Performance Optimization for Interference Alignment Based on MMSE Prediction to Time-Variant Channels

Momiao Zhou and Feifei Zhao
School of Telecommunications Engineering, Xidian University, Xi'an 710071, China
Email: {mmzhou_work, ffzhao_work}@hotmail.com

Abstract —Interference Alignment (IA) is a precoding technique that achieves the maximum multiplexing gain over an interference channel when perfect Channel State Information (CSI) is available at transmitters. Most of IA researches assume channels remain static for a period but vary independently from block to block, which neglects the temporal correlation of time-variant channels. In this paper, we propose a novel scheme that transmitters utilize a number of samples to predict CSI instead of obtaining CSI through feedback all the time. By making full use of the correlation of time-variant channels, our proposed scheme is able to reduce overhead and compensate for the feedback error due to low feedback Signal-Noise Ratio (SNR). Furthermore, we find an optimized prediction horizon achieving the maximum sum rate of our system, which is the best tradeoff between prediction error and overhead length. Simulation results verify that our scheme outperforms the traditional non-predictive feedback scheme.

Index Terms—Precoding technique, feedback, correlation, time-variant channels, overhead, prediction

I. Introduction

Interference Alignment (IA) is a precoding strategy that interference can be aligned to a small subspace at each receiver so that the number of interference-free dimensions remaining for the desired signal can be maximized. IA can achieve the maximum degrees of freedom (DoF) in a K-user interference channel [1]. However, this result is based on the assumption that perfect Channel State Information (CSI) is available at all nodes, which is impossible in practical systems. The performance of IA with imperfect CSI over Multiple-Input-Multiple-Output (MIMO) interference channel was investigated in [2], [3]. Reference [2] provided an approximate closed-form signal-to-interference-plus-noise-ratio (SINR) expression for IA with imperfect CSI. The upper and lower bounds of sum mutual information were derived for the case of imperfect channel knowledge in [3]. In practical systems, a number of training sequences and feedback bits are needed for terminals to learn the global CSI [4]. The channel training and feedback strategy seriously affects the accuracy of CSI. Quantized feedback was first considered in [5] and it has been shown that the full DoF may be preserved only if the number of feedback bits scales fast enough with SNR [5], [6], which results in a considerable large-sized codebook [7]. To solve this problem, analog feedback strategy was proposed in [8]. Instead of quantizing the CSI, analog feedback directly transmits the channel coefficients as uncoded symbols. Reference [8] proved that IA's multiplexing gain is preserved as long as the forward and reverse link SNRs scale together. All these feedback strategies are based on the assumption that channels remain static for a long period, which neglects the variance of channels caused by some environmental factors such as users' mobility. The average CSI and precoder feedback bit rates were derived over time-variant MIMO interference channel in [9]. The robustness of IA in time-variant channels was assessed through a joint optimization of the pilot overhead and the IA update interval in [10]. However, the derivations only accommodate forward and reverse links with asymmetric power levels.

The main contribution of this paper is that we propose a new IA scheme that transmitters utilize a number of samples to predict CSI instead of obtaining CSI through feedback all the time. First-order Gauss-Markov channel model is employed to characterized the change of channels. By making full use of the correlation of time-variant channels, our proposed scheme is able to reduce overhead and compensate for the feedback error, especially when reverse link power is at a lower level than forward link power. We quantify the error of the prediction algorithm and optimize the average sum-rate of the system. Furthermore, we demonstrate with simulation results that the performance of our proposed scheme could be affected by many factors such as Doppler shift and the ratio of forward and reverse link power. If users move slowly or the reverse link is of poor quality, our proposed IA scheme achieves an outstanding performance.

Throughout this paper, we use the following notations: $\mathbf{A}$ denotes a matrix; $(\cdot)^*$ denotes the conjugate transpose; $|a|$ is the absolute value of a; $\mathbf{I}_N$ is the $N \times N$ identity matrix; $\mathcal{CN}(a, \mathbf{A})$ denotes a complex Gaussian random vector with mean and covariance matrix $\mathbf{A}$; δ_{ik} is the unit impulse function; $E[\cdot]$ denotes expectation.

II. System Model

A standard downlink of a cellular system is considered where Orthogonal Frequency Division Multiple-Access (OFDMA) is applied, such as TD-LTE [11]. Without loss of generality, we consider only one frequency band, with K base-stations, each serving one user, over a frequency-

Manuscript received November 6, 2015; revised May 11, 2016.
Corresponding author email: mmzhou_work@hotmail.com.
doi:10.12720/jcm.11.5.446-454

flat MIMO channel. Base-stations and users are all equipped with M antennas. Each base-station delivers d independent data streams to the target user. Fig. 1 depicts the scenario. To ensure the feasibility of interference alignment, we let $d \le 2M/(K+1)$ [12]. The received signal at user j is composed of interference from other cells and its intended signal, which can be written as

$$\mathbf{Y}_j = \sum_{i=1}^{K} \sqrt{\frac{P_f}{d}} \mathbf{H}_{j,i} \mathbf{V}_i \mathbf{s}_i + \mathbf{Z}_j \tag{1}$$

where $\mathbf{H}_{j,i} \in \mathbb{C}^{M \times M}$ represents the narrowband channel matrix from base-station i to user j. $\mathbf{H}_{j,i}$ is assumed to be independent across users and with i.i.d $\mathcal{CN}(0,1)$. $\mathbf{V}_i = [\mathbf{v}_{i,1},...,\mathbf{v}_{i,d}] \in \mathbb{C}^{M \times d}$, represents the precoding matrix of base-station i. $\mathbf{s}_i \in \mathbb{C}^{d \times 1}$ is the data symbol vector of base-station i satisfying $E(\mathbf{s}_i \mathbf{s}_i^*) = \mathrm{I}_d$, and is the circularly symmetric additive white Gaussian noise with the distribution of $\mathbf{Z}_j \in \mathbb{C}^{M \times 1}$. At user j, the received signal $\mathbf{Y}_j$ is filtered by a unitary interference suppression matrix $\mathbf{U}_j = [\mathbf{u}_{j,1},...,\mathbf{u}_{j,d}] \in \mathbb{C}^{M \times d}$. P_f is the forward link power.

For the network defined above, interference alignment is feasible if there exists a set of precoders $\mathbf{V}_i$ and decoders $\mathbf{U}_j$ satisfying

$$\mathbf{U}_j^* \mathbf{H}_{j,i} \mathbf{V}_i = \mathbf{0}, \forall\ j \ne i \tag{2}$$

$$rank(\mathbf{U}_j^* \mathbf{H}_{j,j} \mathbf{V}_j) = d, \forall\ j \tag{3}$$

where (2) ensures interference from other cells being aligned in a reduced subspace, while (3) guarantees the required dimensionality for the desired signal space.

By considering i.i.d Gaussian input signaling, the average sum-rate can be written as

$$R_{sum} = \sum_{j=1}^{K} \sum_{l=1}^{d} \log_2 \left(1 + \frac{\frac{P_f}{d} \left| \mathbf{u}_{j,l}^* \mathbf{H}_{j,j} \mathbf{v}_{j,l} \right|^2}{J_{j,l} + \sigma^2} \right) \tag{4}$$

where the leakage interference is treated as noise, and

$$J_{j,l} = \sum_{\substack{k=1 \\ k \ne j}}^{K} \sum_{m=1}^{d} \frac{P_f}{d} \left| \mathbf{u}_{j,l}^* \mathbf{H}_{j,k} \mathbf{v}_{k,m} \right|^2 + \sum_{\substack{m=1 \\ m \ne l}}^{d} \frac{P_f}{d} \left| \mathbf{u}_{j,l}^* \mathbf{H}_{j,j} \mathbf{v}_{j,m} \right|^2 \tag{5}$$

If perfect CSI is available, condition (2) is satisfied, that is $J_{j,1} = 0$. The effective direct channels $\mathbf{u}_{j,l}^* \mathbf{H}_{j,j} \mathbf{v}_{j,l}$ are complex Gaussian with unit variance since $\mathbf{v}_{j,l}$ and $\mathbf{u}_{j,l}$ are unitary and independent of $\mathbf{H}_{j,j}$. Therefore, (4) can be expressed as [13]

$$R_{sum} = \sum_{j=1}^{K} \sum_{l=1}^{d} \log_2 \left(1 + \rho \left| \mathbf{u}_{j,l}^* \mathbf{H}_{j,j} \mathbf{v}_{j,l} \right|^2 \right) \tag{6}$$

$$= Kd \log_2 (\mathrm{e}) e^{1/\rho} E_1 \left(\frac{1}{\rho} \right) \tag{7}$$

where $\rho = P_f / (d\sigma^2)$ represents the average per-stream SNR. (7) is the exponential integral form of (6), where $E_1(\eta) = \int_1^\infty t^{-1} e^{-\eta t} dt$.

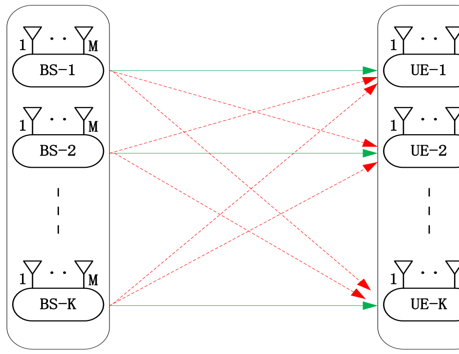

Fig. 1. A system model with K base-stations, each base-station serving one user

To study the influence of channel variation and measure the performance of our prediction scheme, a time-varying channel model is needed. In this paper, we use the first-order Gauss-Markov channel model, which has widely been used as a model for Rayleigh-fading time-varying channels [14], [15]. For simplicity, we make two assumptions as follows.

Assumption 1: Channel variation is time-stationary, which means that the values of channel correlation coefficients depend on relative time only.

Assumption 2: Channel correlation coefficients are entirely identical for different antenna pairs.

The channel gain between base-station i and user j at time k is given by

$$\mathbf{H}_{j,i}(k) = \mathbf{R}[k] \mathbf{H}_{j,i}(0) + \sqrt{1 - \mathbf{R}[k]^2} \boldsymbol{\varphi}_{j,i} \tag{8}$$

where $\mathbf{R}[k] = \mathrm{E}\{\mathbf{H}(n)\mathbf{H}(n+k)^*\}$ is the channel correlation coefficient during a time period of k. $\mathbf{H}_{j,i}(0)$ represents the channel matrix at time block 0, and $\boldsymbol{\varphi}_{j,i} \sim \mathrm{CN}(0,\mathbf{I})$ is the innovation between time block 0 and n.

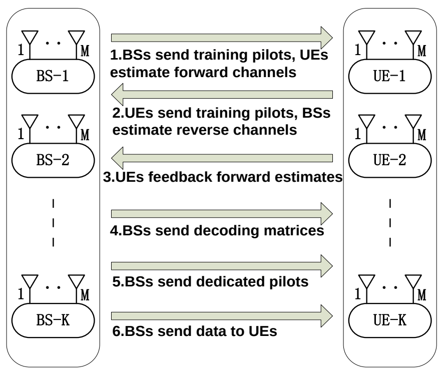

Fig. 2(a) CSI transfer of conventional IA scheme

III. Pilot Overhead Model

We split the transmission procedure into two phases. For the first P time blocks, the system operates the conventional training and feedback scheme [16], which is depicted in Fig. 2 (a); for the next Q time blocks, reverse channel training and feedback are omitted owing to our channel prediction algorithm based on autoregressive model [17], which is depicted in Fig. 2 (b).

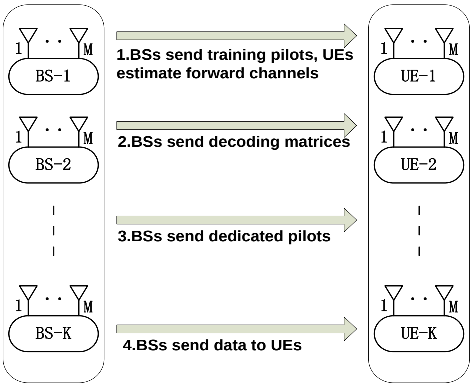

Fig. 2(b) CSI transfer of IA scheme with channel prediction algorithm

A. *First P Blocks*

Fig. 3 shows the pilot overhead model of conventional training and feedback. We assume that over each time block, fading remains constant and N_b symbols can be delivered. IA solution is recomputed every time block. Among the N_b symbols, N_{ft}, N_{rt}, N_{fb}, N_{su} and N_{sv} are respectively reserved for forward training, reverse training, CSI feedback, decoding matrices delivery and dedicated pilots delivery; the rest is for payload data. It's obvious that if the overhead occupies too many symbols, time for payload data is little left.

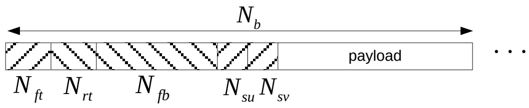

Fig. 3. Pilot overhead model of conventional IA scheme

1) *Forward and reverse channel training*

In the first place, each base-station i broadcasts a $M \times N_{ft}$ orthogonal pilot sequence matrix $\mathbf{\Phi}_i$, such that $\mathbf{\Phi}_i\mathbf{\Phi}_k^* = \delta_{ik}\mathbf{I}_M$. To guarantee the pilot orthogonality, we let $N_{ft} = KM$. The observation at user j is

$$\mathbf{Y}_j = \sqrt{\frac{N_{ft}P_f}{M}}\sum_{i=1}^{K}\mathbf{H}_{j,i}\mathbf{\Phi}_i + \mathbf{Z}_j, \forall j \tag{9}$$

where $\mathbf{Z}_j \in \mathbb{C}^{M\times N_{ft}}$ is a matrix of noise terms. User j then calculates the minimum-mean-square error (MMSE) estimate of the forward channel $\mathbf{H}_{j,i}$ by

$$\mathbf{H}_{j,i} = \frac{\sqrt{\frac{N_{ft}P_f}{M}}}{\sigma^2 + \frac{N_{ft}P_f}{M}}\mathbf{Y}_j\mathbf{\Phi}_i^*, \forall j \tag{10}$$

with entries of

$$\mathbf{H}_{j,i} \sim \mathcal{CN}(0, \frac{N_{ft}P_f/M}{\sigma^2 + N_{ft}P_f/M}) \tag{11}$$

with corresponding errors

$$\mathbf{H}_{j,i} \sim \mathcal{CN}(0, \frac{\sigma^2}{\sigma^2 + N_{ft}P_f/M}) \tag{12}$$

It's similar to perform reverse channel training. Each user broadcasts an orthogonal pilot sequence matrix over $N_{rt} = KM$ symbols. Base-stations then respectively calculate MMSE estimate with entries of

$$\mathbf{H}_{j,i}^r \sim \mathcal{CN}(0, \frac{N_{rt}P_r/M}{\sigma^2 + N_{rt}P_r/M}) \tag{13}$$

while the error matrix

$$\mathbf{H}_{j,i}^r \sim \mathcal{CN}(0, \frac{\sigma^2}{\sigma^2 + N_{rt}P_r/M}) \tag{14}$$

where P_r represents the reverse link power.

2) *Analog CSI feedback*

After forward and reverse channel training, each user j feedbacks its forward channel estimates to all base-stations over $N_{fb} = K^2M$ symbols. The base-stations then recompute MMSE estimates of the forward channels by utilizing their received feedback matrices. For simplicity, we make the same assumption as [8]: at the end of the feedback step, the base-stations cooperate by sharing their rows of the received feedback matrix, which enables them to form a unified estimate of the forward channels $\mathbf{H}_{j,i}$. Borrowing the derivations in [13], the final expression for the variance of the channel feedback error is

$$\sigma_{\mathbf{H}}^2 = \frac{M}{\rho d N_{ft}} + \frac{1}{\rho d(K-1)M}\left(\frac{M^2}{\beta N_{rt}} + \frac{KM^2}{\beta N_{fb}}\right) \tag{15}$$

$$= \frac{1}{Kd\rho} + \frac{2}{Kd\rho\beta(K-1)} \tag{16}$$

where $\beta = P_r/P_f$, is the ratio of reverse link power and forward link power.

After channel estimation, base-stations calculate precoding and decoding matrices following principle (2) and (3). To achieve interference alignment, user j must have the knowledge of the decoding matrix $\mathbf{U}_j$ and its precoded channel $\mathbf{H}_{j,j}\mathbf{V}_j$.

3) *Decoding matrices delivery*

To separate the desired signal and the interference, decoding matrix is required at users. Decoding matrices are delivered by analog signal based on non-codebook. Each base-station i broadcasts its decoding matrix $\mathbf{U}_i$ by multiplying an orthogonal pilot sequence matrix $\mathbf{\Psi}_i \in \mathbb{C}^{d\times N_{su}}$ over $N_{su} = Kd$ symbols. The observed signal at user j is given by

$$\mathbf{Y}_j = \sqrt{\frac{N_{su}P_f}{d}}\sum_{i=1}^{K}\mathbf{H}_{j,i}\mathbf{U}_i\mathbf{\Psi}_i + \mathbf{Z}_j \tag{17}$$

User j calculates a MMSE estimate of $\mathbf{U}_j$, that is

$$\mathbf{U}_j = \frac{\sqrt{\frac{N_{su}P_f}{d}}}{\sigma^2 + \frac{N_{su}P_f}{d}}\mathbf{H}^{*}_{j,j}\mathbf{Y}_j\mathbf{\Psi}^{*}_j \tag{18}$$

where the error matrix $\mathbf{U}_{j,j} = \mathbf{U}_{j,j} - \mathbf{U}_{j,j}$ is complex Gaussian with variance

$$\sigma^2_{\mathbf{U}_j} = 1 - \frac{N_{su}P_f/d}{\sigma^2 + N_{su}P_f/d} \bullet \frac{N_{ft}P_f/M}{\sigma^2 + N_{ft}P_f/M} \tag{19}$$

$$= 1 - \left(\frac{Kd\rho}{1+Kd\rho}\right)^2 \tag{20}$$

4) Dedicated pilots delivery

To enable each user j to learn its precoded channel, each base-station i broadcasts its precoding matrix $\mathbf{V}_i$ by multiplying an orthogonal pilot sequence matrix $\mathbf{\Omega}_i \in \mathbb{C}^{d\times N_{sv}}$ over $N_{sv} = Kd$ symbols. The observed signal at user j is given by

$$\mathbf{Y}_j = \sqrt{\frac{N_{sv}P_f}{d}}\sum_{i=1}^{K}\mathbf{H}_{j,i}\mathbf{V}_i\mathbf{\Omega}_i + \mathbf{Z}_j \tag{21}$$

The MMSE estimate of $\mathbf{A}_{j,j} = \mathbf{H}_{j,j}\mathbf{V}_j$ is then

$$\mathbf{A}_{j,j} = \frac{\sqrt{\frac{N_{sv}P_f}{d}}}{\sigma^2 + \frac{N_{sv}P_f}{d}}\mathbf{Y}_j\mathbf{\Omega}^{*}_j, \tag{22}$$

where the error matrix $\mathbf{A}_{j,j} = \mathbf{A}_{j,j} - \mathbf{A}_{j,j}$ is complex Gaussian with variance

$$\sigma^2_{\mathbf{A}} = \frac{\sigma^2}{\sigma^2 + N_{sv}P_f/d} = \frac{1}{1+Kd\rho} \tag{23}$$

B. Next Q Blocks

Fig. 4 shows the pilot overhead model of our proposed IA scheme based on channel prediction. In this phase, reverse channel training and feedback are omitted owing to our channel prediction based on previous P CSI samples obtained in phase A. Therefore, Among the N_b symbols, only N_{ft}, N_{su} and N_{sv} are required respectively for forward training, decoding matrices delivery and dedicated pilots delivery; the rest of the symbols are for payload data. These steps are identical with those in phase A. It's clear that the introduction of channel prediction greatly reduces the length of overhead. Furthermore, since conventional channel feedback is replaced by channel prediction, the channel feedback error in phase A is of course replaced by channel prediction error. When the SNR of reverse link is relatively low, the channel feedback error is much larger than the channel prediction error, which can be seen from our simulation in section V. Next we introduce the channel prediction algorithm.

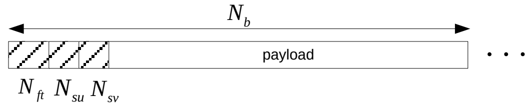

Fig. 4. Pilot overhead model of IA scheme with channel prediction algorithm

The number of samples predictor saved is flagged as P, denoting the *prediction order*; While Q, representing the number of blocks the channel is predicted ahead, denotes the *prediction horizon* [18]. The model is depicted by Fig. 5. We employ the channel prediction algorithm based on autoregressive model. Autoregressive model can represent a type of time-varying random process, and it specifies that the output signal depends linearly on its own previous values [19], [20]. It can be expressed as

$$x(n+1) = \sum_{m=1}^{p} a_m x(n-m+1) + \varepsilon(n+1) \tag{24}$$

where $a_1,\ldots,a_p$ are the parameters of the model, and $\varepsilon(n+1)$ is a white noise process with zero mean and constant variance σ^2_ε. When σ^2_ε is small, we can obtain the p-order autoregressive-model prediction as

$$x(n+1) = \sum_{m=1}^{p} a_m x(n-m+1) \tag{25}$$

where $x(n+1)$ is the predicted value of $x(n+1)$ based on the autoregressive model with the previous values $x(n),\ldots,x(n-p-1)$. Thus the main task of the autoregressive-model prediction is to calculate of the parameters of the autoregressive model.

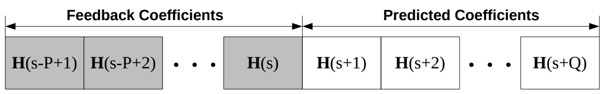

Fig. 5. Channel prediction algorithm model

Since we have made the assumption that all base-stations cooperate with each other, the predictor can be placed at any base-station. We use $\overline{\mathbf{H}}^{s+n}_{j,i}$ to denote the predictive channel coefficient and $\mathbf{H}^{s+n}_{j,i}$ the actual channel coefficient at time $s+n$ between base-station i and user j. The computation of predictive channel coefficient is given by

$$\overline{\mathbf{H}}^{s+n}_{j,i} = \sum_{m=0}^{P-1} W_m\left[\mathbf{H}^{s-m}_{j,i} + \mathbf{E}^{s-m}_{j,i}\right], n \in [1,Q] \tag{26}$$

where $\mathbf{H}^{s-m}_{j,i} + \mathbf{E}^{s-m}_{j,i}$ is the estimated channel coefficient and $\mathbf{E}^{s-m}_{j,i}$ is the channel estimation error at time $s-m$. We aim at designing the prediction weight vector $\mathbf{W}_n = [W_0, W_1, \ldots, W_{P-1}]$ to minimize the mean square error between predictive and actual channel coefficients.

Therefore, the objective function is given by

$$\min\left\{E\left[(\overline{\mathbf{H}}_{j,i}^{s+n}-\mathbf{H}_{j,i}^{s+n})(\overline{\mathbf{H}}_{j,i}^{s+n}-\mathbf{H}_{j,i}^{s+n})^*\right]\right\} \quad (27)$$

Bringing (26) into this function and according to the orthogonality principle, we can get the point that the optimum $\mathbf{W}_n$ must satisfy (for $m=0,\ldots,P-1$)

$$E\left\{(\overline{\mathbf{H}}_{j,i}^{s+n}-\mathbf{H}_{j,i}^{s+n})(\mathbf{H}_{j,i}^{s-m}+\mathbf{E}_{j,i}^{s-m})^*\right\}=0 \quad (28)$$

with (26) and (28), this yields the Wiener-Hopf equations

$$\sum_{m'=0}^{P-1} W_{m'}(\mathbf{R}[m-m']+\sigma_{\mathbf{H}}^2\delta[m-m']\mathbf{I})=\mathbf{R}[n+m] \quad (29)$$

where $\mathbf{R}[m]=\mathrm{E}\{\mathbf{H}_{j,i}^{s}\bullet(\mathbf{H}_{j,i}^{s-m})^*\}$ is the correlation matrix of interval m and we assume that this information is available to all base-stations . $\sigma_{\mathbf{H}}^2=E\{\mathbf{E}_{j,i}^{s}\bullet(\mathbf{E}_{j,i}^{s})^*\}$, is the error variance of the channel samples, i.e., the variance of feedback error calculated in phase A, step 2). (29) can also be written in matrix form as [21]

$$\mathbf{W}_n(\overline{\mathbf{R}}+\sigma_{\mathbf{H}}^2\mathbf{I})=\mathbf{\Pi} \quad (30)$$

where

$$\mathbf{\Pi}=\left[\mathbf{R}[n],\mathbf{R}[n+1],\ldots,\mathbf{R}[n+P-1]\right],$$

$$\overline{\mathbf{R}}=\begin{bmatrix}\mathbf{R}[0] & \mathbf{R}[1] & \cdots & \mathbf{R}[P-1]\\ \mathbf{R}[-1] & \mathbf{R}[0] & \cdots & \mathbf{R}[P-2]\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{R}[1-P] & \mathbf{R}[2-P] & \cdots & \mathbf{R}[0]\end{bmatrix}$$

Thus, the MMSE-optimum predictor coefficient matrix $\mathbf{W}_n$ is given by

$$\mathbf{W}_{\mathrm{opt}}=\mathbf{\Pi}(\overline{\mathbf{R}}+\sigma_{\mathrm{H}}^2\mathbf{I})^{-1} \quad (31)$$

The mean square error obtained with the optimum coefficient $\mathbf{W}_{\mathrm{opt}}$ can be shown to be (for $n=1,\ldots,Q$)

$$\begin{aligned}\sigma_{\mathbf{B}}^2[n]&=E\left\{(\overline{\mathbf{H}}_{j,i}^{s+n}-\mathbf{H}_{j,i}^{s+n})(\overline{\mathbf{H}}_{j,i}^{s+n}-\mathbf{H}_{j,i}^{s+n})^*\right\}\\&=1-\mathbf{\Pi}(\overline{\mathbf{R}}+\sigma_{\mathbf{H}}^2\mathbf{I})^{-1}\mathbf{\Pi}^*\end{aligned} \quad (32)$$

This depends on the channel correlation $\mathbf{R}[m]$ and the channel estimation error variance $\sigma_{\mathbf{H}}^2$. Although (32) is difficult to interpret, based on the conclusions derived by [21], it can be proved that if $\sigma_{\mathbf{H}}^2$ is fixed, with the increase of n, the correlation between predicted channels and channel samples decays, and $\sigma_{\mathbf{B}}^2[n]$ increases as a result.

IV. Optimum Prediction Horizon

According to (4) and (5), the average sum-rate of each block can be expressed as

$$C_{sum}=\sum_{j=1}^{K}\sum_{l=1}^{d}E\left[\log_2\left(1+\rho_{eff}\right)\right]$$

where

$$\begin{aligned}\rho_{eff}&=\frac{\frac{P_f}{d}\left|\mathbf{u}_{j,1}^*\mathbf{H}_{j,j}\mathbf{v}_{j,1}\right|^2}{\sigma^2+\sum_{k=1}^{K}\sum_{m=1}^{d}E\left[\frac{P_f}{d}\left|(\mathbf{u}_{j,1}^*-\overline{\mathbf{u}}_{j,1}^*)(\mathbf{H}_{j,k}+\overline{\mathbf{H}}_{j,k})\mathbf{v}_{k,m}\right|^2\right]}\\&=\frac{\frac{P_f}{d}\left|\overline{\mathbf{u}}_{j,1}^*\mathbf{H}_{j,j}\mathbf{v}_{j,1}\right|^2}{\sigma^2+\sum_{k=1}^{K}\sum_{m=1}^{d}E\left[\frac{P_f}{d}\left|(\mathbf{u}_{j,1}^*\mathbf{H}_{j,k}-\overline{\mathbf{u}}_{j,1}^*\mathbf{H}_{j,k})\mathbf{v}_{k,m}\right|^2\right]}\end{aligned}$$

For phase A, since we do channel estimation and feedback for each block, the effective average per-stream SINR keeps constant, and it is given by [10]

$$\rho_P=\frac{\rho\sigma_{\mathbf{A}}^2\sigma_{\mathbf{U}}^2}{1+\rho d[(K-1)(\sigma_{\mathbf{H}}^2\sigma_{\mathbf{U}}^2+\sigma_{\bar{\mathbf{U}}}^2\sigma_{\mathbf{H}}^2)+\sigma_{\bar{\mathbf{A}}}^2\sigma_{\mathbf{U}}^2+\sigma_{\bar{\mathbf{U}}}^2\sigma_{\mathbf{A}}^2]}$$

For phase B and $n\in[1,Q]$, with the increase of n, the prediction accuracy drops off. The effective average per-stream SINR is shown to be

$$\rho_Q[n]=\frac{\rho\sigma_{\mathbf{A}}^2\sigma_{\mathbf{U}}^2}{1+\rho d[(K-1)(\sigma_{\mathbf{B}}^2[n]\sigma_{\mathbf{U}}^2+\sigma_{\bar{\mathbf{U}}}^2\sigma_{\mathbf{B}}^2[n])+\sigma_{\bar{\mathbf{A}}}^2\sigma_{\mathbf{U}}^2+\sigma_{\bar{\mathbf{U}}}^2\sigma_{\mathbf{A}}^2]}$$

Then the effective sum-rate of phase A is

$$\overline{R}_P=\frac{N_b-N_P}{N_b}\bullet Kd\log_2(e)\mathrm{e}^{1/\rho_P}E_1\left(\frac{1}{\rho_P}\right)$$

And the effective sum-rate of phase B is

$$\overline{R}_Q[Q]=\begin{cases}0, & Q=0\\ \frac{N_b-N_Q}{N_b}\bullet\frac{\sum_{n=1}^{Q}Kd\log_2(e)\mathrm{e}^{1/\rho_Q[n]}E_1\left(\frac{1}{\rho_Q[n]}\right)}{Q}, & Q>0\end{cases}$$

where $N_P=N_{ft}+N_{rt}+N_{fb}+N_{su}+N_{sv}$ is the total overhead of IA with conventional channel estimation and feedback; while $N_Q=N_{ft}+N_{su}+N_{sv}$ is the total overhead of IA with channel prediction. Furthermore, the effective sum-rate of the whole process is

$$R_{sum}[Q]=\frac{P}{P+Q}\overline{R}_P+\frac{Q}{P+Q}\overline{R}_Q[Q]$$

Let P be a certain value that never varies, in Appendix A, we prove that under some suitable conditions, there must exist a value of Q that maximizes $R_{sum}[Q]$.

Having quantified IA average sum-rate as a function of prediction horizon Q, we redefine the optimization problem as (33), shown at the bottom of this page.

In fact, it's hard to get the closed-form solution for the function above. Since this is a unconstrained nonlinear integer optimization problem, one-dimension search method is suitable to solve it. To achieve maximum search efficiency, D.S.C interpolation is adopted [22]. Specific algorithm is as follows.

Algorithm: D.S.C Interpolation

Initialization:

Define Q_0 as the initial value, and step size $\Delta Q = 1$.

Iteration:

1) Let $k = 0$

  if $R_{sum}[Q_0 + \Delta Q] \geq R_{sum}[Q_0]$

    turn to 2);

  else

  let $\Delta Q = -\Delta Q$ and turn to 2);

  end if

2) Let $Q_{k+1} = Q_k + \Delta Q$ and calculate $R_{sum}[Q_{k+1}]$;

3) If $R_{sum}[Q_{k+1}] \geq R_{sum}[Q_k]$

    let $\Delta Q = 2\Delta Q$, $k = k+1$ and turn to 2);

  else

    let $\Delta Q = \Delta Q/2$, $Q_m = Q_{k+1}$, $Q_{m+1} = Q_m - \Delta Q$,

    $Q_{m-1} = Q_k$, $Q_{m-2} = Q_{k-1}$;

  end if

4) For set $\{Q_{m-2}, Q_{m-1}, Q_{m+1}, Q_m\}$

  if $R_{sum}[Q_{m+1}] < R_{sum}[Q_{m-1}]$

    remove Q_m and let $Q_b = Q_{m-1}$;

  else

    remove Q_{m-2} and let $Q_b = Q_{m+1}$;

  end if

5) Let $Q_a = Q_b - \Delta Q$, $Q_c = Q_b + \Delta Q$, and apply quadratic

  interpolation($\lfloor \bullet \rfloor$ denotes rounding down)

$$\overline{Q} = \left\lfloor Q_b + \frac{(R_{sum}[Q_a] - R_{sum}[Q_c]) \bullet \Delta Q}{2(R_{sum}[Q_a] - 2R_{sum}[Q_b] + R_{sum}[Q_c])} + 0.5 \right\rfloor;$$

6) If $R_{sum}[\overline{Q}] \geq \max\left\{R_{sum}[\overline{Q}-1], R_{sum}[\overline{Q}+1]\right\}$

  stop iteration and let $Q_{opt} = \overline{Q}$;

  else if $R_{sum}[\overline{Q}] < R_{sum}[\overline{Q}-1]$ && $R_{sum}[\overline{Q}] \geq R_{sum}[\overline{Q}+1]$

    make $\overline{Q}$ as the new initial state, and let $Q_0 = \overline{Q}$, $\Delta\overline{Q} = -1$,

    then turn to 1);

  else

    make $\overline{Q}$ as the new initial state, and let $Q_0 = \overline{Q}$, $\Delta\overline{Q} = 1$,

    then turn to 1);

  end if

Next section we will make performance analysis to the IA scheme with channel prediction by matlab simulation.

V. Simulation Results

Consider an IA cluster with three base-stations, each serving only one user. All the base-stations and users are equipped with two antennas and all channel coefficients are i.i.d zero mean unit variance circularly symmetric complex Gaussian. Suppose each base-station conveys one spatial stream, i.e. $d = 1$. The OFDM symbol period $T_s = 66.7\mu s$ is chosen according to the 3GPP LTE standard [23]. "Time block" mentioned above denotes $10ms$. We make the assumption that during each time block, channels remain static. According to [23], 7 symbols can be delivered during one time slot ($0.5ms$). Hence, one time block consists of 140 symbols. Channels vary based on first-order Gauss-Markov channel model from block to block, which is given by (8). We invoke the standard Clarke-Jakes correlation function $\rho[k] = J_0(2\pi f_m k T_s)$, where $J_0(\cdot)$ is the 0-th order bessel function of the first kind and f_m denotes the normalized Doppler frequency. $f_m = v/\lambda$, where v denotes the velocity of the users. In this paper we let the carrier frequency be 2.4GHz, with the corresponding wavelength $\lambda = 0.125m$.

We illustrate the mean square error of the channel coefficients based on non-predictive scheme as well as predictive scheme in Fig. 6 respectively for $f_m = 15Hz$, $f_m = 30Hz$ and $f_m = 45Hz$. The non-predictive scheme means that the latest estimated CSI is used in all the upcoming time blocks without updating. Without loss of generality, prediction order $P = 5$ is adopted. As time goes on, the mean square error of the two schemes both increase. Nevertheless, the mean square error of the predictive scheme is below that of the non-predictive scheme all along, which fully embodies the accuracy of our prediction algorithm.

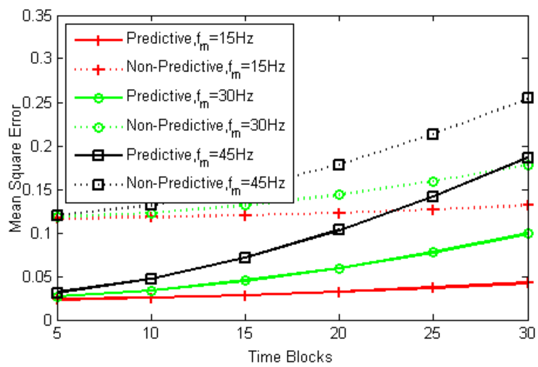

Fig. 6. Mean Square Error v. Time blocks with $SNR = 15dB$, $\beta = 0.1$

$$Q_{opt} = \arg\max_{Q \in N^+} \left(\frac{P}{P+Q} \frac{N_b - N_P}{N_b} \bullet Kd \log_2(e) e^{1/\rho_P} E_1\left(\frac{1}{\rho_P}\right) + \frac{Q}{P+Q} \frac{N_b - N_Q}{N_b} \bullet \frac{\sum_{n=1}^{Q} Kd \log_2(e) e^{1/\rho_Q[n]} E_1\left(\frac{1}{\rho_Q[n]}\right)}{Q} \right) \quad (33)$$

Fig. 7 plots the optimum prediction horizon versus prediction order respectively for $f_m = 20Hz$ and $f_m = 60Hz$. As anticipated, with the growth of prediction order, the optimum prediction horizon is also on a rising trend.

Respectively for $SNR = P_f/\sigma^2 = 10dB$ and $SNR = P_f/\sigma^2 = 15dB$, we illustrate the optimum prediction horizon versus Doppler shift in Fig. 8. Without loss of generality, prediction order $P = 5$ is adopted. As anticipated, with the increase of Doppler shift, the optimum prediction horizon decreases.

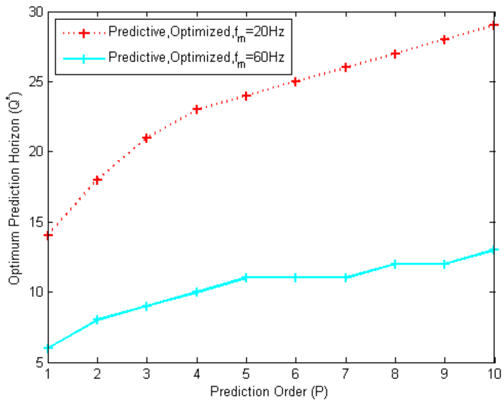

Fig. 7. Optimum Prediction Horizon v. Prediction Order with $SNR = 15dB$, $\beta = 0.1$

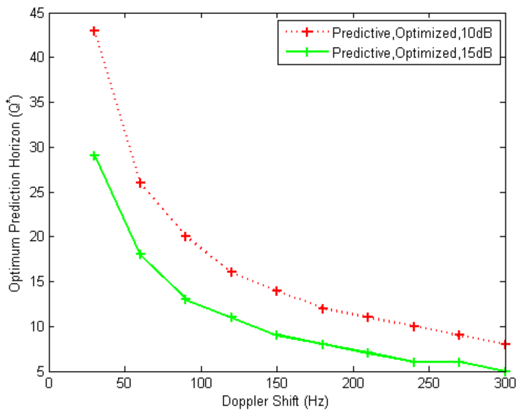

Fig. 8. Optimum Prediction Horizon v. Doppler Shift with $\beta = 0.1$

Respectively for $SNR = 10\text{dB}$ and $SNR = 15\text{dB}$, we illustrate the effective sum-rate versus the ratio of forward and reverse link power in Fig. 9. We compare three different schemes: “NP-NO”, “NP-O” and “P-O”. The “NP-NO” scheme is the conventional IA scheme that channel estimation and feedback is required at each time block with neither prediction nor optimization. While the “NP-O” scheme is an IA scheme of which the optimization objective is the interval to do channel estimation and feedback that achieves the maximum sum-rate, but channel prediction is unconsidered [10]. The “P-O” scheme is our proposed scheme in this paper. Without loss of generality, prediction order $P = 5$ is adopted. As observed, if the reverse link power decays, the average sum-rate of the three schemes all decrease. The reason is that when SNR of the reverse link becomes lower, the channel estimation accuracy degrades. Meanwhile, it’s clear that the downtrend of the “P-O” scheme is slower than the other two schemes. This is because the “P-O” scheme partially omits the reverse channel estimation and feedback phase, which relieves its sensitivity to the reverse link quality. When $1/\beta$ is quite low, the “NP-O” scheme may be a good choice, but when $1/\beta$ is high enough, the “P-O” scheme is sure to perform better.

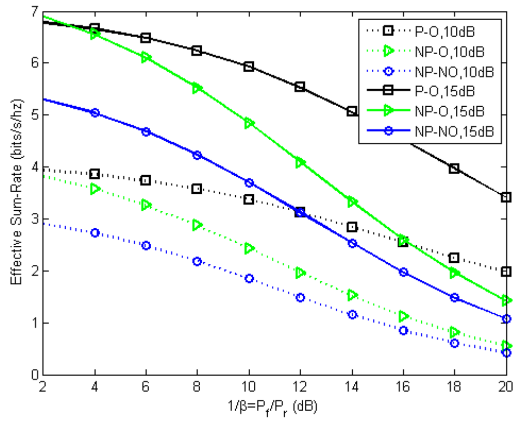

Fig. 9. Effective Sum-Rate v. $1/\beta$ with $f_m = 20\text{Hz}$

Fig. 10 compares the effective sum-rate of the three schemes referred above versus Doppler shift. Without loss of generality, prediction order $P = 5$ is adopted. As observed, the effective sum-rate of the “NP-NO” scheme remains constant no matter how Doppler shift varies. While the other two schemes perform worse with the increase of Doppler shift, which is consistent with our anticipation. Fig. 10 also shows that the “P-O” scheme is more sensitive to Doppler shift. When Doppler shift is relatively small, the “P-O” scheme achieves a much better performance than the other two schemes.

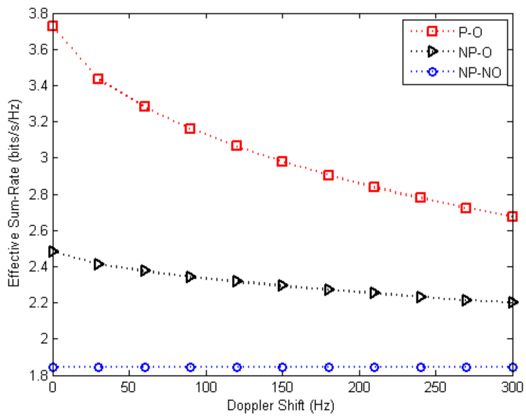

Fig. 10. Effective Sum-Rate v. Doppler Shift with $SNR = 10\text{dB}$, $\beta = 0.1$

VI. Conclusion

It’s known that perfect channel state information is needed to achieve the maximum gain of IA. We analyzed how imperfect CSI impacts on IA performance caused by additive white Gaussian noise. Then a classic time-varying channel model was introduced. We showed that if channel is time-varying, channel estimation and feedback would become more frequent, which results in a huge rise of overhead. We proposed a new IA scheme comprising two phases. Phase A performs conventional channel estimation and analog feedback, while phase B executes the channel prediction algorithm. We quantified the error of the prediction algorithm as well as the average sum-rate of the system. We confirmed one-dimension search method could be used to find the optimum prediction horizon in order to achieve the maximum average sum-rate. Finally, simulation results showed that the performance of our proposed scheme could be affected by many factors such as Doppler shift and the ratio of forward and reverse link power. If users move at a low speed or the reverse link is of poor quality, our proposed IA scheme achieves a much better performance. Our derived analysis can also be of great significance to the further work on interference alignment.

APPENDIX A

Assumption: $\begin{cases} \rho > \sqrt{2}+1 \\ (1+\rho_P)^{N_b-N_P} > (1+\rho_Q[+\infty])^{N_b-N_Q} \end{cases}$

Proposition: In the definition domain, there must exist a block Q that is a local maximum point of $R_{sum}[Q]$, and also it is the global optimal solution.

Proof: The assumption $\rho > \sqrt{2}+1$ indicates that $\sigma_U^2 < 0.5$, which can be used to prove that $\rho_Q[n]$ is monotone decreasing with n .Therefore, $\overline{R}_Q[Q]$ is monotone decreasing with Q. Suppose $\exists N > 0$ such that $R_{sum}[N+1] - R_{sum}[N] < 0$.

Define $\alpha = (N_b - N_Q)/N_b$. Then

$$\begin{aligned} & R_{sum}[N+2] - R_{sum}[N+1] \\ &= \frac{\alpha\left[(P+N+1)\overline{R}_Q[N+2] - \sum_{n=0}^{N+1}\overline{R}_Q[n]\right] - P\overline{R}_P}{(P+N+2)(P+N+1)} \\ &< \frac{\alpha\left[(P+N+1)\overline{R}_Q[N+1] - \sum_{n=0}^{N+1}\overline{R}_Q[n]\right] - P\overline{R}_P}{(P+N+2)(P+N+1)} \\ &= \frac{P+N}{P+N+2}\bullet(R_{sum}[N+1] - R_{sum}[N]) \\ &< 0. \end{aligned}$$

Therefore, we come to the conclusion that once there exist a block N for which $R_{sum}[N] < R_{sum}[N-1]$, then for the following blocks $Q > N$, $R_{sum}[Q]$ is monotone decreasing with Q.

While since $(1+\rho_P)^{N_b-N_P} > (1+\rho_Q[+\infty])^{N_b-N_Q}$, then $\overline{R}_P > \overline{R}_Q[+\infty]$, meaning that $R_{sum}[0] > R_{sum}[+\infty]$. Therefore, in the definition domain, $R_{sum}[Q]$ is monotone decreasing or first increasing and then decreasing. Summing up the above, there must exist a block Q that is a local maximum point of $R_{sum}[Q]$, and also it is the global optimal solution. The proposition is fully attained.

REFERENCES

[1] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425-3441, 2008.

[2] B. Nosrat-Makouei, J. G. Andrews, and R. W. Heath, "MIMO interference alignment over correlated channels with imperfect CSI," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2783-2794, 2011.

[3] T. Roland and M. Guillaud, "Cellular interference alignment with imperfect channel knowledge," in *Proc. IEEE International Conference on Communications Workshops*, 2009.

[4] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005, ch. 4.

[5] B. Helmut and J. Thukral, "Interference alignment with limited feedback," in *Proc. IEEE International Symposium on Information Theory*, 2009.

[6] R. T. Krishnamachari and M. K. Varanasi, "Interference alignment under limited feedback for MIMO interference channels," *IEEE Transactions on Signal Processing*, vol. 61, no. 15, pp. 3908-3917, 2013.

[7] C. Giuseppe, N. Jindal, and S. Shamai, "On the required accuracy of transmitter channel state information in multiple antenna broadcast channels," in *Proc. Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, 2007.

[8] A. O. El and R. W. Heath, "Interference alignment with analog channel state feedback," *IEEE Transactions on Wireless Communications*, vol. 11, no. 2, pp. 626-636, 2012.

[9] Z. Yang, *et al*., "How much feedback is required for interference alignment over K-user time-variant channel?" in *Proc. International Conference on Communications in China*, 2013.

[10] R. K. Mungara, G. George, and A. Lozano, "Overhead and spectral efficiency of pilot-assisted interference alignment in time-selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4884-4895, 2014.

[11] 3GPP, "Technical Specification Group Radio Access Network; Requirements for Further Advancements for E-UTRA (LTE-Advanced)," 3GPP, Tech. Rep. 36.913 V8.0.0, June 2008.

[12] L. Ruan, V. K. N. Lau, and M. Z. Win, "The feasibility conditions for interference alignment in MIMO networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2066-2077, 2013.

[13] O. El Ayach, A. Lozano, and R. W. Heath, "On the overhead of interference alignment: Training, feedback, and cooperation," *IEEE Transactions on Wireless Communications*, vol. 11, no. 11, pp. 4192-4203, 2012.

[14] L. Junse, H. Yu, and Y. Sung, "Beam tracking for interference alignment in time-varying mimo interference channels: A conjugate-gradient-based approach," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 958-964, 2014.

[15] H. Yu, *et al*., "Beam tracking for interference alignment in slowly fading MIMO interference channels: A perturbations approach under a linear framework," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1910-1926, 2012.

[16] R. K. Mungara, G. George, and A. Lozano, “Overhead and spectral efficiency of pilot-assisted interference alignment in time-selective fading channels,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4884-4895, 2014.

[17] N. Zhao, *et al*, “Interference alignment with delayed channel state information and dynamic AR-Model channel prediction in wireless networks,” *Wireless Networks*, vol. 21, no. 4, pp. 1227-1242, 2015.

[18] A. Duel-Hallen, “Fading channel prediction for mobile radio adaptive transmission systems,” *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2299-2313, 2007.

[19] A. S. Khrwat, “Channel prediction for limited feedback precoded MIMO-OFDM systems,” in *Proc. 19th International Conference on Telecommu-nications*, 2012.

[20] A. Heidari, A. K. Khandani, and D. McAvoy, “Adaptive modelling and long-range prediction of mobile fading channels,” *IET Communications*, vol. 4, no. 1, pp. 39-50, 2010.

[21] D. Schafhuber and G. Matz, “MMSE and adaptive prediction of time-varying channels for OFDM systems,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 593-602, 2005.

[22] R. K. Sundaram, *A First Course in Optimization Theory*, Cambridge University Press, 1996, ch. 2.

[23] 3GPP. 3GPP TS 36.211 version 12.6.0 Release 12, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical Channels and Modulation. 2015.

Momiao Zhou received the BE degree in Communication Engineering from Xidian University, China. He is currently working towards the phD degree at the School of Telecommunications Engineering in Xidian University. His research interests include cooperative communication, relay network and MIMO systems. Now he is working on the topic of interference management in Ultra Dense Network (UDN).

Feifei Zhao received her BE degree in Electronic Information Science and Technology from West Anhui University, China in 2008, and she is currently studying towards her PhD degree at the School of Telecommunications Engineering in Xidian University. Her research interests include wireless resource allocation, interference management and D2D communications.