposed converges within 10 iterations, which confirms the practicality of our beamforming design. Moreover, it is observed that both algorithms take more iterations to converge as the SNR increases.

From Figs. 5 and 6, we can observe the averaged sum-rate performance of the system under different SNRs for different beamforming cases. Firstly, it is easy to notice the performance gain achieved by the channel prediction, which grows as the SNR increases. With the predicted CSI, WMMSE cannot attain equivalent gain compared with our proposed method that takes the error into consideration, due to the imperfection of CSI brought by the prediction errors which



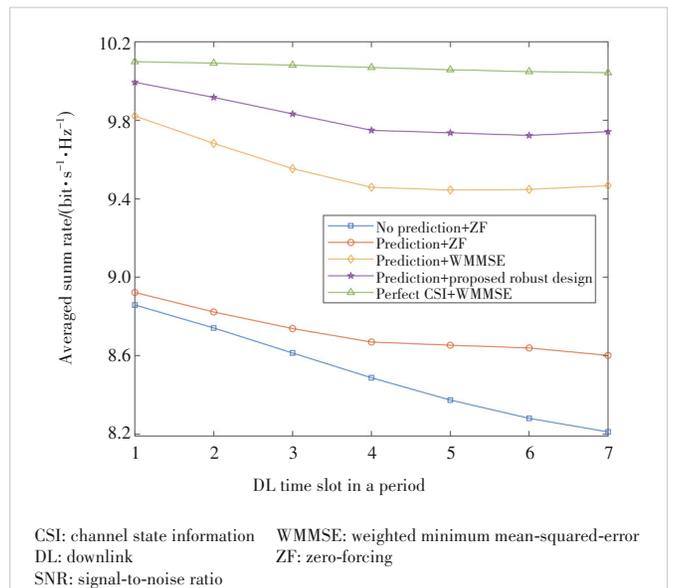
▲ Figure 5. Averaged sum rate versus SNR under different beamforming methods with  $v=60$  km/h



▲ Figure 6. Averaged sum rate versus SNR under different beamforming methods with  $v=120$  km/h

cause the performance limitation. The proposed beamforming design outperforms WMMSE more as the SNR goes higher. Comparing these two figures, we can find that the overall performance goes down as the UE velocity increases, with the performance gaps from the ideal perfect case to others becoming larger, but our robust method can maintain certain gains under such a severe time-varying effect. The proposed design can even achieve more gain against WMMSE with a higher velocity.

Fig. 7 shows the sum-rate performance comparisons during the DL time slot within the half-frame period. The statistics of each DL time slot plotted in the figure are obtained by averaging 100 periods. As time goes by, the outdated of acquired



▲ Figure 7. Sum rate versus DL time slots under different beamforming methods with  $v=60$  km/h

CSI is more serious, so the averaged sum rates decrease during the DL transmission period. However, the channel prediction can combat this fading, with more obvious gain in the last few time slots as shown in this figure. By considering the channel prediction errors and utilizing the statistical characteristics, the proposed beamforming can even achieve better performance which is close to the ideal case.

Fig. 8 presents the fall of the performance during the DL time slots within the half-frame period under different UE velocities. Considering the advantage of joint channel prediction and proposed robust beamforming design, we take the WMMSE method with no prediction as the benchmark. The averaged sum rates decrease more rapidly with the increasing of velocity. This is because the velocity is a sign of channel variation. Though the achievable gain of channel prediction gets smaller as the velocity increase, which can be observed by comparing Figs. 5 and 6, the joint robust design can implement more substantial performance gain as the time-varying effects become severer.

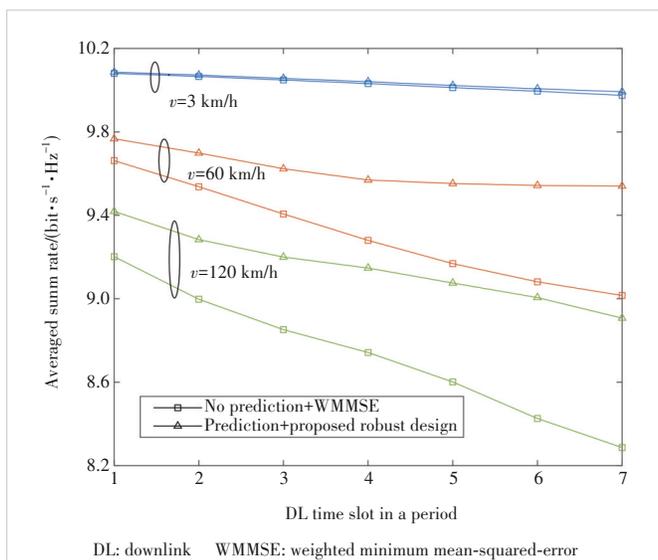
## 5 Conclusions

In this paper, we investigate the beamforming design under channel prediction errors in a time-varying MIMO system. By collecting the prediction errors and exploiting the statistical characteristics of the error samples, we propose a robust beamforming design to further combat the detriment caused by channel aging based on the channel prediction system. Due to the uncertainty brought by the error part, we exploit deterministic equivalents to obtain the closed-form expression to derive the robust beamforming matrix. Our simulation results show the effectiveness of the proposed beamforming design and reveal that the achievable gain even grows as UE velocities increase. The proposed design can

maintain the outperformance during the DL transmission time while the channels vary fast.

## References

- [1] NADEEM Q U A, KAMMOUN A, ALOUINI M S. Elevation beamforming with full dimension MIMO architectures in 5G systems: a tutorial [J]. IEEE communications surveys & tutorials, 2019, 21(4): 3238 - 3273. DOI: 10.1109/COMST.2019.2930621
- [2] LARSSON E G, EDFORS O, TUFVESSON F, et al. Massive MIMO for next generation wireless systems [J]. IEEE communications magazine, 2014, 52(2): 186 - 195. DOI: 10.1109/MCOM.2014.6736761
- [3] CASTAÑEDA E, SILVA A, GAMEIRO A, et al. An overview on resource allocation techniques for multi-user MIMO systems [J]. IEEE communications surveys & tutorials, 2017, 19(1): 239 - 284. DOI: 10.1109/COMST.2016.2618870
- [4] CHRISTENSEN S S, AGARWAL R, DE CARVALHO E, et al. Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design [J]. IEEE transactions on wireless communications, 2008, 7(12): 4792 - 4799. DOI: 10.1109/T-WC.2008.070851
- [5] SHI Q J, RAZAVIYAYN M, LUO Z Q, et al. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel [J]. IEEE transactions on signal processing, 2011, 59(9): 4331 - 4340. DOI: 10.1109/TSP.2011.2147784
- [6] WU Y P, JIN S, GAO X Q, et al. Transmit designs for the MIMO broadcast channel with statistical CSI [J]. IEEE transactions on signal processing, 2014, 62(17): 4451 - 4466. DOI: 10.1109/TSP.2014.2336637
- [7] WU D Y, ZHANG H T. Tractable modelling and robust coordinated beamforming design with partially accurate CSI [J]. IEEE wireless communications letters, 2021, 10(11): 2384 - 2387. DOI: 10.1109/LWC.2021.3101037
- [8] ZHENG J K, ZHANG J Y, BJÖRNSON E, et al. Impact of channel aging on cell-free massive MIMO over spatially correlated channels [J]. IEEE transactions on wireless communications, 2021, 20(10): 6451 - 6466. DOI: 10.1109/TWC.2021.3074421
- [9] LIU L H, FENG H, YANG T, et al. MIMO-OFDM wireless channel prediction by exploiting spatial-temporal correlation [J]. IEEE transactions on wireless communications, 2014, 13(1): 310 - 319. DOI: 10.1109/TWC.2013.112613.130455
- [10] YIN H F, WANG H Q, LIU Y Z, et al. Addressing the curse of mobility in massive MIMO with prony-based angular-delay domain channel predictions [J]. IEEE journal on selected areas in communications, 2020, 38(12): 2903 - 2917. DOI: 10.1109/JSAC.2020.3005473
- [11] WU C, YI X P, ZHU Y M, et al. Channel prediction in high-mobility massive MIMO: from spatio-temporal autoregression to deep learning [J]. IEEE journal on selected areas in communications, 2021, 39(7): 1915 - 1930. DOI: 10.1109/JSAC.2021.3078503
- [12] THEODORIDIS S. Machine learning: a Bayesian and optimization perspective [M]. Orlando, USA: Academic press, 2015
- [13] YUAN J D, NGO H Q, MATTHAIIOU M. Machine learning-based channel prediction in massive MIMO with channel aging [J]. IEEE transactions on wireless communications, 2020, 19(5): 2960 - 2973. DOI: 10.1109/TWC.2020.2969627
- [14] XU K, SHEN Z X, WANG Y R, et al. Location-aided mMIMO channel tracking and hybrid beamforming for high-speed railway communications: an angle-domain approach [J]. IEEE systems journal, 2020, 14(1): 93 - 104. DOI: 10.1109/JSYST.2019.2911296
- [15] XIA X C, XU K, ZHAO S B, et al. Learning the time-varying massive MIMO channels: robust estimation and data-aided prediction [J]. IEEE transactions on vehicular technology, 2020, 69(8): 8080 - 8096. DOI: 10.1109/TVT.2020.2968637
- [16] ENGEL Y, MANNOR S, MEIR R. The kernel recursive least-squares algorithm [J]. IEEE transactions on signal processing, 2004, 52(8): 2275 - 2285. DOI: 10.1109/TSP.2004.830985



▲ Figure 8. Sum rate versus DL time slots under different beamforming methods with  $v=60$  km/h

- [17] LIU W F, PARK I, PRINCIPE J C. An information theoretic approach of designing sparse kernel adaptive filters [J]. IEEE transactions on neural networks, 2009, 20(12): 1950 - 1961. DOI: 10.1109/TNN.2009.2033676
- [18] MIN C, CHANG N, CHA J, et al. MIMO-OFDM downlink channel prediction for IEEE802.16e systems using Kalman filter [C]//IEEE Wireless Communications and Networking Conference. IEEE, 2007: 942 - 946. DOI: 10.1109/WCNC.2007.179
- [19] RAMYA T R, BHASHYAM S. On using channel prediction in adaptive beamforming systems [C]//2nd International Conference on Communication Systems Software and Middleware. IEEE, 2007: 1 - 6. DOI: 10.1109/COMSWA.2007.382451
- [20] MARTOS-NAYA E, PARIS J F, FERNANDEZ-PLAZAOLA U, et al. Exact BER analysis for M-QAM modulation with transmit beamforming under channel prediction errors [J]. IEEE transactions on wireless communications, 2008, 7(10): 3674 - 3678. DOI: 10.1109/T-WC.2008.070192
- [21] PARIS J F, MARTOS-NAYA E, FERNANDEZ-PLAZAOLA U, et al. Analysis of adaptive MIMO transmit beamforming under channel prediction errors based on incomplete lipschitz - hankel integrals [J]. IEEE transactions on vehicular technology, 2009, 58(6): 2815 - 2824. DOI: 10.1109/TVT.2008.2011990
- [22] SUN Y, BABU P, PALOMAR D P. Majorization-minimization algorithms in signal processing, communications, and machine learning [J]. IEEE transactions on signal processing, 2017, 65(3): 794 - 816. DOI: 10.1109/TSP.2016.2601299
- [23] LU A N, GAO X Q, XIAO C S. Free deterministic equivalents for the analysis of MIMO multiple access channel [J]. IEEE transactions on information theory, 2016, 62(8): 4604 - 4629. DOI: 10.1109/TIT.2016.2573309
- [24] TULINO A M, VERDÚ S. Random matrix theory and wireless communications [M]. Norwell, USA: Now Publishers Inc, 2004. DOI: 10.1561/9781933019505
- [25] 3GPP. Study on channel model for frequencies from 0.5 to 100 GHz: TR 38.901 [S]. 2019

### Biographies

**ZHU Yuting** received her bachelor's degree in communication engineering from Beijing University of Posts and Telecommunications (BUPT), China in 2022. She is currently working toward a master's degree in communication and information engineering at the School of Artificial Intelligence, BUPT. Her research interests include the emerging technologies of 5G wireless communication networks.

**LI Zeng** (li.zeng@zte.com.cn) received his master's degree in information and communication engineering from Harbin Institute of Technology, China in 2016. He is currently working at the Algorithm Department, Wireless Product R&D Institute, Wireless Product Operation Division, ZTE Corporation. His research interests include 5G wireless communications and signal processing.

**ZHANG Hongtao** received his PhD degree in communication and information systems from Beijing University of Posts and Telecommunications (BUPT), China in 2008. He is currently a full professor with BUPT. He has authored or co-authored more than 100 articles on international journals and conferences, and has filed more than 50 patents. He is also the author of ten technical books. His research interests include 5G wireless communications and signal processing.