Multicarrier-Division Duplex for Solving the Channel Aging Problem in Massive MIMO Systems

Bohan Li, Lie-Liang Yang, *Fellow, IEEE*, Robert G. Maunder, *Senior Member, IEEE*, Songlin Sun, *Senior Member, IEEE*, Pei Xiao, *Senior Member, IEEE*

Abstract—The separation of training and data transmission as well as the frequent uplink/downlink (UL/DL) switching make time-division duplex (TDD)-based massive multiple-input multiple-output (mMIMO) systems less competent in fast time-varying scenarios due to the resultant severe channel aging. To this end, a multicarrier-division duplex (MDD) mMIMO scheme associated with two types of well-designed frame structures are introduced for combating channel aging when communicating over fast time-varying channels. For comparison, the corresponding TDD frame structures related to the 3rd Generation Partnership Project (3GPP) standards and their variant forms are presented. The MDD-specific general Wiener predictor and decision-directed Wiener predictor are introduced to predict the channel state information, respectively, in the time domain based on UL pilots and in the frequency domain based on the detected UL data, considering the impact of residual self-interference (SI). Moreover, by applying the zero-forcing precoding and maximum ratio combining, the closed-form approximations for the lower bounded rate achieved by TDD and MDD frame structures over time-varying channels are derived. Our main conclusion from this study is that the MDD, endowed with the capability of full-duplex but less demand on SI cancellation than in-band full-duplex (IBFD), outperforms both the conventional TDD and IBFD in combating channel aging.

Index Terms—Massive multiple-input multiple-output, channel aging, time-varying, multicarrier-division duplex, time-division duplex, full-duplex, Wiener predictor, performance analysis.

I. INTRODUCTION

Massive multiple-input multiple-output (mMIMO) has been regarded as one of the most promising technologies for significantly improving the spectral efficiency (SE) in the fifth-generation (5G) wireless systems [1]. However, due to the large number of antennas equipped at base station (BS), the acquisition of channel state information (CSI) becomes highly challenging. As the result, time-division duplex (TDD) is deemed to be the only feasible mode in practical mMIMO systems to date, which exploits the channel reciprocity to acquire downlink (DL) CSI from uplink (UL) pilots [2]. As its

counterpart, in frequency-division duplex (FDD) systems, the overhead for training/feedback is proportional to the size of the antenna arrays at BS, resulting in less resources for sending desired data in one coherence block [3]. On the other hand, with the growing demand on the mobile services operated in ground vehicles, high-speed trains, unmanned aerial vehicles (UAVs), etc., the performance of TDD systems is no longer dominated by pilot contamination but the channel aging problem. In other words, due to fast time-varying, the channel at the time when it is estimated via UL training may be very different from that at the time when it is applied for DL preprocessing and UL detection. Moreover, the UL/DL switching intervals required by TDD systems exacerbate the situation. This CSI mismatch may lead to significant degradation of the SE of both UL and DL.

In order to support the robust communications over fast time-varying channels, channel prediction based on the Wiener filter or Kalman filter has been studied. In [4], the authors considered the channel aging problem in mMIMO systems and applied the Wiener predictor (WP) to predict future CSI. The results show that using channel prediction can partially relieve the channel aging effect and render an increase of sum rate. The authors of [5] analyzed the effect of WP for time-varying channels on the system performance. However, in these two papers, the one-step WP relies on the latest channels estimated using the UL pilots that are always available in the front of the symbol to be predicted. Unfortunately, in practical TDD systems, UL pilots are unavailable during DL transmission and hence, the accuracy of channel prediction degrades with time, causing the performance degradation. For this sake, in [6], the authors considered a more practical scenario, where no UL pilots are sent during the Kalman filter assisted channel prediction for DL transmission. The results show that the Kalman predictor is capable of improving performance when compared with the case without prediction. However, due to the time-varying effect, the achievable DL rate reduces quickly with time as the result of the fact that the predicted CSI becomes less and less accurate. Recently in [7], a learning-based approach was introduced, showing that it can outperform the conventional non-linear Kalman predictor in dealing with the channel aging problem. Again, the prediction accuracy of CSI deteriorates as the interval between the channel being predicted and the pilot symbols becomes larger.

In contrast to the half-duplex (HD) mode of TDD, in-band full-duplex (IBFD) has the inherent advantage for solving the channel aging problem. This is because in IBFD systems, DL and UL occur concurrently, which enables to acquire the

latest CSI during data transmission without invoking DL/UL switching. However, to the best of our knowledge, there are no open references which have considered the IBFD-relied transmission design for the multicarrier mMIMO systems communicating over fast time-varying channels. One conceivable reason may be that the self-interference (SI) problem in IBFD-relied systems is still intractable. This becomes even more challenging in the high-mobility communication scenarios. Note that, although channel aging is not considered, in [8], the channel acquisition relying on a hybrid IBFD and TDD mode was demonstrated to be efficient. With this hybrid scheme, partial DL transmissions are activated during UL training, leading to an increased SE. Moreover, the effect of SI cancellation (SIC) on the performance of channel estimation was studied with the IBFD systems, showing that if SIC is insufficient, IBFD is unable to outperform the conventional TDD. Similar observations were also obtained in [9].

To overcome the weakness of both IBFD and TDD over fast time-varying communication channels, inspired by [10], the multicarrier division duplex (MDD) associated with the dedicated frame design is proposed so as to relieve the channel aging problem, which constitutes the main motivation of this paper. The rationale can be briefly explained as follows. On the one hand, in MDD systems, both DL and UL transmissions can occur concurrently within the same frequency band but on different subcarriers. Hence, when needed, UL pilots can be continuously (or frequently) transmitted during DL transmissions. Therefore, CSI can be updated in time and does not become outdated as in TDD systems. On the other hand, MDD is capable of circumventing the stringent requirement for SIC in IBFD systems. According to [11], in practical IBFD systems, in addition to the SIC in the propagation- and analog-domain, the digital-domain SIC has to cancel the main SI signal component by at least 30 dB. Achieving this is very power-consuming and technically demanding, especially for the relatively small-sized mobile terminals (MTs), e.g., UAVs and smartphones. By contrast, in the MDD-assisted systems, this amount of SIC can be attained at nearly no extra cost of system resources, owing to the embedded fast Fourier transform (FFT) operation [9], which allows to ideally separate UL signals from DL signals in the digital-domain.

In literature, there are several papers having studied MDD from the perspective of SI mitigation [9, 12, 13], flexible subcarrier allocation [13, 14], and the application in cell-free mMIMO systems [15]. However, the advanced frame design for MDD-assisted systems and MDD's potential for solving the channel aging problem have not been studied yet. To fill the research gap, in this paper, we address comprehensively the implementation issue of MDD in fast time-varying communication scenarios, so as to unleash the advantages of MDD over the conventional TDD and IBFD systems. Our main contributions can be summarized as follows:

- *Firstly*, a model for the MDD-based multiuser mMIMO systems communicating over time-varying channels is presented. To alleviate the channel aging problem, two types of frame structures dedicated for MDD are proposed, which enable UL pilots (or UL data transmissions) to occur concurrently with DL transmissions, so that CSI

can be promptly updated whenever needed. For comparison, the corresponding TDD frame structures related to the 3rd Generation Partnership Project (3GPP) standards [16, 17] as well as their modified forms for supporting different mobility scenarios are introduced.

- *Secondly*, to operate with the proposed frame structures, two finite impulse response WPs that consider both channel estimation and residual SI errors, namely the general WP and the decision-directed WP (DD-WP), are proposed for MDD frame structures. To be more specific, the general WP directly predicts the time-domain CSI based on the observations collected from the UL channel estimation. By contrast, the DD-WP leverages both the UL pilots and the detected UL data symbols to predict the frequency-domain CSI. Along with these WPs, the impact of residual SI and the order of WPs on the performance of MDD systems are studied and compared.

- *Thirdly*, the closed-form expressions for approximating the lower bounded average sum rates of both the MDD and TDD systems with our proposed frame structures are derived, when assuming that the zero-forcing (ZF) precoding and maximum ratio combining (MRC) are respectively used for DL transmission and UL detection.

Our studies and simulation results show that TDD systems suffer from the channel aging problem, whose performance degrades significantly with time, when channels vary fast. By contrast, the MDD systems endowed with the full duplex (FD) capability can effectively mitigate the channel aging problem and hence, are capable of significantly outperforming their TDD counterparts, when communicating over fast time-varying channels. Moreover, the studies demonstrate that the SIC in FD systems plays a paramount role in channel prediction. In the case of imperfect SIC, MDD becomes more competent than IBFD for operation in high-mobility communications scenarios.

The rest of the paper is organized as follows. In Section II, the model for the MDD-based system is described. In Section III, two general frame structures for TDD and MDD systems are introduced. Section IV presents the principles of channel estimation and two approaches for channel prediction. Section V analyzes the lower bounded average sum rates of MDD and TDD systems. Section VI provides the performance results to compare TDD, MDD and IBFD systems. Finally, conclusions are drawn in Section VII.

Throughout the paper, the following notations are used: $\boldsymbol{A}$, $\boldsymbol{a}$ and a stand for matrix, vector, and scalar, respectively; $\mathcal{A}$ and $|\mathcal{A}|$ represent the set and the cardinality of set, respectively; $(\boldsymbol{a})_i$ denotes the i-th element of $\boldsymbol{a}$; $\boldsymbol{A}^{(i,:)}$, $\boldsymbol{A}^{(:,j)}$ and $(\boldsymbol{A})_{i,j}$ denote the i-th row, the j-th column and the (i,j)-th element of $\boldsymbol{A}$, respectively; Furthermore, $|\boldsymbol{A}|$, $\boldsymbol{A}^*$, $\boldsymbol{A}^T$, $\boldsymbol{A}^{-1}$ and $\boldsymbol{A}^H$ represent, respectively, the determinant, complex conjugate, transpose, inverse and Hermitian transpose of $\boldsymbol{A}$; The Euclidean norm of a vector and the Frobenius norm of a matrix are denoted as $\|\cdot\|_2$ and $\|\cdot\|_F$, respectively; $\boldsymbol{I}_N$ denotes a $(N \times N)$ identity matrix; $\mathcal{CN}(\boldsymbol{0}, \boldsymbol{A})$ represents the zero-mean complex Gaussian distribution with covariance matrix $\boldsymbol{A}$; Furthermore, $\text{Tr}(\cdot)$, $\log(\cdot)$ and $\mathbb{E}[\cdot]$ denote the trace, logarithmic and expectation operators, respectively.

II. SYSTEM MODEL

Consider a single-cell mMIMO orthogonal frequency-division multiplexing (OFDM) system having a base station (BS) equipped with N antennas and D single-antenna MTs randomly distributed in the cell. The system is operated in the MDD mode, allowing DL and UL to communicate concurrently in the same frequency band but on different subcarriers. Following the concept of mMIMO, the number of antennas at BS is assumed to be much larger than the number of served MTs [3], i.e., $N \gg D$. Furthermore, based on the principles of MDD [9], subcarriers are divided into two mutually exclusive subsets, namely a DL subcarrier subset $\mathcal{M}$ with M subcarriers, and a UL subcarrier subset $\bar{\mathcal{M}}$ with $\bar{M}$ subcarriers, i.e., $|\mathcal{M}| = M$ and $|\bar{\mathcal{M}}| = \bar{M}$. The total number of subcarriers is expressed as $M_{\text{sum}} = M + \bar{M}$. Similar to the other works on resource-allocation, e.g., [18–20], the DL/UL subcarrier allocation (SA) is carried out at BS based on the CSI and the requirements of MTs' quality of services. The allocation results are informed to MTs through control channels during the initial access procedure. Below it is assumed that the DL/UL SA results have been obtained, and that each MT knows its assigned subcarriers. The readers interested in the SA in MDD systems are referred to [14] for details.

A. Channel Model

The channels are assumed to be frequency-selective in terms of the M_{sum} subcarriers but each subcarrier experiences flat fading. To consider the influence of channel aging, it is assumed that the channel coefficients do not change within one OFDM symbol, but vary from one symbol to the next. Therefore, the multipath fading channel can be modeled by a tapped-delay-line filter [21], with the L-tap channel impulse response (CIR) between the d-th MT and the n-th BS antenna over the i-th OFDM symbol duration expressed as

$$\boldsymbol{g}_{n,d}[i] = [g_{n,d}[i,1], ..., g_{n,d}[i,l], ..., g_{n,d}[i,L]]^T, \quad (1)$$

where $g_{n,d}[i,l] = \nu_{n,d}[i,l]\sqrt{\beta_d}$ and $\nu_{n,d}[i,l] \sim \mathcal{CN}(0, 1/L)$ is the small-scale fading, while β_d represents the large-scale fading, which only depends on the distance between MT d and BS and is assumed to remain constant over one communication frame. For any user-antenna pair, the channels of different taps are assumed to be independent. Hence we have $\boldsymbol{R}_g^d = \mathbb{E}\left[\boldsymbol{g}_{n,d}[i]\boldsymbol{g}_{n,d}^H[i]\right] = \frac{\beta_d}{L}\boldsymbol{I}_L$.

According to the principles of OFDM [10], the frequency-domain CSI $\boldsymbol{h}_{n,d}[i]$ over the i-th OFDM symbol duration can be obtained as

$$\boldsymbol{h}_{n,d}[i] = \boldsymbol{F}\boldsymbol{\Psi}\boldsymbol{g}_{n,d}[i], \quad (2)$$

where $\boldsymbol{F} \in \mathbb{C}^{M_{\text{sum}} \times M_{\text{sum}}}$ is the FFT matrix with $(\boldsymbol{F})_{p,q} = \frac{1}{\sqrt{M_{\text{sum}}}}e^{-j2\pi(p-1)(q-1)/M_{\text{sum}}}$, $\boldsymbol{\Psi} \in \mathbb{C}^{M_{\text{sum}} \times L}$ is constructed by the first L columns of $\boldsymbol{I}_{M_{\text{sum}}}$. Furthermore, the DL subchannels $\boldsymbol{h}_{n,d}^{\text{DL}}[i]$ and UL subchannels $\boldsymbol{h}_{n,d}^{\text{UL}}[i]$ between the n-th antenna

at BS and the d-th MT over the i-th OFDM symbol can be obtained from (2), which can be expressed as

$$\boldsymbol{h}_{n,d}^{\text{DL}}[i] = \boldsymbol{\Phi}_{\text{DL}}\boldsymbol{h}_{n,d}[i]$$
$$= [h_{n,d}[i,1], ..., h_{n,d}[i,m], ..., h_{n,d}[i,M]]^T, \quad (3)$$
$$\boldsymbol{h}_{n,d}^{\text{UL}}[i] = \boldsymbol{\Phi}_{\text{UL}}\boldsymbol{h}_{n,d}[i]$$
$$= [h_{n,d}[i,1], ..., h_{n,d}[i,\bar{m}], ..., h_{n,d}[i,\bar{M}]]^T, \quad (4)$$

where $\boldsymbol{\Phi}_{\text{DL}} = \boldsymbol{I}_{M_{\text{sum}}}^{(\mathcal{M},:)}$ and $\boldsymbol{\Phi}_{\text{UL}} = \boldsymbol{I}_{M_{\text{sum}}}^{(\bar{\mathcal{M}},:)}$ are the mapping matrices, constructed from $\boldsymbol{I}_{M_{\text{sum}}}$ by choosing its rows corresponding to the particular subcarriers assigned to DL and UL, respectively.

Since in the analog-domain the MDD systems are operated in the FD mode, there is SI at both BS and MTs. In particular, although the MDD-based systems employ the inherent advantage to suppress SI in the digital-domain by the FFT operation[1], the DL and UL subcarrier signals are coupled in the analog-domain before the ADC at the BS receiver. As in the traditional IBFD systems [23], without sufficient SIC in the propagation- and analog-domain prior to ADC, the high-power SI imposed by the DL subcarriers from the BS transmitter may overwhelm the desired UL signals from MTs, which leads to the inefficient operation of ADC and consequently significant quantization error in digital-domain. Due to the short separation between transmitter and receiver, the SI channels are assumed to be flat fading and identical to all subcarriers, and also quasi-static within one frame. Specifically, the SI channels between the transmitter and receiver at BS and MTs follow the distributions of $(\boldsymbol{H}_{\text{SI}})_{i,j} \sim \mathcal{CN}(0,1)$, where $\boldsymbol{H}_{\text{SI}} \in \mathbb{C}^{N \times N}$, and $h_{\text{SI}} \sim \mathcal{CN}(0,1)$, respectively.

B. Channel Aging

As stated in Section I, the relative mobility between BS and MTs leads to time-varying channels, causing that the channel varies between the time when it is estimated and the time when the estimated CSI is applied for transmissions. This is the well-known channel aging problem. To model the channel aging, the Jakes model is introduced, which has the normalized discrete-time autocorrelation function at the BS given by [7]

$$R[k] = \text{J}_0(2\pi f_D T_s |k|), \quad (5)$$

where $\text{J}_0(\cdot)$ is the zeroth-order Bessel function of the first kind, T_s is the OFDM symbol duration, f_D is the maximum Doppler frequency shift, and $|k|$ is the delay in terms of the number of symbols.

Furthermore, for the sake of analyzing MDD frame structures with WP and DD-WP in the following sections, we adopt the autoregressive model of order 1, denoted as AR(1), for approximating the temporal correlation between adjacent

---

[1]In IBFD-based systems, in order to mitigate the digital-domain SI, the receiver has to estimate the channel between the adigital-to-nalog (DAC) at transmitter and the analog-to-digital converter (ADC) at receiver so as to reconstruct the transmitted signal, which is then subtracted from the received signal [22]. Explicitly, this process imposes a heavy burden on system overhead, especially when a system with a large number of subcarriers is considered.

symbols. In this case, the channel coefficient $\boldsymbol{g}_{n,d}[i]$ of (1) can be formulated as [24]

$$\boldsymbol{g}_{n,d}[i] = \alpha\boldsymbol{g}_{n,d}[i-1] + \boldsymbol{v}_{n,d}[i], \tag{6}$$

where $\alpha = \mathrm{J}_0(2\pi f_D T_s)$ given by $R[1]$ in (5), which is assumed to be perfectly known at BS, $\boldsymbol{v}_{n,d}[i]$ is the uncorrelated complex white Gaussian noise vector with the distribution of $\mathcal{CN}\left(\boldsymbol{0}, (1-\alpha^2)\boldsymbol{R}_g^d\right)$. According to the channel model of (6), $\boldsymbol{g}_{n,d}[i]$ is a stationary Markov random process, and has the auto-correlation of $\mathbb{E}\left[\boldsymbol{g}_{n,d}[i-q]\boldsymbol{g}_{n,d}^H[i-p]\right] = \alpha^{|p-q|}\boldsymbol{R}_g^d$. Moreover, according to (3) and (4), the DL subchannels satisfying

$$h_{n,d}[i,m] = \alpha h_{n,d}[i-1,m] + \underbrace{\boldsymbol{\psi}_m \boldsymbol{v}_{n,d}[i]}_{\tilde{v}_{n,d}^m[i]} \tag{7}$$

can be obtained, where $\boldsymbol{\psi}_m = \boldsymbol{I}_{\mathrm{sum}}^{(m,:)}\boldsymbol{F}\boldsymbol{\Psi}$ and $\tilde{v}_{n,d}^m[i] \in \mathcal{CN}\left(0, (1-\alpha^2)R_h^d\right)$ with $R_h^d = \mathbb{E}\left[h_{n,d}[i,m]h_{n,d}^H[i,m]\right] = \beta_d/M_{\mathrm{sum}}$. Similarly, the UL subchannels $h_{n,d}[i,\bar{m}]$ can be generated.

Remark 1. This paper considers the time-domain correlation between different OFDM symbols, as shown in (6), and assumes that the multipath fading channel within each symbol duration is time-invariant. The reason for us to make these assumptions is as follows. According to the 5G NR standard [17], in an OFDM system having a carrier frequency of 5 GHz and a subcarrier spacing of 15 kHz, and supporting a relative velocity of 500 km/h, the maximum Doppler shift is $f_D = 2.315$ kHz. Hence, the channel's coherence time is about $1/(4f_D) = 108$ μs [25], which is much larger than one OFDM symbol duration (about 80 μs including cyclic prefix). With this regard and considering that our focus is on the channel prediction between adjacent OFDM symbols, the multipath fading channel within one OFDM symbol duration can be deemed as time-invariant and hence, the inter-carrier interference generated by time-varying can also be neglected.

C. Transmission Model

1) Downlink Transmission: For DL transmission, the signal $\boldsymbol{s}_{\mathrm{DL}}[i,m] \in \mathbb{C}^{N \times 1}$ transmitted by BS on the m-th subcarrier in the i-th OFDM symbol duration is

$$\boldsymbol{s}_{\mathrm{DL}}[i,m] = \sqrt{p_{\mathrm{DL}}}\boldsymbol{F}^{\mathrm{ZF}}[i,m]\boldsymbol{x}[i,m], \tag{8}$$

where p_{DL} is the power budget per DL subcarrier and the total BS transmit power is $P_{\mathrm{DL}} = p_{\mathrm{DL}}M$, $\boldsymbol{x}[i,m] = [x_1[i,m], ..., x_D[i,m]]^T$ contains the information symbols normalized to satisfy $\mathbb{E}\left[\boldsymbol{x}[i,m]\boldsymbol{x}[i,m]^H\right] = \boldsymbol{I}_D$. It is assumed that the ZF precoding is applied at transmitter, giving $\boldsymbol{F}^{\mathrm{ZF}}[i,m] = \boldsymbol{H}_{\mathrm{DL}}^H[i,m]\left(\boldsymbol{H}_{\mathrm{DL}}[i,m]\boldsymbol{H}_{\mathrm{DL}}^H[i,m]\right)^{-1}$, where $\boldsymbol{F}^{\mathrm{ZF}}[i,m] = [\boldsymbol{f}_1[i,m], ..., \boldsymbol{f}_D[i,m]]$ with the constraint of $\|\boldsymbol{f}_d[i,m]\|_2 = 1/\sqrt{D}$, such that $\mathbb{E}\left[\|\boldsymbol{s}_{\mathrm{DL}}[i,m]\|_2^2\right] = p_{\mathrm{DL}}$, and $\boldsymbol{H}_{\mathrm{DL}}[i,m] = [\boldsymbol{h}_1[i,m], ..., \boldsymbol{h}_d[i,m], ...\boldsymbol{h}_D[i,m]]^H$, where $\boldsymbol{h}_d[i,m] = [h_{1,d}[i,m], h_{2,d}[i,m], ..., h_{N,d}[i,m]]^T$. It is noteworthy that the analysis in this paper is based on the equal power allocation among MTs. Intuitively, an appropriate power allocation method, e.g., that in our previous work [14], will

increase the achievable sum rates, owing to both user and subcarrier diversity being exploited. The received signal at the d-th MT can be expressed as

$$y_d[i,m] = \sqrt{p_{\mathrm{DL}}}\boldsymbol{h}_d^H[i,m]\boldsymbol{f}_d[i,m]x_d[i,m]$$
$$+ \sqrt{p_{\mathrm{DL}}}\sum_{k=1, k\neq d}^{D}\boldsymbol{h}_d^H[i,m]\boldsymbol{f}_k[i,m]x_k[i,m] + z_d^{\mathrm{SI}}[i] + z_d[i,m], \tag{9}$$

where $z_d^{\mathrm{SI}}[i] \sim \mathcal{CN}\left(0, \xi_{\mathrm{MT}}p_{\mathrm{UL}}\bar{M}\right)$ denotes the residual SI[2] signal imposed by the UL transmission of MT d with p_{UL} being the power budget per UL subcarrier, such that the total MT transmit power is $P_{\mathrm{UL}} = p_{\mathrm{UL}}\bar{M}$. Note that ξ_{MT} in (9) and ξ_{BS} in (10) denote the SIC capability that is provided by the pathloss of SI channels and the existing SIC techniques, such as, the antenna circulator [11], spatial beamforming [12], dual-port polarized antenna [27], multi-tap RF canceller [28], etc., implemented in propagation- and analog-domain, as well as the FFT operation operated in digital-domain. In (9), the covariance of $z_d^{\mathrm{SI}}[i]$ is calculated in detail in Appendix A. In addition, $z_d[i,m] \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white Gaussian noise. Note also that, the interference caused by the other MTs' UL transmission on the DL reception is ignored. The reason is as follows. First, in analogy to the conventional orthogonal frequency-division multiplexing access (OFDMA) systems [18], the inter-MT interference (IMI) and the desired DL signal are located on the mutually orthogonal subcarriers. Secondly, the large-scale fading of MT-MT link and the MT's relatively lower transmit power allow the received signal of an MT to efficiently pass the ADC. Consequently, the IMI can be readily removed by the FFT operation in the digital-domain, when all MTs and the BS are assumed to be synchronized within an allowable time window [13].

2) Uplink Transmission: For the UL transmission, the signal received at BS over the $\bar{m}$-th UL subcarrier and i-th OFDM symbol duration can be expressed as

$$\boldsymbol{s}_{\mathrm{UL}}[i,\bar{m}] = \sum_{d=1}^{D}\boldsymbol{h}_d[i,\bar{m}]\sqrt{p_{\mathrm{UL}}}x_d[i,\bar{m}] + \boldsymbol{z}^{\mathrm{SI}}[i] + \boldsymbol{z}[i,\bar{m}], \tag{10}$$

where $\boldsymbol{h}_d[i,\bar{m}] = [h_{1,d}[i,\bar{m}], h_{2,d}[i,\bar{m}], ..., h_{N,d}[i,\bar{m}]]^T$, $\boldsymbol{z}[i,\bar{m}] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_N)$. $\boldsymbol{z}^{\mathrm{SI}}[i]$ is the SI generated by BS transmissions, which is modeled as $\boldsymbol{z}^{\mathrm{SI}}[i] \sim \mathcal{CN}\left(\boldsymbol{0}, \xi_{\mathrm{BS}}p_{\mathrm{DL}}M\boldsymbol{I}_N\right)$. The covariance of $\boldsymbol{z}^{\mathrm{SI}}[i]$ is derived in detail in Appendix A. In this paper, it is assumed that the MRC is used for signal detection. Then we have $\boldsymbol{W}^{\mathrm{MRC}}[i,\bar{m}] = \boldsymbol{H}_{\mathrm{UL}}[i,\bar{m}]$, where $\boldsymbol{H}_{\mathrm{UL}}[i,\bar{m}] = [\boldsymbol{h}_1[i,\bar{m}], ..., \boldsymbol{h}_D[i,\bar{m}]]$. Let us write $\boldsymbol{W}^{\mathrm{MRC}}[i,\bar{m}] = [\boldsymbol{w}_1[i,\bar{m}], ..., \boldsymbol{w}_D[i,\bar{m}]]$. Then the decision variable for MT d can be expressed as

$$y_d[i,\bar{m}] = \sqrt{p_{\mathrm{UL}}}\boldsymbol{w}_d^H[i,\bar{m}]\boldsymbol{h}_d[i,\bar{m}]x_d[i,\bar{m}]$$
$$+ \sqrt{p_{\mathrm{UL}}}\sum_{k=1, k\neq d}^{D}\boldsymbol{w}_d^H[i,\bar{m}]\boldsymbol{h}_k[i,\bar{m}]x_k[i,\bar{m}] \tag{11}$$
$$+ \boldsymbol{w}_d^H[i,\bar{m}]\boldsymbol{z}^{\mathrm{SI}}[i] + \boldsymbol{w}_d^H[i,m]\boldsymbol{z}[i,\bar{m}],$$

---

[2]The residual SI consists of the combined effect of the additive noise introduced by automatic gain control (AGC), non-linearity of ADC and the phase noise generated by oscillator due to RF imperfection [26].
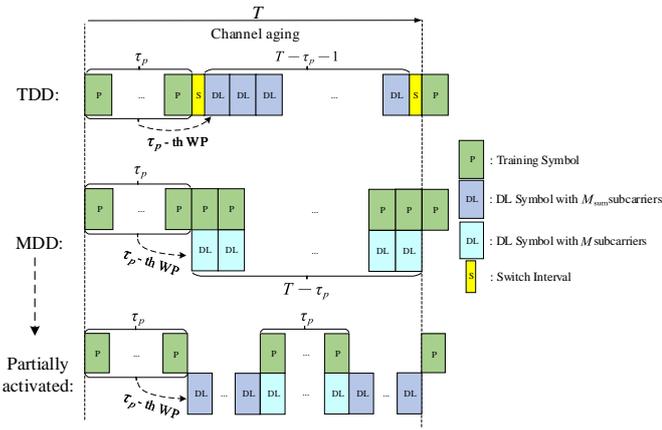
Fig. 1. Type I frame structure for the conventional TDD and proposed MDD systems.

where the second term at the right-hand side is multiuser interference, while the third term is due to SI.

III. FRAME STRUCTURE DESIGN

Due to the effect of channel aging, channel coefficients may vary from symbol to symbol, leading to poor system performance in high-mobility communication scenarios [4]. To overcome the time-varying fading problem, in conventional TDD mode, channel prediction is usually implemented after the training phase to reduce CSI error. However, when the relative velocity between transmitter and receiver increases, the channel becomes less correlated, increasing the prediction error, which in turn renders the later detected symbols less reliable. One way to mitigate this problem is to employ extra training symbols for updating the CSI during data transmission. Following this observation, if the channel aging problem becomes severe, more training symbols are inserted, which results in more frequent switching between UL training and DL data transmission, as well as the decrease of useful data rate. To this end, our main motivation of this paper is to demonstrate that the MDD mode is capable of efficiently solving the channel aging problem in fast fading communication scenarios. For demonstration and comparison, in the following subsections, we will present two types of frame structures for coping with the channel aging problem and also for analyzing the performance of MDD systems in terms of sum-rate.

A. Frame Structure: Type I

In this subsection, a simple scenario is considered, where the frame structure only includes UL training and DL transmission, as shown in Fig. 1. The length of one frame[3] is assumed to be equal to T OFDM symbol durations.

---

[3]Note that one frame defined in this paper may include several subframes used in 4G LTE systems or slots in 5G NR systems, where each subframe or slot contains 14 OFDM symbols. Specifically in 4G LTE, one frame consists of 10 subframes and two of them are used for training [16]. The frame structure in 5G NR is more flexible than that in 4G LTE [17].

1) TDD: As shown in Fig. 1, in the TDD mode, a frame consists of τ_p training symbols transmitted by UL and $T - \tau_p - 1$ symbols used for DL transmission. The switching from UL training to DL data transmission or from DL data transmission to the next training phase incurs a fixed cost, which is assumed to be half of an OFDM symbol.[4] To cope with channel aging, the channel prediction is implemented by a τ_p-th order WP[5] after the training phase, as shown in Fig. 1, and all the precoding/detection are operated at BS based on the predicted CSI. It is noteworthy that in practice, the τ_p training symbols can be first disassembled into several groups, which are then evenly distributed over one frame, so as to relieve the effect of accumulated prediction errors. However, the accompany of the increased switching intervals results in the reduced time for DL transmission and hence the decreased system's efficiency.

2) MDD: In the context of MDD, as UL/DL transmissions are on different subcarriers, CSI estimation and prediction can be implemented in the FD mode. That is to say, UL pilots can be kept active to update CSI in real time within one OFDM symbol. In detail, as shown in Fig. 1, UL training is always activated and hence, the channels for DL transmission can be predicted by a WP based on the newest τ_p pilots. However, in low-speed scenarios, the excessive transmission of UL pilots inevitably increases energy consumption and system complexity, which is unnecessary as the channels change slowly. Moreover, in the case where only the DL transmission is needed, continuous UL training on $\bar{M}$ subcarriers leads to only M of the M_{sum} subcarriers being used for the DL transmission. Therefore, in practice, the pilots within one MDD frame can be partially transmitted (as illustrated by the example presented in Fig. 1) so that some UL subcarriers can be 'borrowed' for DL data transmission to reduce the overhead of SI mitigation. In Section VI, the TDD/MDD frame structures with different distributions of pilots will be investigated.

B. Frame Structure: Type II

A more general frame structure consisting of the UL training followed by both UL transmission and DL transmission is considered for both TDD and MDD, as shown in Fig. 2.

1) TDD: In the TDD mode, as shown in Fig. 2, τ_p pilot symbols are transmitted in Phase I for initializing the transmission. During Phase II, the detected UL data symbols and possibly a part of pilots are used to predict channels by a τ_p-th order decision-directed (DD)-WP. With this setting, during the UL transmission, the channel knowledge can be continuously updated with the DD prediction principle. By contrast, during the DL transmission, channel updating is unavailable as there is no UL transmission. Consequently, the transmissions of the $T - \tau_p - \tau_u - 1$ DL symbols are based on the outdated

---

[4]The length of the switching period in LTE can be a different value, such as 1, 2, 3, 4, 9 or 10 OFDM symbols [16]. We choose 0.5 to minimize the influence of the switching interval on the evaluation of TDD system's performance.

[5]To ease analysis and facilitate comparison, only WP is considered in this paper. The Kalman predictor and the state-of-the-art deep learning approaches will be studied in our future work.
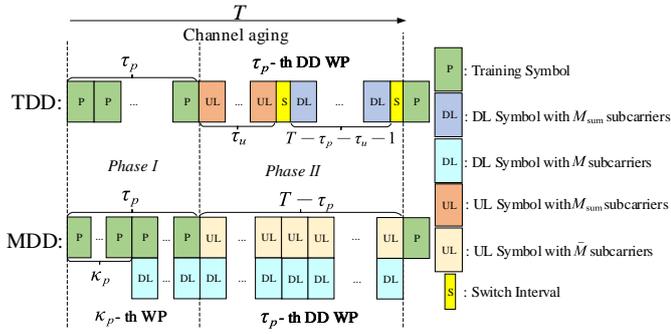
Fig. 2. Type II frame structure for conventional TDD and proposed MDD systems.

CSI, which is similar to the situation with the Type I frame structure.

2) MDD: As shown in Fig. 2, the Type II frame structure in MDD systems can be designed in a very efficient way. Firstly, for fair comparison with TDD, in Phase I, τ_p pilot symbols are still transmitted for the implementation of a τ_p-th order DD-WP in Phase II. However, different from TDD, during the transmission of pilot symbols, DL symbols can also be transmitted with the aid of the CSI predicted by a κ_p-th order WP, where $1 \leq \kappa_p \leq \tau_p$. In the extreme case, when $\kappa_p = 1$, $\tau_p - 1$ more DL symbols than TDD can be transmitted in Phase I, which may significantly increase the SE. During Phase II, both DL and UL can transmit $T - \tau_p$ symbols. Furthermore, a continuously updated τ_p-th order DD-WP can be implemented to predict the channels using the pilots and the detected UL symbols.

IV. CHANNEL ESTIMATION AND PREDICTION IN MDD SYSTEMS

In this section, the channel estimation and prediction in MDD systems are studied, when the proposed frame structures in Section III are introduced.

A. Channel Estimation

As the fading experienced by different subcarriers in an OFDM system is usually highly correlated, the CSI of all subcarriers can be acquired via transmitting pilots on a fraction of subcarriers [29]. With this regard, let us assume that all MTs synchronously transmit their frequency-domain pilot sequences (FDPS) over $\bar{M}$ UL subcarriers. Specifically, the FDPS transmitted by the d-th MT is expressed as $\boldsymbol{x}_{p,d} = \left[x_{p,d}[1], ..., x_{p,d}[\bar{m}], ..., x_{p,d}[\bar{M}] \right]^T$. Then, the received signal at the n-th antenna of BS corresponding to the i-th symbol is given by

$$\boldsymbol{y}_{p,n}[i] = \sqrt{p_{\text{UL}}} \sum_{d=1}^{D} \boldsymbol{X}_{p,d} \boldsymbol{h}_{n,d}^{\text{UL}}[i] + \boldsymbol{z}_n^{\text{SI}}[i] + \boldsymbol{z}_{p,n}[i]$$

$$= \sqrt{p_{\text{UL}}} \sum_{d=1}^{D} \boldsymbol{X}_{p,d} \boldsymbol{\Phi}_{\text{UL}} \boldsymbol{F} \boldsymbol{\Psi} \boldsymbol{g}_{n,d}[i] + \boldsymbol{z}_n^{\text{SI}}[i] + \boldsymbol{z}_{p,n}[i], \tag{12}$$

where $\boldsymbol{X}_{p,d} = \text{diag}\{\boldsymbol{x}_{p,d}\}, \boldsymbol{z}_{p,n}[i] \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{\bar{M}})$. Note that, in MDD systems, BS may experience SI during UL training, if DL symbols are transmitted at the same time. For example, in the frame structure as shown in Fig. 1, when $1 \leq i \leq \tau_p$, we have $\boldsymbol{z}_n^{\text{SI}}[i] = 0$, as there is no DL transmissions. By contrast, when $\tau_p < i \leq T$, $\boldsymbol{z}_n^{\text{SI}}[i] \sim \mathcal{CN}(0, \xi_{\text{BS}} p_{\text{DL}} M \boldsymbol{I}_{\bar{M}})$.

According to our previous study [12], the FDPS of MT d can be designed as

$$\boldsymbol{x}_{p,d} = \left[1, e^{j2\pi \frac{(d-1)\xi}{\bar{M}}}, e^{j2\pi \frac{2(d-1)\xi}{\bar{M}}}, \cdots, e^{j2\pi \frac{(\bar{M}-1)(d-1)\xi}{\bar{M}}} \right]^T, \tag{13}$$

where $\xi = \left\lfloor \frac{\bar{M}}{D} \right\rfloor$. With the assumption that $\bar{M} \geq DL$ and the $\bar{M}$ number of UL subcarriers are evenly distributed, $\boldsymbol{P}_d = \boldsymbol{X}_{p,d} \boldsymbol{\Phi}_{\text{UL}} \boldsymbol{F} \boldsymbol{\Psi}$ for $d = 1, ..., D$ are mutually orthogonal, i.e.,

$$\begin{cases} \boldsymbol{P}_d^H \boldsymbol{P}_d = \frac{\bar{M}}{M_{\text{sum}}} \boldsymbol{I}_L, \\ \boldsymbol{P}_d^H \boldsymbol{P}_k = \boldsymbol{0}_L, \ \forall \ d \neq k \end{cases}. \tag{14}$$

Then, the noisy observation of $\boldsymbol{g}_{n,d}[i]$ is given by

$$\tilde{\boldsymbol{y}}_{p,n}^d[i] = \boldsymbol{P}_d^H \boldsymbol{y}_{p,n}[i]$$

$$= \frac{\sqrt{p_{\text{UL}}} \bar{M}}{M_{\text{sum}}} \boldsymbol{g}_{n,d}[i] + \boldsymbol{P}_d^H (\boldsymbol{z}^{\text{SI}}[i] + \boldsymbol{z}_{p,n}[i]). \tag{15}$$

Correspondingly, the minimum mean square error (MMSE) estimate to $\boldsymbol{g}_{n,d}[i]$ is given by

$$\hat{\boldsymbol{g}}_{n,d}[i] = \frac{\frac{\sqrt{p_{\text{UL}}} \bar{M}}{M_{\text{sum}}} \frac{\beta_d}{L}}{\frac{p_{\text{UL}} \bar{M}^2}{M_{\text{sum}}^2} \frac{\beta_d}{L} + \frac{\xi_{\text{BS}} p_{\text{DL}} M \bar{M}}{M_{\text{sum}}} + \frac{\bar{M}}{M_{\text{sum}}} \sigma^2} \tilde{\boldsymbol{y}}_{p,n}^d[i]. \tag{16}$$

B. Channel Prediction

1) General Wiener Predictor: In this subsection, we focus on the prediction of $\boldsymbol{g}_{n,d}[i+1]$ based on the current and previous received training signals, which are expressed as $\bar{\boldsymbol{y}}_{p,n}^d[i] = \left[\tilde{\boldsymbol{y}}_{p,n}^{d,H}[i], \tilde{\boldsymbol{y}}_{p,n}^{d,H}[i-1], ..., \tilde{\boldsymbol{y}}_{p,n}^{d,H}[i+1-\tau_p] \right]^H$, where $\tilde{\boldsymbol{y}}_{p,n}^d[i]$ is given by (15). A τ_p-th order WP is first assumed, with the weights $\boldsymbol{V}_{n,d} = \left[\boldsymbol{V}_{n,d}^1, ..., \boldsymbol{V}_{n,d}^q, ..., \boldsymbol{V}_{n,d}^{\tau_p} \right], \boldsymbol{V}_{n,d}^q \in \mathbb{C}^{L \times L}$. Then, the predictor can be formulated as

$$\check{\boldsymbol{g}}_{n,d}^{\text{WP}}[i+1] = \sum_{q=1}^{\tau_p} \boldsymbol{V}_{n,d}^q \tilde{\boldsymbol{y}}_{p,n}^d[i+1-q] = \boldsymbol{V}_{n,d} \bar{\boldsymbol{y}}_{p,n}^d[i]. \tag{17}$$

According to the principles of Wiener filter [4, 30], the solution to $\boldsymbol{V}_{n,d}$ is

$$\boldsymbol{V}_{n,d} = \boldsymbol{R}_{g\bar{y}}^d[1] \left(\boldsymbol{R}_{\bar{y},i}^d[0] \right)^{-1}, \tag{18}$$

where the cross-correlation between the real channel of the next symbol and the observations is

$$\boldsymbol{R}_{g\bar{y}}^d[1] = \mathbb{E}\left[\boldsymbol{g}_{n,d}[i+1] \bar{\boldsymbol{y}}_{p,n}^{d,H}[i] \right] = \frac{\sqrt{p_{\text{UL}}} \bar{M}}{M_{\text{sum}}} [\boldsymbol{\delta}(\tau_p, \alpha) \otimes \boldsymbol{R}_g^d], \tag{19}$$

where $\boldsymbol{\delta}(\tau_p, \alpha) = [\alpha, \alpha^2, ..., \alpha^{\tau_p}]$. In (18), the autocorrelation matrix of the training signals is

$$\boldsymbol{R}_{\bar{y},i}^d[0] = \bar{\boldsymbol{y}}_{p,n}^d[i] \bar{\boldsymbol{y}}_{p,n}^{d,H}[i]$$

$$= \boldsymbol{\xi}(\tau_p, \alpha) \otimes \left(\frac{p_{\text{UL}} \bar{M}^2}{M_{\text{sum}}^2} \boldsymbol{R}_g^d \right) + \frac{\bar{M} \sigma^2}{M_{\text{sum}}} \boldsymbol{I}_{L\tau_p} + \boldsymbol{\Xi}_{\text{SI}}[i], \tag{20}$$

where

$$\boldsymbol{\xi}(\tau_p, \alpha) = \begin{bmatrix} 1 & \alpha & \cdots & \alpha^{\tau_p-1} \\ \alpha & 1 & \cdots & \alpha^{\tau_p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{\tau_p-1} & \alpha^{\tau_p-2} & \cdots & 1 \end{bmatrix}. \quad (21)$$

It can be observed that in the Type I frame structure for fast-fading channels, the amount of SI is dependent on the observation time, which leads to

$$\boldsymbol{\Xi}_{\mathrm{SI}}[i] = \begin{cases} \boldsymbol{0}_{L\tau_p \times L\tau_p}, & i = \tau_p \\ \begin{bmatrix} \frac{\xi_{\mathrm{BS}}p_{\mathrm{DL}}M\bar{M}}{M_{\mathrm{sum}}}\boldsymbol{I}_{(i-\tau_p)L} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, & \tau_p < i \le 2\tau_p \\ \frac{\xi_{\mathrm{BS}}p_{\mathrm{DL}}M\bar{M}}{M_{\mathrm{sum}}}\boldsymbol{I}_{\tau_p L}, & i > 2\tau_p \end{cases} . \quad (22)$$

According to the properties of WP, $\boldsymbol{g}_{n,d}[i+1]$ can be orthogonally decomposed into

$$\boldsymbol{g}_{n,d}[i+1] = \check{\boldsymbol{g}}_{n,d}^{\mathrm{WP}}[i+1] + \check{\boldsymbol{v}}_{n,d}[i+1], \quad (23)$$

where $\check{\boldsymbol{v}}_{n,d}[i+1] \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_g^d - \boldsymbol{\Upsilon}_{d,i})$ is the un-correlated channel prediction error vector with $\boldsymbol{\Upsilon}_{d,i} = \boldsymbol{R}_{g\check{y}}^d[1]\boldsymbol{R}_{\check{y},i}^{d,-1}[0]\boldsymbol{R}_{g\check{y}}^{d,H}[1]$.

Similar to (7), the predicted DL subchannels can be attained, which are expressed as

$$h_{n,d}[i+1, m] = \check{h}_{n,d}^{\mathrm{WP}}[i+1, m] + \underbrace{\boldsymbol{\psi}_m \check{\boldsymbol{v}}_{n,d}[i+1]}_{\check{v}_{n,d}^{\mathrm{WP}}[i+1, m]}, \ m = 1, ..., M, \quad (24)$$

where $\check{h}_{n,d}^{\mathrm{WP}}[i+1, m] \sim \mathcal{CN}\left(0, \sigma_{\check{h}_{d,i,m}^{\mathrm{WP}}}^2\right)$ with $\sigma_{\check{h}_{d,i,m}^{\mathrm{WP}}}^2 = \boldsymbol{\psi}_m \boldsymbol{\Upsilon}_{d,i}\boldsymbol{\psi}_m^H$, and $\check{v}_{n,d}^{\mathrm{WP}}[i+1, m] \sim \mathcal{CN}(0, \sigma_{\check{v}_{d,i,m}^{\mathrm{WP}}}^2)$ with $\sigma_{\check{v}_{d,i,m}^{\mathrm{WP}}}^2 = R_h^d - \sigma_{\check{h}_{d,i,m}^{\mathrm{WP}}}^2$. Moreover, as the channels of different antennas are assumed to be independent, the channels between BS antennas and the d-th MT can be collected to a vector as

$$\check{\boldsymbol{h}}_d^{\mathrm{WP}}[i+1, m] = \boldsymbol{h}_d[i+1, m] + \check{\boldsymbol{v}}_d^{\mathrm{WP}}[i+1, m], \quad (25)$$

where $\check{\boldsymbol{h}}_d^{\mathrm{WP}}[i+1, m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma_{\check{h}_{d,i,m}^{\mathrm{WP}}}^2 \boldsymbol{I}_N)$ and $\check{\boldsymbol{v}}_d^{\mathrm{WP}}[i+1, m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma_{\check{v}_{d,i,m}^{\mathrm{WP}}}^2 \boldsymbol{I}_N)$.

2) Decision-Directed Wiener Predictor: In the proposed Type II frame structure, as UL transmission is activated, channels can be predicted in the decision-directed (DD) principles [31]. Assume that a τ_p-th order DD-WP is performed based on the τ_p latest detected symbols received by BS. According to (10), the received signal by the n-th antenna of BS over the $\bar{m}$-th UL subcarrier and the i-th symbol duration is

$$s_n[i, \bar{m}] = \sqrt{p_{\mathrm{UL}}}\boldsymbol{x}^T[i, \bar{m}]\boldsymbol{h}_n[i, \bar{m}] + z^{\mathrm{SI}}[i] + z[i, \bar{m}], \quad (26)$$

where $\boldsymbol{x}[i, \bar{m}] = [x_1[i, \bar{m}], ..., x_D[i, \bar{m}]]^T$, $\boldsymbol{h}_n[i, \bar{m}] = [h_{n,1}[i, \bar{m}], ..., h_{n,D}[i, \bar{m}]]^T$ and $z^{\mathrm{SI}}[i] \sim \mathcal{CN}(0, \xi_{\mathrm{BS}}p_{\mathrm{DL}}M)$.

Based on the DD signals of (26), a τ_p-th order DD-WP is employed to predict the channels $\boldsymbol{h}_n[i+1, \bar{m}]$ as $\check{\boldsymbol{h}}_n^{\mathrm{DD}}[i+1, \bar{m}] = \boldsymbol{V}_{\mathrm{DD}}\bar{\boldsymbol{s}}_n[i, \bar{m}]$, where $\bar{\boldsymbol{s}}_n[i, \bar{m}] = [s_n[i, \bar{m}], ..., s_n[i+1-\tau_p, \bar{m}]]^T$, and $\boldsymbol{V}_{\mathrm{DD}}$ can be expressed as

$$\boldsymbol{V}_{\mathrm{DD}} = \boldsymbol{R}_{h\bar{s}}[1, \bar{m}]\boldsymbol{R}_{\bar{s},i}^{-1}[0, \bar{m}], \quad (27)$$

with $\boldsymbol{R}_{h\bar{s}}[1, \bar{m}] = \mathbb{E}\left[\boldsymbol{h}_n[i+1, \bar{m}]\bar{\boldsymbol{s}}_n^H[i, \bar{m}]\right]$ and $\boldsymbol{R}_{\bar{s},i}[0, \bar{m}] = \mathbb{E}\left[\bar{\boldsymbol{s}}_n[i, \bar{m}]\bar{\boldsymbol{s}}_n^H[i, \bar{m}]\right]$, which are detailed below.

First, the cross-correlation matrix in (27) can be obtained with the aid of (7) and (26), which is

$$\begin{aligned} \boldsymbol{R}_{h\bar{s}}[1, \bar{m}] &= \mathbb{E}\left[\boldsymbol{h}_n[i+1, \bar{m}]\bar{\boldsymbol{s}}_n^H[i, \bar{m}]\right] \\ &= \sqrt{p_{\mathrm{UL}}}\left[\alpha\boldsymbol{R}_h\boldsymbol{x}^*[i, \bar{m}], ..., \alpha^{\tau_p}\boldsymbol{R}_h\boldsymbol{x}^*[i+1-\tau_p, \bar{m}]\right] \quad (28) \\ &= \sqrt{p_{\mathrm{UL}}}\boldsymbol{A}(\tau_p, \alpha)\boldsymbol{Q}(\tau_p, D)\boldsymbol{B}^H(\tau_p, D), \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{R}_h &= \mathrm{diag}(R_h^1, ..., R_h^D), \ \boldsymbol{Q}(\tau_p, D) = \boldsymbol{R}_h \otimes \boldsymbol{I}_{\tau_p}, \\ \boldsymbol{A}(\tau_p, \alpha) &= \boldsymbol{I}_D \otimes \boldsymbol{\delta}(\tau_p, \alpha), \\ \boldsymbol{B}(\tau_p, D) &= [\mathrm{diag}(\hat{\boldsymbol{x}}_1[\bar{m}]), ..., \mathrm{diag}(\hat{\boldsymbol{x}}_d[\bar{m}]), ..., \mathrm{diag}(\hat{\boldsymbol{x}}_D[\bar{m}])], \end{aligned} \quad (29)$$

and in $\boldsymbol{B}(\tau_p, D)$, $\hat{\boldsymbol{x}}_d[\bar{m}] = [\hat{x}_d[i, \bar{m}], ..., \hat{x}_d[i+1-\tau_p, \bar{m}]]^T$ are UL detected symbols. To obtain (28), each element in $\hat{\boldsymbol{x}}_d[\bar{m}]$ needs to be correctly detected. This can be achieved by measuring the reliability of each detection in terms of the likelihood value of a detected symbol. If the likelihood value is larger than a preset threshold resulting in high reliability, the symbol is then included for DD prediction. Otherwise, the symbol will not be used for DD channel prediction.

The auto-correlation matrix $\boldsymbol{R}_{\bar{s},i}[0, \bar{m}]$ in (27) is given by

$$\begin{aligned} \boldsymbol{R}_{\bar{s},i}[0, \bar{m}] &= \mathbb{E}\left[\bar{\boldsymbol{s}}_n[i, \bar{m}]\bar{\boldsymbol{s}}_n^H[i, \bar{m}]\right] \\ &= p_{\mathrm{UL}}\boldsymbol{B}(\tau_p, D)\boldsymbol{C}(\tau_p, D)\boldsymbol{Q}(\tau_p, D)\boldsymbol{B}^H(\tau_p, D) \\ &\quad + \sigma^2\boldsymbol{I}_{\tau_p} + \boldsymbol{\Xi}_{\mathrm{SI}}^{\mathrm{DD}}[i], \end{aligned} \quad (30)$$

where $\boldsymbol{C}(\tau_p, D) = \boldsymbol{I}_D \otimes \boldsymbol{\xi}(\tau_p, \alpha)$. When the proposed Type II frame structure is employed, SI is also present at the observation time, which leads to

$$\boldsymbol{\Xi}_{\mathrm{SI}}^{\mathrm{DD}}[i] = \begin{cases} \begin{bmatrix} \xi_{\mathrm{BS}}p_{\mathrm{DL}}M\boldsymbol{I}_{\tau_p-1} & \boldsymbol{0} \\ \boldsymbol{0} & 0 \end{bmatrix}, & i = \tau_p, \\ \xi_{\mathrm{BS}}p_{\mathrm{DL}}M\boldsymbol{I}_{\tau_p}, & i > \tau_p. \end{cases} \quad (31)$$

With the aid of the orthogonal decomposition, we have the expression

$$\boldsymbol{h}_n[i+1, \bar{m}] = \check{\boldsymbol{h}}_n^{\mathrm{DD}}[i+1, \bar{m}] + \check{\boldsymbol{e}}_n^{\mathrm{DD}}[i+1, \bar{m}], \quad (32)$$

where $\check{\boldsymbol{e}}_n^{\mathrm{DD}}[i+1, \bar{m}] \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_h - \boldsymbol{\Theta}_i[\bar{m}])$ is the un-correlated channel prediction error vector, and $\boldsymbol{\Theta}_i[\bar{m}] = \boldsymbol{R}_{h\bar{s}}[1, \bar{m}]\boldsymbol{R}_{\bar{s},i}^{-1}[0, \bar{m}]\boldsymbol{R}_{h\bar{s}}^H[1, \bar{m}]$, which is the covariance matrix of $\check{\boldsymbol{h}}_n^{\mathrm{DD}}[i+1, \bar{m}]$. Furthermore, we can write $\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1] = \left[\check{h}_{n,d}^{\mathrm{DD}}[i+1, 1], ..., \check{h}_{n,d}^{\mathrm{DD}}[i+1, \bar{M}]\right]^T$, which follows the distribution of $\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1] \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{\Gamma}_{d,i})$, with $\boldsymbol{\Gamma}_{d,i}$ derived in detail in Appendix B.

Having obtained the frequency-domain prediction of $\boldsymbol{h}_{n,d}[i+1]$ over UL subcarriers, the corresponding time-domain CSI can be found as

$$\check{\boldsymbol{g}}_{n,d}^{\mathrm{DD}}[i+1] = \boldsymbol{J}\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1], \quad (33)$$

provided that $\bar{M} \ge L$. In (33), $\boldsymbol{J} = \left((\boldsymbol{\Phi}_{\mathrm{UL}}\boldsymbol{F}\boldsymbol{\Psi})^H(\boldsymbol{\Phi}_{\mathrm{UL}}\boldsymbol{F}\boldsymbol{\Psi})\right)^{-1}(\boldsymbol{\Phi}_{\mathrm{UL}}\boldsymbol{F}\boldsymbol{\Psi})^H$. Finally, upon following the same approach described in (23)-(24), the DL

$$R_{d,i,m}^{\text{MDD-1}} \geq R_{d,i,m}^{\text{MDD-1-LB}} = \log_2 \left(1 + \frac{p_{\text{DL}} \left| \mathbb{E} \left[\omega_{d,d}^{\text{ZF}}[i,m] \right] \right|^2}{p_{\text{DL}} \text{var} \left\{ \omega_{d,d}^{\text{ZF}}[i,m] \right\} + p_{\text{DL}} \sum_{k=1, k \neq d}^{D} \mathbb{E} \left[\left| \omega_{d,k}^{\text{ZF}}[i,m] \right|^2 \right] + \text{var} \left\{ z_d^{\text{SI}} \right\} + \sigma^2} \right). \tag{36}$$

$$R_{d,i,m}^{\text{MDD-1-LB}} \approx \log_2 \left(1 + \frac{p_{\text{DL}} (N - D + 1) \boldsymbol{\psi}_m \boldsymbol{\Upsilon}_{d,i} \boldsymbol{\psi}_m^H}{0.25 p_{\text{DL}} \boldsymbol{\psi}_m \boldsymbol{\Upsilon}_{d,i} \boldsymbol{\psi}_m^H + p_{\text{DL}} D \left(\frac{\beta_d}{M_{\text{sum}}} - \boldsymbol{\psi}_m \boldsymbol{\Upsilon}_{d,i} \boldsymbol{\psi}_m^H \right) + \xi_{\text{MT}} p_{\text{UL}} D \bar{M} + D \sigma^2} \right). \tag{37}$$

channel vector between BS and the d-th MT with respect to the m-th DL subcarrier and the $(i + 1)$ symbol duration can be predicted as $\check{\boldsymbol{h}}_d^{\text{DD}}[i+1,m]$, which follows

$$\check{\boldsymbol{h}}_d^{\text{DD}}[i+1,m] = \boldsymbol{h}_d[i+1,m] + \check{\boldsymbol{v}}_d^{\text{DD}}[i+1,m], \ m = 1,...,M, \tag{34}$$

where $\check{\boldsymbol{h}}_d^{\text{DD}}[i + 1, m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma_{\check{h}_{d,i,m}^{\text{DD}}}^2 \boldsymbol{I}_N)$, $\sigma_{\check{h}_{d,i,m}^{\text{DD}}}^2 = \boldsymbol{\psi}_m \boldsymbol{J} \boldsymbol{\Gamma}_{d,i} \boldsymbol{J}^H \boldsymbol{\psi}_m^H$, and $\check{\boldsymbol{v}}_d^{\text{DD}}[i + 1, m] \sim \mathcal{CN}(\boldsymbol{0}, \sigma_{\check{v}_{d,i,m}^{\text{DD}}}^2 \boldsymbol{I}_N)$, $\sigma_{\check{v}_{d,i,m}^{\text{DD}}}^2 = R_h^d - \sigma_{\check{h}_{d,i,m}^{\text{DD}}}^2$.

V. Performance Analysis Over fast-fading channels

In this section, the ergodic rates attainable by the MDD and TDD systems with the proposed channel acquisition are derived, where the two types of frame structures introduced in Section III are respectively applied.

A. Type I Frame Structure

The ergodic achievable rate of the MDD systems with the Type I frame structure shown in Fig. 1 can be expressed as

$$R^{\text{MDD-1}}(\tau_p) = \frac{1}{T M_{\text{sum}}} \sum_{d=1}^{D} \sum_{i=\tau_p+1}^{T} \sum_{m=1}^{M} R_{d,i,m}^{\text{MDD-1}}, \tag{35}$$

where $R_{d,i,m}^{\text{MDD-1}}$ denotes the rate obtained by the d-th MT over the m-th DL subcarrier during the i-th symbol duration. Note that since no DL pilots are employed, only the expected effective channel gain is available at MTs. Hence, the mean effective channel gain is treated as the CSI for signal detection at MTs' sides. The effective channel gain is $\omega_{d,d}^{\text{ZF}}[i,m] = \boldsymbol{h}_d^H[i,m] \check{\boldsymbol{f}}_d[i,m]$, where $\check{\boldsymbol{f}}_d[i,m]$ is derived by the ZF principle with the predicted channel $\check{\boldsymbol{h}}_d^{\text{WP}}[i+1,m]$ obtained in (25). Then, the lower bounded achievable ergodic rate of $R_{d,i,m}^{\text{MDD-1-LB}}$ can be derived as shown in (36).

According to Lemma 4 in [32], the expectation and variance of the effective channel gain $\omega_{d,d}^{\text{ZF}}[i,m]$ can be approximated as

$$\mathbb{E}\left[\omega_{d,d}^{\text{ZF}}[i,m]\right] = \mathbb{E}\left[1/(\sqrt{D} \left\| \check{\boldsymbol{f}}_d[i,m] \right\|)\right]$$
$$\approx \sqrt{\frac{N-D+1}{D}} \sqrt{\sigma_{\check{h}_{d,i,m}^{\text{WP}}}^2} \tag{38}$$

and

$$\text{var}\left\{\omega_{d,d}^{\text{ZF}}[i,m]\right\} \approx \frac{1}{D}\left(\frac{1}{4}\sigma_{\check{h}_{d,i,m}^{\text{WP}}}^2 + \sigma_{\check{v}_{d,i,m}^{\text{WP}}}^2\right), \tag{39}$$

respectively, where $\check{\boldsymbol{f}}_d[i,m]$ is the unnormalized column of the matrix $\check{\boldsymbol{F}}^{\text{ZF}}[i,m]$ seen in (8). For the interference signal from the other MTs, as seen in the denominator of (36), we have

$$\mathbb{E}\left[\left|\omega_{d,k}^{\text{ZF}}[i,m]\right|^2\right]$$
$$= \mathbb{E}\left[\left|\check{\boldsymbol{h}}_d^{\text{WP},H}[i,m]\check{\boldsymbol{f}}_k[i,m] + \check{\boldsymbol{v}}_d^{\text{WP},H}[i+1,m]\check{\boldsymbol{f}}_k[i,m]\right|^2\right]$$
$$= \mathbb{E}\left[\left|\check{\boldsymbol{v}}_d^{\text{WP},H}[i+1,m]\check{\boldsymbol{f}}_k[i,m]\right|^2\right]$$
$$= \frac{1}{D}\sigma_{\check{v}_{d,i,m}^{\text{WP}}}^2. \tag{40}$$

Consequently, upon substituting (24), (38), (40) and (48) into (36), the approximation to the lower bounded achievable ergodic rate of the d-th MT on the m-th DL subcarrier within the i-th OFDM symbol duration can be expressed as (37). Finally, $R^{\text{MDD-1-LB}}(\tau_p)$ is obtained by substituting (37) into (35).

On the other hand, the achievable ergodic rate of the TDD systems with the Type I frame structure shown in Fig. 1 can be written as

$$R^{\text{TDD-1}}(\tau_p) = \frac{1}{T M_{\text{sum}}} \sum_{d=1}^{D} \left(\sum_{i=\tau_p+2}^{T-1} \sum_{m=1}^{M_{\text{sum}}} R_{d,i,m}^{\text{TDD-1}} \right.$$
$$\left. + \frac{1}{2} \sum_{m=1}^{M_{\text{sum}}} \left(R_{d,\tau_p+1,m}^{\text{TDD-1}} + R_{d,T,m}^{\text{TDD-1}} \right) \right). \tag{41}$$

Furthermore, the approximated lower bounded achievable ergodic rate of TDD systems can be derived by following nearly the same steps as above for the MDD systems, except only the following two points. The first one is that all the channels within one frame are predicted relying on the first τ_p training symbols sent at the start of the frame, yielding $\boldsymbol{\delta}(\tau_p, \alpha)$ in (19) being changed to $[\alpha^{i-\tau_p}, \alpha^{i+1-\tau_p}, ..., \alpha^{i-1}]$. The second point is that the TDD systems do not experience SI. Therefore, we have $\boldsymbol{\Xi}_{\text{SI}}[i] = 0, \forall i$ in (22).

B. Type II Frame Structure

The ergodic achievable rate of the MDD systems with the Type II frame structure shown in Fig. 2 can be expressed as

$$R^{\text{MDD-2}}(\tau_p, \kappa_p) = \frac{1}{T M_{\text{sum}}} \sum_{d=1}^{D} \left(\sum_{i=\kappa_p+1}^{\tau_p} \sum_{m=1}^{M} R_{d,i,m}^{\text{MDD-1}} \right.$$
$$\left. + \sum_{i=\tau_p+1}^{T} \sum_{m=1}^{M} R_{d,i,m}^{\text{MDD-2-DL}} + \sum_{i=\tau_p+1}^{T} \sum_{\bar{m}=1}^{\bar{M}} R_{d,i,\bar{m}}^{\text{MDD-2-UL}} \right), \tag{42}$$

$$R_{d,i,\bar{m}}^{\text{MDD-2-UL}} \geq R_{d,i,\bar{m}}^{\text{MDD-2-UL-LB}} = \log_2 \left(1 + \frac{p_{\text{UL}} \left| \mathbb{E}\left[\omega_{d,d}^{\text{MRC}}[i,\bar{m}] \right] \right|^2}{p_{\text{UL}} \sum\limits_{k=1,k\neq d}^{D} \mathbb{E}\left[\left| \omega_{d,k}^{\text{MRC}}[i,\bar{m}] \right|^2 \right] + (\text{var}\{\boldsymbol{z}^{\text{SI}}[i]\} + \sigma^2) \mathbb{E}\left[\|\check{\boldsymbol{w}}_d[i,\bar{m}]\|_2^2 \right]} \right)$$

$$\overset{(a)}{=} \log_2 \left(1 + \frac{p_{\text{UL}} \left| \mathbb{E}\left[\check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \boldsymbol{h}_d[i,\bar{m}] \right] \right|^2}{p_{\text{UL}} \sum\limits_{k=1,k\neq d}^{D} \mathbb{E}\left[\left| \check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \boldsymbol{h}_k[i,\bar{m}] \right|^2 \right] + (\xi_{\text{BS}} p_{\text{DL}} M + \sigma^2) \mathbb{E}\left[\left\| \check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}] \right\|_2^2 \right]} \right), \tag{43}$$

where $R_{d,i,m}^{\text{MDD-1}}$ denotes the rate attained during Phase I, $R_{d,i,m}^{\text{MDD-2-DL}}$ and $R_{d,i,\bar{m}}^{\text{MDD-2-UL}}$ denote respectively the DL and UL rates achievable during Phase II. To compute the lower bounded rate $R^{\text{MDD-2-LB}}(\tau_p.\kappa_p)$, in (42), the lower bound for $R_{d,i,m}^{\text{MDD-1}}$ is given by (37), while the lower bound for $R_{d,i,m}^{\text{MDD-2-DL}}$ can be obtained from the approach for computing $R_{d,i,m}^{\text{MDD-1-LB}}$ by replacing $\check{h}_{d,i,m}^{\text{WP}}$ using $\check{h}_{d,i,m}^{\text{DD}}$. Therefore, only the lower bound for the achievable UL ergodic rate needs to be derived, i.e., $R_{d,i,\bar{m}}^{\text{MDD-2-UL}}$. Denote the effective channel gain as $\omega_{d,d}^{\text{MRC}}[i,\bar{m}] = \check{\boldsymbol{w}}_d^H[i,\bar{m}] \boldsymbol{h}_d[i,\bar{m}]$. Then, we have (43) at the top of next page, where (a) is obtained by using $\check{\boldsymbol{w}}_d[i,\bar{m}] = \check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}]$ to implement the MRC receiver. In particular, the expectation term in the numerator of (43) can be derived as

$$\left| \mathbb{E}\left[\check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \boldsymbol{h}_d[i,\bar{m}] \right] \right|^2$$
$$= \left| \mathbb{E}\left[\check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \left(\check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}] + \check{\boldsymbol{e}}_d^{\text{DD}}[i,\bar{m}] \right) \right] \right|^2$$
$$\overset{(b)}{=} \left| \mathbb{E}\left[\check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}] \right] \right|^2$$
$$= \left| \mathbb{E}\left[\left\| \check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}] \right\|_2^2 \right] \right|^2 \tag{44}$$

where (b) is due to the fact that $\check{\boldsymbol{e}}_d^{\text{DD}}[i,m]$ has zero mean and is independent of $\check{\boldsymbol{h}}_d[i,\bar{m}]$. The first expectation in the denominator of (43) can be derived as

$$\mathbb{E}\left[\left| \check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \boldsymbol{h}_k[i,\bar{m}] \right|^2 \right]$$
$$= \mathbb{E}\left[\left| \check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \check{\boldsymbol{h}}_k^{\text{DD}}[i,\bar{m}] + \check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \check{\boldsymbol{e}}_k^{\text{DD}}[i,\bar{m}] \right|^2 \right]$$
$$\overset{(c)}{\leq} \mathbb{E}\left[\left| \check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \check{\boldsymbol{h}}_k^{\text{DD}}[i,\bar{m}] \right|^2 \right] + \mathbb{E}\left[\left| \check{\boldsymbol{h}}_d^{\text{DD},H}[i,\bar{m}] \check{\boldsymbol{e}}_k^{\text{DD}}[i,\bar{m}] \right|^2 \right]$$
$$\overset{(d)}{=} R_h^k \mathbb{E}\left[\left\| \check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}] \right\|_2^2 \right], \tag{45}$$

where (c) and (d) follow from the fact that $\check{\boldsymbol{h}}_d^{\text{DD}}[i,\bar{m}]$, $\check{\boldsymbol{h}}_k^{\text{DD}}[i,\bar{m}]$ and $\check{\boldsymbol{e}}_k^{\text{DD}}[i,\bar{m}]$ are independent vectors and $\check{\boldsymbol{e}}_k^{\text{DD}}[i,\bar{m}]$ has zero mean. Consequently, upon substituting (32), (44) and (45) into (43), the approximated lower-bounded UL rate can be

TABLE I
SIMULATION PARAMETERS

Default parameters	Value
Distance between BS and MTs (D_d)	$[50, 100]$ m
Number of BS antennas (N)	32
Number of MTs (D)	8
Number of DL/UL subcarriers for TDD (M_{sum})	96
Number of DL/UL subcarriers for MDD ($M/\bar{M}$)	64 / 32
BS transmit power ($p_{\text{DL}} M$)	30 dBm
MT transmit power ($p_{\text{UL}} M$)	20 dBm
Noise power	-94 dBm
Large-scale fading (β_d)	$(D_d)^{3.8}$
Delay taps (L)	4
Carrier center frequency	5 GHz
Subcarrier bandwidth	15 kHz
OFDM symbol duration (T_s)	66.67 μs
Total symbols in one frame (T)	28
Switching interval in TDD	0.5
Digital modulation	16 QAM
SIC capability at BS (ξ_{BS})	130 dB
SIC capability at MT (ξ_{MT})	120 dB

obtained, which is given by

$$R_{d,i,\bar{m}}^{\text{MDD-2-UL-LB}} \approx \log_2 \left(1 + \frac{p_{\text{UL}} N \left(\boldsymbol{\Theta}_i[\bar{m}] \right)_{d,d}}{p_{\text{UL}} \sum\limits_{k=1,k\neq d}^{D} \frac{\beta_k}{M_{\text{sum}}} + \xi_{\text{BS}} p_{\text{DL}} M + \sigma^2} \right). \tag{46}$$

In the context of TDD, the achievable ergodic rate in the case of the Type II frame structure shown in Fig. 2 can be expressed as

$$R^{\text{TDD-2}}(\tau_p, \tau_u) = \frac{1}{TM_{\text{sum}}} \sum_{d=1}^{D} \left(\sum_{i=\tau_p+1}^{\tau_p+\tau_u} \sum_{m=1}^{M_{\text{sum}}} R_{d,i,m}^{\text{TDD-2-UL}} \right.$$
$$+ \sum_{i=\tau_p+\tau_u+2}^{T} \sum_{m=1}^{M_{\text{sum}}} R_{d,i,m}^{\text{TDD-2-DL}}$$
$$\left. + \frac{1}{2} \sum_{m=1}^{M_{\text{sum}}} \left(R_{d,\tau_p+\tau_u+1,m}^{\text{TDD-2-UL}} + R_{d,T,m}^{\text{TDD-2-DL}} \right) \right). \tag{47}$$

Based on (47), the approximated lower bound $R^{\text{TDD-2-LB}}(\tau_p, \tau_u)$ can be obtained in the way similar to that for MDD systems. Therefore, the derivation is omitted to avoid redundancy.
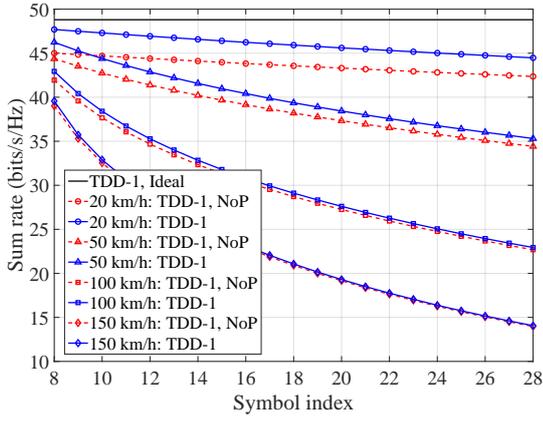
Fig. 3. Sum rate versus OFDM symbol index of the TDD systems with the different assumptions about channel acquisition and relative velocity.

VI. SIMULATION RESULTS AND DISCUSSION

In this section, the numerical results for the performance comparison of the MDD and TDD systems with two types of frame structures are present, as shown in Fig. 1 and Fig. 2, respectively. Also, the performance of IBFD and MDD systems with the proposed frame structures is compared, when SIC capability is limited. Furthermore, the analytical results derived in Section V is validated by Monte Carlo simulations over 1000 channel realizations. Throughout this section, a single-cell network with one BS and multiple single-antenna MTs is assumed. Some key parameters are listed in Table I. It is noteworthy that to demonstrate the tendency of the rate change within one frame, the switching intervals in the TDD mode are not considered, while the X-axis represents the index of OFDM symbol. By contrast, the switching intervals are taken into account in evaluating the average sum rate over one frame, as shown in (41) and (47).

A. Type I Frame Structure: TDD Vs MDD

In this subsection, TDD and MDD systems are compared under the Type I frame structure. For TDD systems, a total of 7 pilots are transmitted within one frame. By contrast, for MDD systems, UL pilots can always be transmitted or be transmitted in part. Note that, in the following figures, 'TDD-1' denotes the scheme that all pilots are consecutively transmitted at the beginning of the frame and a 7-th order WP is applied. The notations of 'TDD-1, Ideal' and 'TDD-1, NoP' indicate respectively that the channel is ideal without the channel aging problem and that the CSI acquisition does not rely on channel prediction. 'MDD-1(z)' denotes the scheme that UL pilots are continuously sent by MTs and a z-th order WP is applied. Furthermore, the results of 'TDD-1' and 'MDD-1(z)' are plotted by using the lower-bounded rates given in (35), (41), (42) and (47), while the results of 'TDD-1, Apx' and 'MDD-1(z), Apx' represent the correspondingly approximated rates derived in Section V.

In Fig. 3, the performance of the TDD systems operated in different mobility scenarios with or without channel prediction is evaluated. It can be observed that when the relative velocity is 20 km/h, the TDD system with WP outperforms the TDD
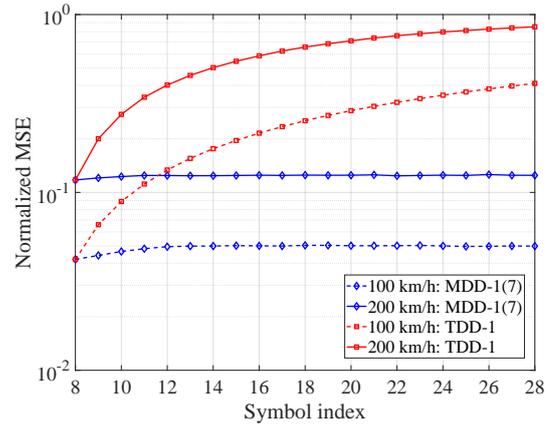
(a) Normalized MSE
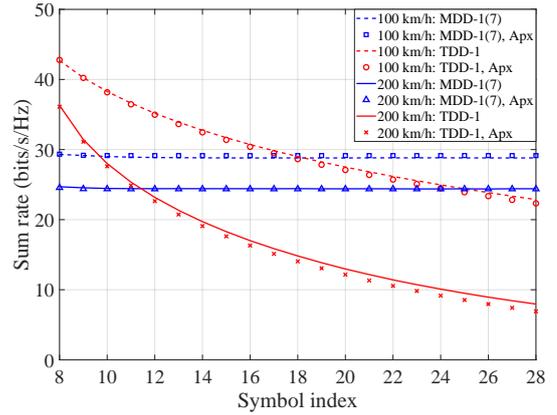
(b) Sum rate

Fig. 4. Performance comparison of the MDD and TDD systems, when Type I frame structure and 7-th order WP are employed.

system without channel prediction. However, as the symbol index increases, meaning that CSI becomes more outdated, the performance gap between the above two strategies becomes narrower. Therefore, in the TDD mode, the latter symbols benefit less from channel prediction, as only 7 symbols are used for prediction and the prediction error increases over time. Additionally, as seen in Fig. 3, as the relative speed increases, the advantage of employing channel prediction diminishes, and the performance drops faster and becomes significantly worse than the system with precise CSI. The results in Fig. 3 explicitly indicate that the traditional TDD mode is not feasible for communication in high-mobility scenarios, even when channel prediction is employed.

Based on the Type I frame structure presented in Fig. 1, the effect of the time-varying channels on the performance of TDD and MDD systems is demonstrated in Fig. 4. It can be observed from Fig. 4(a) that the accuracy of channel prediction in MDD systems is nearly constant after 7 pilot symbols, regardless of the relative traveling velocity of 100 km/h or 200 km/h. This is because in MDD systems, the UL pilots can continuously provide BS with the most updated CSI for DL transmission, which hence guarantees a stable system sum rate, as shown in Fig. 4(b). Moreover, the difference between the sum rates achieved by the MDD systems at

100 km/h and 200 km/h is relatively small, meaning that MDD systems are robust to time-varying fading channels. By contrast, although the sum rate of TDD systems is higher than that of MDD systems within the first several symbols, for both the velocities considered, it reduces quickly with time in terms of the OFDM symbol indices, due to the reduced accuracy of channel prediction with time. Furthermore, when the velocity is changed from 100 km/h to 200 km/h, the sum rate deteriorates significantly. Therefore, the performance of TDD systems is sensitive to the mobility of wireless channels. The rationale is that in TDD systems, UL training and DL have to be performed in an alternative way. Once it is switched to DL transmission, the CSI has to be predicted based on the old data, which may be outdated, if mobility is high. Additionally, in Fig. 4, it is demonstrated that the approximations of the lower bounded ergodic rates derived in Section V match closely to the lower bounded sum rates attained via Monte Carlo simulations.
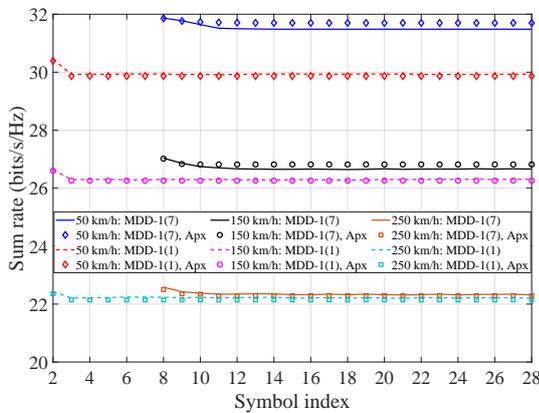
(a) Sum rate versus OFDM symbol index.
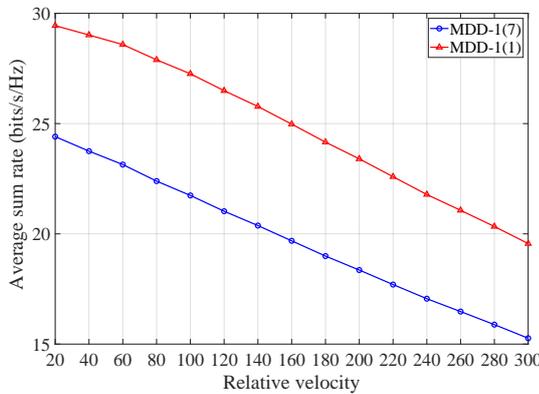
(b) Average sum rate within one frame versus relative velocity.

Fig. 5. Performance comparison of the MDD systems with Type I frame structure and different orders of WPs.

In Fig. 5, the performance of MDD systems is investigated, when Type I frame structure and different orders of WPs are employed. According to Fig. 1, a higher order WP uses more channel observations received in the past to predict channels. It is observed from Fig. 5(a) that at the relative speed of 50 km/h, the 7-th order WP in general significantly outperforms the 1-st order WP. However, when the speed is increased to 250 km/h, the 1-st order WP achieves nearly the same performance as the

(a) Evenly spaced pilots in TDD Type I frame structure.

(b) Two groups of pilots in TDD Type I frame structure.
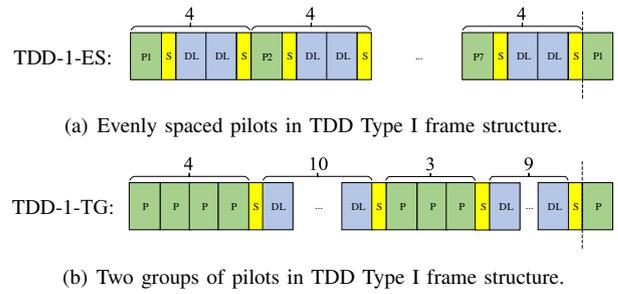
Fig. 6. Alternative pilot distribution methods for the Type I frame structure in TDD mode, when the total number of pilots is 7.

7-th order WP. The reason behind is that when channel varies fast, the prediction of the current channel is more dependent on the nearest pilot symbols and the observations in the past become less correlated to the prediction. Fig. 5(b) further shows that when pilots are continuously transmitted under the Type I frame structure, the 1-st order WP attains a better sum rate performance averaged over one frame, owing to the fact that in this case, 6 more DL symbols can be transmitted, even though the slightly reduced prediction accuracy results in some performance loss.

Next, to comprehensively compare the performance of the TDD and MDD systems employing the Type I frame, two more pilot distribution methods for TDD are considered, namely the TDD-1-ES and TDD-1-TG, as shown in Fig. 6. Specifically, with the TDD-1-ES, 7 training symbols are evenly distributed within one frame. By contrast, with the TDD-1-TG, 7 training symbols are first divided into two groups, which are then inserted into the frame. Note that the TDD-1-TG is similar to the frame structure applied in 3GPP 36.211, where two pilot subframes are evenly distributed within one frame [33]. For MDD systems, an alternative Type I frame structure denoted as MDD-1-PA is introduced, as depicted in Fig. 1. With the MDD-1-PA, $\tau_p = 1$, and a total of 7 pilots are activated, which are evenly distributed within one frame.

In Fig. 7, the average sum rate of the MDD and TDD systems with the various Type I frame structures is plotted, and when two frame lengths are considered. Note that, as the frame length increases, the number of pilots is set to remain the same in MDD-1-PA and three TDD structures, while the number of DL symbols increases. It can be observed from Fig. 7 that MDD-1-PA outperforms all the other arrangements, when the relative speed is lower than 280 km/h in Fig. 7(a) and 210 km/h in Fig. 7(b). This is because when only a fraction of pilots are transmitted, the saved UL resources (subcarriers) can be exploited to transmit DL data, provided that the CSI updating is enough for protecting the DL transmission. By contrast, MDD-1(1) is superior to the MDD-1-PA in the high-speed case. In this case, although there are only M subcarriers used for DL transmission, the continuously transmitted pilots on the $\bar{M}$ UL subcarriers guarantees the reliable channel prediction, which the MDD-1-PA structure is however unable to provide. With respect to TDD, the advantages of the three frame structures are dependent on the relative velocity and the length of frames. As shown in Fig. 7(a) and 7(b), in general,

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2022.3209799

12
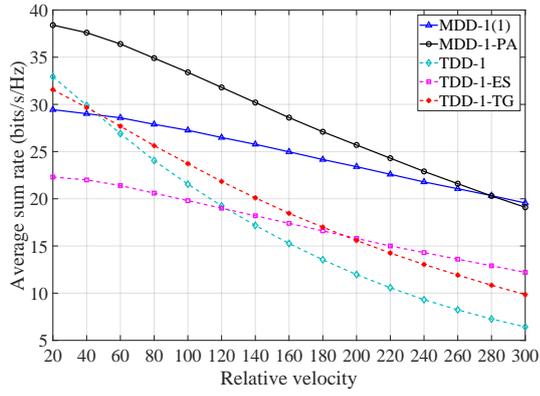
(a) $T = 28$ symbols per frame
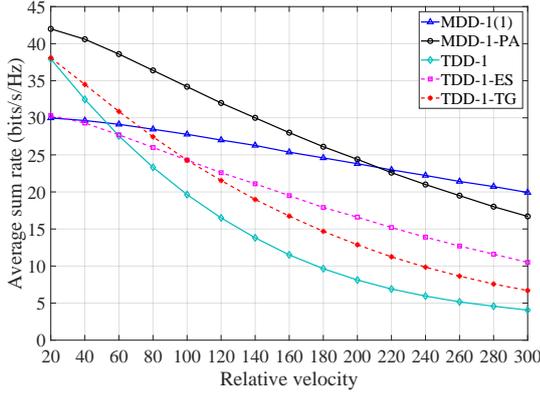
(b) $T = 56$ symbols per frame

Fig. 7. Average sum rate versus relative velocity, when the Type I frame structure is used.

(a) Normalized MSE

(b) Sum rate

Fig. 8. Performance comparison of the MDD and TDD systems, when Type II frame structure is applied.

TDD-1 and TDD-TG have relatively better performance in terms of average sum rates within one frame, when the relative velocity is low, owing to the use of less switching intervals. By contrast, TDD-1-ES benefits from the frequent transmission of training symbols within one frame and hence achieves higher average sum rates in the high-speed long-frame case.

B. Type II Frame Structure: TDD Vs MDD

In this subsection, the performance comparison of MDD and TDD systems operated with the Type II frame structure is presented. In the following figures, 'TDD-2' and 'MDD-2' denote respectively the TDD and MDD systems employing the proposed frame structure as shown in Fig. 2 with $\tau_p = 7$, $\tau_u = 7$ and $\kappa_p = 1$.

The performance comparison of MDD and TDD with respect to symbol index is plotted in Fig. 8, when Type II frame structure is considered. In conjunction with Fig. 2, it can be seen that, as the MDD system is able to operate DL and UL simultaneously during Phase I, DL data symbols can be transmitted while collecting the channel observations for prediction in Phase II. From Fig. 8(a), it can be observed that the WP in Phase I yields less prediction errors when compared with the DD-WP in Phase II. This is as expected, since the pilots used for WP in Phase I are known to BS. In Phase II, although the accuracy of channel prediction is higher without the effect of SI, TDD system only activates the UL

transmission from the 8-th to the 14-th symbol, leading to lower sum rate when compared with the MDD system, which is capable of transmitting data over both DL and UL. At the 15-th symbol, both MDD and TDD systems apply the 7-th DD-WP and the TDD system attains a higher sum rate, as it uses all the subcarriers for DL transmission. However, similar to the Type I frame structure, new UL data symbols are not available for the TDD system to update channel causing the degraded accuracy of channel prediction. Consequently, its sum rate decreases, as symbol index increases.

In Fig. 9, the average sum rate versus the relative speed of the MDD and TDD systems is plotted, when the Type II frame structure is employed. Here, an optional frame structure for TDD systems is also considered, which is shown in Fig. 10, where two groups of pilots are separately inserted into one frame. It can be observed from Fig. 9 that owing to the DL transmission occurred in both Phase I and the start of Phase II, MDD significantly outperforms TDD over the whole velocity, ranging from 20 km/h to 300 km/h, for both the frame-length considered. Specifically, for the TDD in low-mobility case, the rate achieved by the TDD-2-TG is considerably lower than that achieved by the TDD-2. However, benefiting from the extra training in the middle of the frame, the TDD-2-TG finally surpasses the TDD-2 at 240 km/h when $T = 28$, and at 140 km/h when $T = 56$. The results in Fig. 9 imply that in the general communication scenarios where UL training, UL

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2022.3209799

13
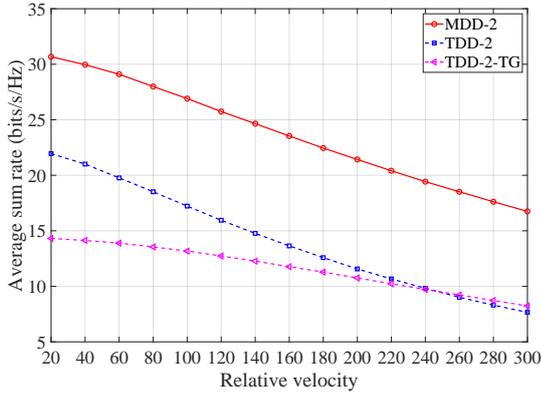
(a) $T = 28$ symbols per frame
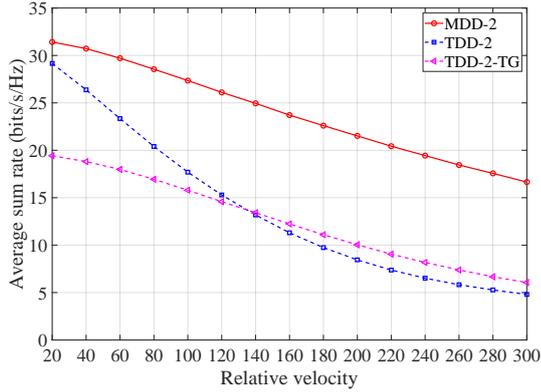
(b) $T = 56$ symbols per frame

Fig. 9. Average sum rate versus the relative velocity, when the Type II frame structure is used.
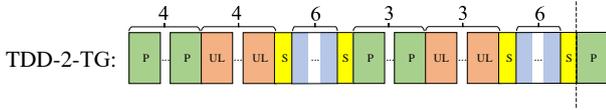
Fig. 10. Distribution of two groups of pilots in TDD Type II frame structure, when the total number of pilots is 7.

transmission and DL transmissions are all required at the same time, MDD shows significant advantages over TDD, when the channel aging problem is encountered.

C. Influence of SIC on IBFD and MDD

In this part, the impact of SIC on the performance of MDD and IBFD systems is demonstrated. To this end, the conventional IBFD with the Type I frame structure denoted as 'IBFD-1' is considered, where UL pilots are unremittingly transmitted over $\bar{M}$ subcarriers and a 7-th order WP is applied for channel prediction. Note that, unlike the MDD-1 that only uses M subcarriers for DL transmission, as stated in Section VI-A, IBFD-1 uses all the M_{sum} subcarriers for the DL transmission, which imposes an extra SIC burden in digital-domain. According to [11], the digital-domain SIC in IBFD systems is expected to cancel the main linear signal component by more than 30 dB. By contrast, MDD systems are almost free from linear SI in the digital-domain owing to the FFT operation, which separates the UL subcarriers from the DL
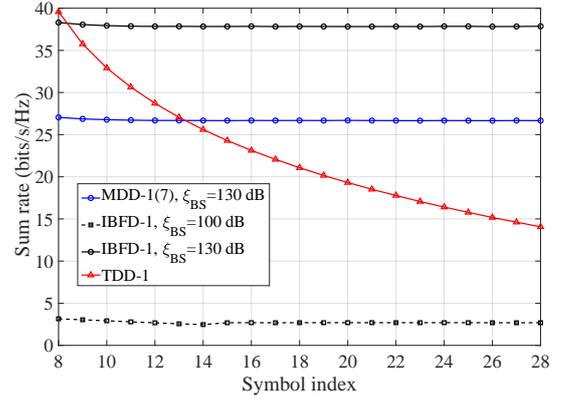
Fig. 11. Sum rate comparison of MDD, IBFD and TDD with respect to OFDM symbol index, when Type I frame structure and the relative velocity of 150 km/h are assumed.

subcarriers. Based on this observation, it is sensible to assume that with the same amount of system resources used for SI suppression, MDD can always achieve 30 dB more SIC than IBFD. From Fig. 11, it is shown that when both IBFD and MDD systems have the same SI mitigation capability[6], IBFD apparently outperforms both TDD and MDD. However, if IBFD can only mitigate the SI by 100 dB[7], the achievable sum rate of IBFD systems drops dramatically, and is much lower than that of MDD systems.

VII. CONCLUSIONS

A MDD-assisted mMIMO has been studied and comprehensively compared with the TDD-relied mMIMO over fast time-varying channels. Two types of MDD frame structures incorporated respectively with WP and DD-WP have been proposed to combat channel aging. For comparison, the frames in TDD mMIMO have been structured according to the 3GPP standards or improved for handling the channel aging problem. MDD has also been compared with the IBFD by considering their SIC capabilities. Moreover, the approximated lower bounds for the achievable rates of the proposed MDD and TDD frame structures have been derived. According to our studies and performance comparison, the following main observations can be obtained. First, when the Type I frame structure is employed, benefiting from the instantly available CSI updating while without involving UL/DL switching, MDD outperforms TDD over time-varying channels, which is more significant when mobility becomes higher. Second, the above performance advantage of MDD over TDD retains under the Type II frame structure, owing to the MDD's merits that it can proactively start DL transmission in Phase I, while simultaneously updating the CSI with the aid of the UL transmission in Phase II. Third, it is shown that SIC is critical to IBFD systems. By taking the advantage of near-free SI in the digital-domain, MDD is capable of achieving higher average sum

---

[6]This means IBFD costs more computation effort and system complexity than MDD for SIC, and hence achieves lower energy efficient [13].

[7]Note that 100 dB of SI is often assumed in IBFD systems, which is regarded as the bottleneck value based on the existing SIC approaches [34, 35].

$$\boldsymbol{\Gamma}_{d,i} = \mathbb{E}\left[\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL}}[i+1]\check{\boldsymbol{h}}_{n,d}^{\mathrm{UL},H}[i+1]\right] = \begin{bmatrix} \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,1]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,1]\right] & \cdots & \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,1]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,\bar{M}]\right] \\ \vdots & \ddots & \vdots \\ \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{M}]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,1]\right] & \cdots & \mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{M}]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,\bar{M}]\right] \end{bmatrix} \tag{51}$$

rates than IBFD within one frame in the presence of channel aging, if SIC in IBFD systems is imperfect. Additionally, the approximated lower bounded rates derived in this paper have been validated by the Monte-Carlo simulations, showing that they agree well with each other.

APPENDIX A
CALCULATION OF SI POWER

1) SI in (9)*:* According to the model for residual SI [26], the covariance of $z_d^{\mathrm{SI}}[i]$ is

$$\mathrm{cov}\left\{z_d^{\mathrm{SI}}[i]\right\} = \mathbb{E}\left[h_{\mathrm{SI}}x_d[i]x_d^H[i]h_{\mathrm{SI}}^H\right] \\ = \xi_{\mathrm{MT}}p_{\mathrm{UL}}\bar{M} \tag{48}$$

when using $\mathbb{E}\left[h_{\mathrm{SI}}h_{\mathrm{SI}}^H\right] = 1$. (48) is obtained by substituting $x_d[i] = \sqrt{p_{\mathrm{UL}}}\sum_{\bar{m}=1}^{\bar{M}}x_d[i,\bar{m}]$.

2) SI in (10)*:* To compute the covariance of $z^{\mathrm{SI}}[i]$ in (10), we again refer to the residual model in [26] and have

$$\mathrm{cov}\left\{\boldsymbol{z}^{\mathrm{SI}}[i]\right\} = \xi_{\mathrm{BS}}\mathrm{diag}\left(\mathbb{E}\left[\boldsymbol{H}_{\mathrm{SI}}\boldsymbol{s}_{\mathrm{DL}}[i]\boldsymbol{s}_{\mathrm{DL}}^H[i]\boldsymbol{H}_{\mathrm{SI}}^H\right]\right) \tag{49}$$

Upon substituting $\boldsymbol{s}_{\mathrm{DL}}[i] = \sum_{m=1}^{M}\sqrt{p_{\mathrm{DL}}}\boldsymbol{F}^{\mathrm{ZF}}[i,m]\boldsymbol{x}[i,m]$ into (49), the SI power at the n-th receive element can be obtained as

$$\mathrm{cov}\left\{\boldsymbol{z}^{\mathrm{SI}}[i]\right\}_{n,n} = \xi_{\mathrm{BS}}\mathbb{E}\left[\boldsymbol{H}_{\mathrm{SI}}^{(n,:)}\boldsymbol{s}_{\mathrm{DL}}[i]\boldsymbol{s}_{\mathrm{DL}}^H[i]\boldsymbol{H}_{\mathrm{SI}}^{(n,:),H}\right] \\ = \xi_{\mathrm{BS}}\mathbb{E}\left[\mathrm{Tr}\left(p_{\mathrm{DL}}\sum_{m=1}^{M}\boldsymbol{H}_{\mathrm{SI}}^{(n,:)}\boldsymbol{F}^{\mathrm{ZF}}[i,m]\boldsymbol{F}^{\mathrm{ZF},H}[i,m]\boldsymbol{H}_{\mathrm{SI}}^{(n,:),H}\right)\right] \\ \overset{(a)}{=} \xi_{\mathrm{BS}}p_{\mathrm{DL}}\sum_{m=1}^{M}\left\|\boldsymbol{F}^{\mathrm{ZF}}[i,m]\right\|_F^2 \\ \overset{(b)}{=} \xi_{\mathrm{BS}}p_{\mathrm{DL}}M \tag{50}$$

where (a) is obtained using $\mathbb{E}[\boldsymbol{H}_{\mathrm{SI}}^{(n,:),H}\boldsymbol{H}_{\mathrm{SI}}^{(n,:)}] = \boldsymbol{I}_N$ and (b) is due to the power normalization $\left\|\boldsymbol{F}^{\mathrm{ZF}}[i,m]\right\|_F^2 = 1$. Consequently, $\mathrm{cov}\left\{\boldsymbol{z}^{\mathrm{SI}}[i]\right\} = \xi_{\mathrm{BS}}p_{\mathrm{DL}}M\boldsymbol{I}_N$.

APPENDIX B
COVARIANCE MATRIX OF PREDICTED UL SUBCARRIER CHANNELS IN (33)

The covariance matrix $\boldsymbol{\Gamma}_{d,i}$ in (33) can be written as (51) on the top of this page. As for the diagonal elements in $\boldsymbol{\Gamma}_{d,i}$, according to (32), we have

$$h_{n,d}[i+1,\bar{m}] = \check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}] + \check{e}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}] \tag{52}$$

Hence, the variance of $\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}]$ is $(\boldsymbol{\Theta}_i[\bar{m}])_{d,d}$, given in (32). For the off-diagonal elements of $\boldsymbol{\Gamma}_{d,i}$, we have

$$\mathbb{E}\left[\check{h}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}_1]\check{h}_{n,d}^{\mathrm{DD},H}[i+1,\bar{m}_2]\right] \\ = \mathbb{E}\left[(h_{n,d}[i+1,\bar{m}_1] - \check{e}_{n,d}^{\mathrm{DD}}[i+1,\bar{m}_1])(h_{n,d}^H[i+1,\bar{m}_2] \\ - \check{e}_{n,d}^{\mathrm{DD},H}[i+1,\bar{m}_2])\right] \\ \overset{(a)}{=} \mathbb{E}\left[h_{n,d}[i+1,\bar{m}_1]h_{n,d}^H[i+1,\bar{m}_2]\right] \\ \overset{(b)}{=} \mathbb{E}\left[\boldsymbol{\psi}_{\bar{m}_1}\boldsymbol{g}_{n,d}[i+1]\boldsymbol{g}_{n,d}^H[i+1]\boldsymbol{\psi}_{\bar{m}_2}^H\right] \\ = \frac{\beta_d}{L}\boldsymbol{\psi}_{\bar{m}_1}\boldsymbol{\psi}_{\bar{m}_2}^H, \ \bar{m}_1 \neq \bar{m}_2 \tag{53}$$

where (a) is obtained based on the fact that the prediction error of the $\bar{m}_1$-th subcarrier channel is independent of the $\bar{m}_2$-th subcarrier channel and vice versa, while (b) can be derived from the FFT operation. In summary, the covariance matrix $\boldsymbol{\Gamma}_{d,i}$ can be expressed as

$$\boldsymbol{\Gamma}_{d,i} = \begin{bmatrix} (\boldsymbol{\Theta}_i[1])_{d,d} & \cdots & \frac{\beta_d}{L}\boldsymbol{\psi}_1\boldsymbol{\psi}_{\bar{M}}^H \\ \vdots & \ddots & \vdots \\ \frac{\beta_d}{L}\boldsymbol{\psi}_{\bar{M}}\boldsymbol{\psi}_1^H & \cdots & (\boldsymbol{\Theta}_i[\bar{M}])_{d,d} \end{bmatrix} \tag{54}$$

REFERENCES

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.

[2] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Comm. Mag.*, vol. 54, no. 2, pp. 114–123, 2016.

[3] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. on Comms.*, vol. 68, no. 1, pp. 232–257, 2019.

[4] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *J. of Comms and Net.*, vol. 15, no. 4, pp. 338–351, 2013.

[5] A. K. Papazafeiropoulos and T. Ratnarajah, "Deterministic equivalent performance analysis of time-varying massive MIMO systems," *IEEE Trans. on Wireless Comms.*, vol. 14, no. 10, pp. 5795–5809, 2015.

[6] S. Kashyap, C. Mollén, E. Björnson, and E. G. Larsson, "Performance analysis of (TDD) massive MIMO with Kalman channel prediction," in *Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, 2017, pp. 3554–3558.

[7] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Machine learning-based channel prediction in massive MIMO with channel aging," *IEEE Trans. on Wireless Comms.*, vol. 19, no. 5, pp. 2960–2973, 2020.

[8] J. Mirza, G. Zheng, K.-K. Wong, S. Lambotharan, and L. Hanzo, "On the performance of multiuser MIMO systems relying on full-duplex CSI acquisition," *IEEE Trans. on Comms.*, vol. 66, no. 10, pp. 4563–4577, 2018.

[9] R. Rajashekar, C. Xu, N. Ishikawa, L.-L. Yang, and L. Hanzo, "Multicarrier division duplex aided millimeter wave communications," *IEEE Access*, vol. 7, pp. 100 719–100 732, 2019.

[10] L.-L. Yang, *Multicarrier Communications*. John Wiley & Sons, 2009.

[11] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4. ACM, 2013, pp. 375–386.

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2022.3209799

15

[12] B. Li, L.-L. Yang, R. G. Maunder, and S. Sun, "Self-interference cancellation and channel estimation in multicarrier-division duplex systems with hybrid beamforming," *IEEE Access*, vol. 8, pp. 160 653–160 669, 2020.

[13] B. Li, L.-L. Yang, R. G. Maunder, P. Xiao, and S. Sun, "Multicarrier-division duplex: A duplexing technique for the shift to 6G wireless communications," *IEEE Veh. Tech. Mag.*, vol. 16, no. 4, pp. 57–67, 2021.

[14] B. Li, L.-L. Yang, R. G. Maunder, and S. Sun, "Resource allocation in millimeter-wave multicarrier-division duplex systems with hybrid beamforming," *IEEE Trans. on Veh. Tech.*, vol. 70, no. 8, pp. 7921–7935, 2021.

[15] B. Li, L.-L. Yang, R. G. Maunder, S. Sun, and P. Xiao, "Spectral-efficiency of cell-free massive mimo with multicarrier-division duplex," *arXiv preprint arXiv:2206.08774*, 2022.

[16] 3GPP, "Physical layer procedures (Release 15), TS 36.213."

[17] ——, "NR; physical channels and modulation (Release 15), TS 38.211."

[18] C. Nam, C. Joo, and S. Bahk, "Joint subcarrier assignment and power allocation in full-duplex OFDMA networks," *IEEE Trans. on Wireless Comms.*, vol. 14, no. 6, pp. 3108–3119, 2015.

[19] X. Li, X. Ge, X. Wang, J. Cheng, and V. C. Leung, "Energy efficiency optimization: Joint antenna-subcarrier-power allocation in OFDM-DASs," *IEEE Trans. on Wireless Comms.*, vol. 15, no. 11, pp. 7470–7483, 2016.

[20] Y. J. Zhang and K. B. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems," *IEEE Trans. on Wireless Comms.*, vol. 3, no. 5, pp. 1566–1575, 2004.

[21] L.-M. Chen and B.-S. Chen, "A robust adaptive dfe receiver for ds-cdma systems under multipath fading channels," *IEEE Trans. on Signal Proc.*, vol. 49, no. 7, pp. 1523–1532, 2001.

[22] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities." *IEEE J. on Select. Areas in Comms.*, vol. 32, no. 9, pp. 1637–1652, 2014.

[23] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Trans. on Microwave Theory and Techniques*, 2019.

[24] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. on Wireless Comms.*, vol. 4, no. 4, pp. 1650–1662, 2005.

[25] Z. Wei, W. Yuan, S. Li, J. Yuan, G. Bharatula, R. Hadani, and L. Hanzo, "Orthogonal time-frequency space modulation: A promising next-generation waveform," *IEEE Wireless Comms*, vol. 28, no. 4, pp. 136–144, 2021.

[26] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. on Select. Areas in Comms.*, vol. 30, no. 8, pp. 1541–1553, 2012.

[27] B. Debaillie, D.-J. van den Broek, C. Lavin, B. van Liempd, E. A. Klumperink, C. Palacios, J. Craninckx, B. Nauta, and A. Pärssinen, "Analog/RF solutions enabling compact full-duplex radios," *IEEE J. on Select. Areas in Comms*, vol. 32, no. 9, pp. 1662–1673, 2014.

[28] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap RF canceller for in-band full-duplex wireless communications," *IEEE Trans. on Wireless Comms.*, vol. 15, no. 6, pp. 4321–4334, 2016.

[29] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. on Broadcasting*, vol. 48, no. 3, pp. 223–229, 2002.

[30] S. S. Haykin, *Adaptive Filter Theory*. Pearson Education India, 2005.

[31] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time-varying channels for OFDM systems," *IEEE Trans. on Wireless Comms.*, vol. 4, no. 2, pp. 593–602, 2005.

[32] A. Khansefid and H. Minn, "Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination," *IEEE Trans. on Comms.*, vol. 63, no. 12, pp. 4849–4864, 2015.

[33] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 8), TS 36.211."

[34] B. Di, S. Bayat, L. Song, Y. Li, and Z. Han, "Joint user pairing, subchannel, and power allocation in full-duplex multi-user OFDMA networks," *IEEE Trans. on Wireless Comms.*, vol. 15, no. 12, pp. 8260–8272, 2016.

[35] D. W. K. Ng, Y. Wu, and R. Schober, "Power efficient resource allocation for full-duplex radio distributed antenna networks," *IEEE Trans. on Wireless Comms.*, vol. 15, no. 4, pp. 2896–2911, 2016.

Bohan-Li received M.Sc. degree from Beijing University of Posts and Telecommunications in 2018. He is currently pursuing the Ph.D. degree with the Next Generation Wireless Research Group, School of Electronics and Computer Science, University of Southampton. His research interests include full-duplex systems, multicarrier communications, massive MIMO and millimeter wave communications.

Lie-Liang Yang (Fellow, IEEE) is the professor of Wireless Communications at University of Southampton. He has research interest in wireless communications, wireless networks and signal processing for wireless communications, as well as molecular communications and nano-networks. He has published 400+ research papers in journals and conference proceedings and authored/co-authored three books, the details about which can be found at https://www.ecs.soton.ac.uk/people/llyang. He is a fellow of both IEEE and IET, and was a distinguished lecturer of IEEE VTS. He served as an associate editor to various journals, and is currently a senior editor to the IEEE Access and a subject editor to the Electronics Letters.

Robert G. Maunder (Senior Member, IEEE) is a professor in the School of Electronics and Computer Science at the University of Southampton. He is the founder and CTO of AccelerComm Ltd, Chilworth, SO16 7NS, U.K., which is commercializing his research as soft-Internet Protocol. He received his Ph.D. degree in telecommunications in 2007. His research interests include joint source/channel coding and the holistic design of algorithms and hardware implementations for wireless communications.

Songlin Sun (Senior Member, IEEE) is a professor in the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. He received his M.S. degree in signal and information processing from Shandong University of Technology in January 2000 and his Ph.D. degree in communication and information systems from the Beijing University of Posts and Telecommunications in 2003. His research interests include signal processing in wireless communication, video code, and multimedia transmission.

Pei Xiao (Senior Member, IEEE) is a professor of wireless communications at the Institute for Communication Systems in the University of Surrey. He is the technical manager of 5G and 6G Innovation Centers, leading the research team in the new physical-layer work area, and coordinating/supervising research activities across all the work areas within 5GIC (www.surrey.ac.uk/5gic/research). He has published extensively in the fields of communication theory, radio frequency, antenna design, and signal processing for wireless communications.