

On the Performance of Multiuser MIMO Systems Relying on Full-Duplex CSI Acquisition

Jawad Mirza, *Member, IEEE*, Gan Zheng, *Senior Member, IEEE*, Kai-Kit Wong, *Fellow, IEEE* Sangarapillai Lambotharan, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

Abstract—In this paper, we propose a combined full duplex (FD) and half duplex (HD) based transmission and channel acquisition model for open-loop multiuser multiple-input multiple-output (MIMO) systems. Assuming residual self interference (SI) at the BS, the idea is to utilize the FD mode during the uplink (UL) training phase in order to achieve simultaneous downlink (DL) data transmission and UL CSI acquisition. More specifically, the BS begins serving a user when its CSI becomes available, while at the same time, it also receives UL pilots from the next scheduled user. We investigate both zero-forcing (ZF) and maximum ratio transmission (MRT) MIMO beamforming techniques for the DL data transmission in the FD mode. The BS switches to the HD mode once it receives the CSI of all users and it employs ZF beamforming for the DL data transmission until the end of the transmission frame. Furthermore, we derive closed-form approximations for the lower bounded ergodic achievable rate relying on the proposed model. Our numerical results show that the proposed FD-HD transmission and channel acquisition approach outperforms its conventional HD counterpart and achieves higher data rates.

I. INTRODUCTION

Full duplex (FD) bi-directional communication systems have gained significant attention recently due to their potential to improve the capacity compared to the conventional half duplex (HD) systems [1], [2]. This improvement in FD systems is achieved by allowing the downlink (DL) and uplink (UL) channels to operate at the same time and frequency [3]–[7]. By contrast, at the time of writing cellular networks employ either time division duplex (TDD) or frequency division duplex (FDD), thus the radio resources are not fully exploited. Another attractive feature of FD radios is that they can also incorporate multiple-input multiple-output (MIMO) techniques which have become a core component of many wireless communication standards.

The major barrier in the implementation of the FD radio is the self interference (SI) imposed by the transmit antennas on

the receive antennas at the FD wireless transceiver. In recent years, many advanced SI cancellation techniques have been developed [3], [8]–[11], which paved the way for sophisticated studies to investigate the FD transmission in practical wireless communication systems. Insightful discussions on recent advances in SI cancellation techniques can be found in [12].

Most of the previous FD MIMO studies are dominated by MIMO-aided relaying [5], [13]–[16] and multiuser (MU) MIMOs [17]–[20]. The FD transmission philosophy is also well investigated in the cognitive radio context [21]–[23]. Recently, FD transmission has been widely investigated in MU systems relying on massive MIMO schemes [24], [25]. Moreover, FD communication has also been exploited for energy harvesting in MIMO systems [26], [27]. It is widely recognized that the performance of MU MIMO cellular DL transmissions heavily depend on the accuracy of channel state information (CSI) available at the base station (BS) [28]. Therefore, to equip the BS with CSI, classic TDD and FDD systems utilize either open-loop [29] or closed-loop [30] training methods, respectively. In closed-loop FDD training, the user quantizes the estimated DL channel and feeds back its quantized version in the UL to the BS. On the other hand, in open-loop TDD training, the BS acquires the CSI from UL training signals by exploiting the channel's reciprocity, even though in high-interference frequency-selective scenarios this has limited accuracy. Recently, simultaneous DL data transmission and UL CSI reception has been introduced at the BS [31]–[33] exploiting the FD capability of the BS. More specifically, the main objective in these studies is to exploit the additional channel created in the FD mode for receiving quantized UL CSI from the users, while at the same time transmitting the DL payload to those users whose CSI has already become available at the BS. Once the BS has obtained the quantized CSI from all the users in the system, it can switch to the classic HD mode for the rest of the transmission duration. Note that ideally the transmission frame length should be adapted to the coherence time and coherence bandwidth of the channel [34]. Therefore, in the classic HD system, a transmission frame typically consists of separate training and payload transmission stages [28], whilst in FD systems, it is possible to accomplish DL data transmission and UL CSI acquisition at the same time [31]. As reported in [31]–[33], [35], this approach improves the achievable rate of the system compared to the classic HD transmission scheme.

In [32], a continuously adaptive beamforming scheme is presented, where simultaneous DL data transmission and quantized UL CSI feedback reception is realized using FD radios

This work is supported in part by EPSRC under grant EP/N008219/1.

J. Mirza, G. Zheng and S. Lambotharan are with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Leicestershire, LE11 3TU, UK (Email: {j.mirza, g.zheng, S.Lambotharan}@lboro.ac.uk). The work of G. Zheng was supported by the UK EPSRC under grant number EP/N007840/1, and the Leverhulme Trust Research Project Grant under grant number RPG-2017-129.

J. Mirza is also with the Department of Electrical Engineering, COMSATS Institute of Information Technology, Islamabad Campus, Islamabad, Pakistan.

K.-K. Wong is with the Department of Electronic and Electrical Engineering, University College London, London, WC1E 6BT, UK (Email: kai-kit.wong@ucl.ac.uk).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (Email: lh@ecs.soton.ac.uk). L. Hanzo would like to gratefully acknowledge the ERC's financial support of his Advanced Fellow Grant.

at both the BS and the users. In the FD mode, the DL data transmission is only commenced for those users whose CSI has already been obtained at the BS, while at the same time receiving quantized CSI from the next scheduled user. Du *et al.* [32] conceived a closed-loop digital feedback technique using random vector quantization codebooks, which provides a beneficial multiplexing gain over the HD transmission scheme at low feedback powers. Then Du *et al.* further extended this [33] by exploiting the channel's reciprocity for open-loop training to demonstrate its advantages over the HD scheme. Approximations of the optimal feedback schemes are also derived in [33] by formulating an optimization problem for maximizing the achievable rate. Similar to [32], [33], sequential training based beamforming was introduced by Du *et al.* in [31] for both open-loop and closed-loop training methods. Another closed-loop training based MU MIMO system was studied by Zhang *et al.* in [35], where the main objective was to improve the energy efficiency using FD transmission. More explicitly, Zhang *et al.* [35] have proposed to maximize the energy efficiency of the FD MU MIMO system by optimizing the partitioning of the UL and DL training phases. However, all the above seminal studies discussed so far have a common limitation, namely that they do not consider the deleterious effects of residual SI in the analysis, even though there is always some residual SI in practical FD systems [36], [37]. To provide users with DL CSI, an accurate CSI acquisition method was proposed in [38] by Arnau and Kountouris for SISO links, where the residual SI is also present in the system, where the BS continuously transmits a pilot sequence using the same frequency band as that dedicated for the UL. It has been reported that the scheme proposed in [38] outperforms its HD counterpart by yielding a 50% throughput improvement.

Although the additional inter-node interference [31] and residual SI [38] still degrade the performance of FD schemes, these techniques tend to outperform the HD scheme in terms of the achievable rate. In the seminal studies by Du *et al.* [31], [32] the DL channel is assumed to be perfectly known at the user. However, in real systems, the user either has no CSI or acquires imperfect CSI via downlink training and again, most of the existing studies [31], [32], [35] on the achievable rate simplify the analysis by neglecting the effects of residual SI. To the best of our knowledge, the dedicated training [28] has not been previously considered in FD based CSI acquisition. Finally, only the zero-forcing (ZF) beamforming scheme is considered for the FD mode in the majority of the existing studies.

Against this background, we propose a DL transmission and UL CSI attainment estimation model for MU MIMO systems, where the BS is equipped with FD capabilities. Unlike the previous studies [31]–[33], [35], we investigate the FD-HD model in the presence of realistic residual SI. An FD assisted BS is utilized for implementing the DL data transmission concurrently with the UL pilot reception using an open-loop training scheme. Hence, we rely on the UL pilot symbols received from the users to provide the BS with the CSI knowledge. In this study, the single-antenna users utilize the conventional HD transmission scheme. We consider the scenario, where no CSI is available at the user, who only has a

statistical characterization of the effective channel information [39] obtained using the specific approximations derived in this paper. In contrast to [31], we consider dedicated training, in order to enable users to coherently detect the signal transmitted in the HD mode. Unlike [31], [32], we comparatively analyze the performance of both ZF and maximum ratio transmission (MRT) beamforming schemes under the FD mode. For the HD mode, we prefer using the ZF beamforming scheme as a benefit of its superiority over the MRT beamforming scheme at moderate to high signal-to-noise ratios (SNRs). By contrast, MRT beamforming is more effective at the low SNRs, despite its lower complexity. In the FD mode, the BS performs simultaneous DL data transmission and UL CSI feedback reception, resulting in SI¹. However, once all the users have fed back their CSI, the BS switches to the HD mode, where only DL data transmission is performed until the end of the transmission frame. Like [28], [38], we also assume that the channel's envelope remains constant over the given transmission frame. We assume that the users share the UL feedback channel in a time division multiplexing access (TDMA) manner for the UL training phase. In a nutshell, the main contributions of the paper are summarized below:

- Using an FD transceiver at the BS, we propose an integrated FD-HD model for a concurrent DL transmission and UL CSI acquisition. In contrast to [31]–[33], [35], realistic residual SI is present at the BS. We also assume that the users operate in the HD mode. We consider open-loop training for CSI acquisition at the BS.
- In the FD assisted CSI acquisition and DL data transmission phase, we propose to use the expected effective channel gain [39]–[41] for the coherent detection of the signal, which helps in significantly reducing the duration of the training period. As the duration of HD DL data transmission is typically higher than the training duration [28], [34], high-quality detection is important. We exploit the instantaneous effective channel gain [28], [42] to support coherent detection in the HD mode. This requires a dedicated training phase before the HD DL data transmission, which has not been considered in previous studies relying on the FD transmission mode.
- We derive closed-form approximations for the lower bounded achievable rates of the proposed scheme. The analysis is performed for both ZF and MRT linear beamforming techniques in the FD mode. We also derive the lower bounded achievable rate approximation for the DL data transmission in the HD mode using the ZF beamforming scheme, which matches the lower bounded achievable rate more closely than the bounded rate derived in [28].

In this study, we specifically focus our attention on TDD-based cellular networks using the FD mode for improving the link capacity. Numerous related studies [43]–[45] have revealed that the use of the FD operation provides a significant throughput gain in LTE-based cellular networks. It can be

¹In this work, we do not consider specific SI cancellation techniques and assume that the SI is suppressed such that there is some residual SI present at the BS.

observed from the numerical results that the ZF beamforming scheme relying on the proposed transmission model has a superior rate both over the MRT beamforming and over the conventional HD schemes. However, the MRT beamforming scheme strikes the best trade-off in terms of the rate versus computational complexity.

The rest of the paper is organized as follows. Section II presents our system model and the proposed DL transmission and UL CSI acquisition model. In Section III, we provide the derivation of the achievable ergodic rate in the FD mode. This is followed by the derivation of the achievable ergodic rate of the HD mode in Section IV. Our numerical results are presented in Section V. Finally, we conclude in Section VI.

Notations: We use $(\cdot)^H$, $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^{-1}$ and $(\cdot)^\perp$ to denote the conjugate transpose, the conjugate, the transpose, the inverse and the pseudoinverse operations, respectively. $\|\cdot\|$ and $|\cdot|$ stand for vector and scalar norms respectively. $\mathbb{E}[\cdot]$ denotes expectation. $\mathbf{A}(:,n)$ represents the n^{th} column of the \mathbf{A} . The complex normal distribution with mean μ and variance σ^2 is denoted by $\mathcal{CN}(\mu, \sigma^2)$.

II. SYSTEM MODEL AND THE PROPOSED CHANNEL ACQUISITION AND TRANSMISSION

We consider an MU MIMO DL system consisting of a FD BS having M antennas, which are used for both transmission and reception [31], [46]. The BS serves K HD single-antenna users employing a linear beamforming technique, where $K \leq M$. Similar to [31], [32], we also assume that the UL training and DL transmission use the same time- and frequency-resource. Furthermore, we consider perfectly reciprocal UL-DL channels, therefore the BS uses the DL channel estimate using the UL pilot signals received from the users. Although various SI cancellation schemes are advocated in [3], [8]–[11] for suppressing the SI within the transceiver, it is not possible to completely eliminate the SI due to unavoidable transceiver non-linearities as well as owing to the IQ mixer amplitude and phase imbalances [11]. Much work has been dedicated to the modeling of the residual SI, which has shown that the residual SI channel undergoes Rician fading. However, in [4], it has been reported that the K -factor of the residual SI channel reduces after the active cancellation process, hence making the Rayleigh distribution also a suitable model.

The main idea in this study is to exploit the additional channel created in the FD mode for achieving simultaneous data transmission and CSI acquisition in DL and UL, respectively. We assume that UL training, DL training and DL data transmission are carried out within the duration of the frame length (or the total number of channel uses), denoted by T . In this study, we assume a block fading channel model, where the channel remains constant over the coherence interval of T . We denote the duration of FD DL data transmission and UL/DL training by T_t . T_d denotes the duration of DL data transmission in the HD mode, so that $T = T_t + T_d$. The user transmits β_{ul} uplink training pilots (i.e., β_{ul} channel uses) to the BS for CSI estimation. Whereas, β_{dt} represents the number of pilots transmitted by the BS in the dedicated training stage.

The proposed transmission model can be partitioned into three phases: i) FD based simultaneous DL data transmission

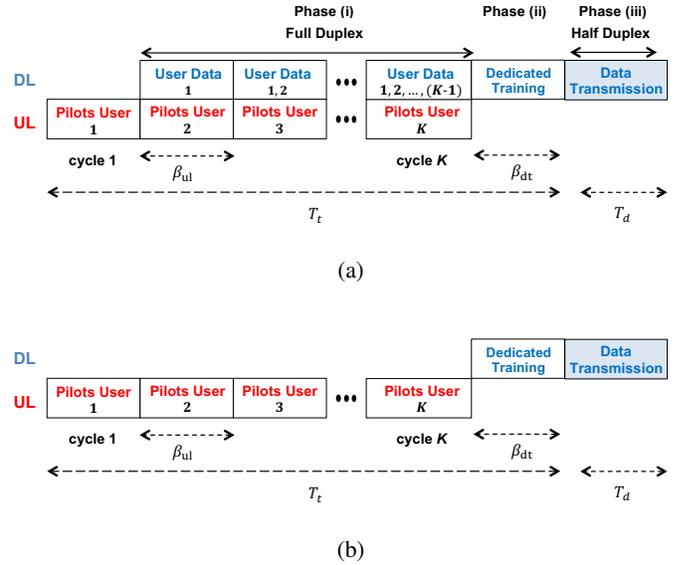


Fig. 1: (a) The proposed UL training and DL data transmission model where users share the UL channel using TDMA. (b) The corresponding HD model. Typically, $T_d \gg T_r$, the durations presented here are for illustration purposes only.

and UL training, ii) dedicated training and iii) HD based DL data transmission, as shown in Fig. 1a. Each single antenna user transmits β_{ul} UL pilot symbols to the BS in order to equip the BS with the estimate of the CSI. As shown in Fig. 1a, while receiving the CSI of the next scheduled user, the BS begins transmitting data symbols to the previously scheduled user(s) using β_{ul} channel uses. It becomes clear from Fig. 1a that the higher the β_{ul} values, the more channel uses will be required for FD transmission. In the HD mode, the CSI of all the users is available at the BS, therefore, it is invoked for all the K users in the system relying on the ZF beamforming scheme. Note that switching between the HD and FD modes or hybrid HD and FD modes has also been considered in many other studies [1], [15], [31]. The traditional HD transmission model is shown in Fig. 1b, where simultaneous DL transmission and UL CSI acquisition is not possible, because the BS does not support FD communication. In the following subsections, we discuss each phase of the proposed model in detail.

A. FD based Simultaneous DL Data Transmission and UL Training

Again, in the first phase, each user sends its UL training signal to the BS as also shown in Fig. 1a. A TDMA technique is employed for the users to access the UL channel. At this stage, the BS takes advantage of the FD mode and starts transmitting its coded data symbols using a linear beamforming scheme to those users who have previously sent their pilot signals to the BS for CSI estimation, while it continues receiving UL pilot symbols from the remaining users in a concurrent manner. In this study, we analyze the performance of the FD mode both with ZF and MRT beamforming schemes. This phase can be divided into two categories detailed in the following subsections.

1) *UL Training*: As shown in Fig. 1a, the first scheduled user transmits β_{ul} UL pilot symbols to the BS over the UL channel. In this phase, we refer to each block of β_{ul} channel uses as a cycle, so that there are a total of K cycles.

The user transmits its pilot sequence to the BS consisting of β_{ul} symbols. We Denote the transmitted pilot sequence by $\mathbf{s} \in \mathbb{C}^{\beta_{\text{ul}} \times 1}$, which is assumed to be a unit-power vector. The signals received at the BS from the k^{th} user is written as [47]

$$\mathbf{Y}_k = \sqrt{\beta_{\text{ul}} p_k} \mathbf{h}_k \mathbf{s}^T + \mathbf{\Psi} + \mathbf{N}, \quad (1)$$

where p_k is the UL transmit power of the k^{th} user and \mathbf{h}_k is an $M \times 1$ channel vector between the k^{th} user and BS, whose entries are distributed according to the uncorrelated complex-valued Gaussian distribution having a zero mean and unit variance, so that $\mathbf{h}_k \sim \mathcal{CN}(0, \mathbf{I})$. The matrix $\mathbf{N} \in \mathbb{C}^{M \times \beta_{\text{ul}}}$ is the noise matrix whose elements are assumed to be distributed according to $\mathcal{CN}(0, 1)$. The residual SI during the β_{ul} channel uses is captured in the matrix $\mathbf{\Psi} = [\psi_1, \dots, \psi_{\beta_{\text{ul}}}]$, where $\psi_n = \mathbf{H}_{\text{SI}} \hat{\mathbf{V}} \mathbf{u}_n$ is the residual SI, when the n^{th} pilot symbol is transmitted. The precoding matrix is $\hat{\mathbf{V}} \in \mathbb{C}^{M \times U}$, where U denotes the number of users that are being served by the BS, such that $U = k - 1$. The vector $\mathbf{u}_n \in \mathbb{C}^U$ consists of the data symbols from the U independently generated Gaussian codewords, such that $\mathbf{u}_n \sim \mathcal{CN}(0, \frac{P}{U} \mathbf{I})$. \mathbf{H}_{SI} is an $M \times M$ matrix representing the residual SI at the BS. As reported in [4], [9], [38], the strongest self-interfering paths are effectively attenuated after the SI cancellation. Therefore, like in [38], we also assume that $\mathbf{H}_{\text{SI}} \sim \mathcal{CN}(0, \nu^2)$. According to the proposed model, during the UL training of the k^{th} user both \mathbf{H}_{SI} and $\hat{\mathbf{V}}$ remain constant. However, the coded data symbols of the user change. The BS carries out the de-spreading of \mathbf{Y}_k , given by

$$\mathbf{r}_k = \mathbf{Y}_k \mathbf{s}^* = \sqrt{\beta_{\text{ul}} p_k} \mathbf{h}_k + \tilde{\boldsymbol{\psi}} + \tilde{\mathbf{n}}_k, \quad (2)$$

where $\tilde{\mathbf{n}}_k = \mathbf{N} \mathbf{s}^* \in \mathcal{CN}(0, N_0)$. The vector $\tilde{\boldsymbol{\psi}}$ is given by $\tilde{\boldsymbol{\psi}} = \mathbf{\Psi} \mathbf{s}^*$. Note that the BS experiences no SI (i.e., also no residual SI) while receiving the pilot symbols from the first scheduled user in the system (i.e., $\tilde{\boldsymbol{\psi}} = 0$ for the first scheduled user).

2) *DL Data Transmission*: After estimating the channel of the first user from the received pilots, the BS starts transmitting the coded data symbols to the first user in the next cycle, while at the same time it also receives the UL pilots from the next scheduled user. Thus, SI is generated at the BS as a result of employing the FD transmission mode. Again, we assume that the BS mitigates the SI by employing both analog and digital cancellation techniques,² so that only reduced residual SI is imposed on the system. During the DL data transmission stage, the received signal by user k (i.e., a single channel use) in a given cycle can be written as

$$y_k^{\text{FD}} = \mathbf{h}_k^H \mathbf{x} + h_{q,k}^{\text{ul}} s_q + n_k, \quad (3)$$

where \mathbf{h}_k is an $M \times 1$ channel vector between the BS and the k^{th} user, and $\mathbf{h}_k \sim \mathcal{CN}(0, \mathbf{I})$. We assume perfect power-control, which eliminates the path loss. The noise at the

k^{th} user is denoted by n_k , which is distributed according to the uncorrelated complex-valued Gaussian distribution with zero mean and variance N_0 , so that $n_k \sim \mathcal{CN}(0, N_0 \mathbf{I})$. The data signal transmitted from the BS is given by $\mathbf{x} = \hat{\mathbf{V}} \mathbf{u}$, where $\hat{\mathbf{V}} \in \mathbb{C}^{M \times j}$ is a linear beamforming matrix and j is the total number of users served by the BS in the given cycle. The vector $\mathbf{u} \in \mathbb{C}^j$ consists of the symbols from the j independently generated Gaussian codewords obeying $\mathbf{u} \sim \mathcal{CN}(0, \frac{P}{j} \mathbf{I})$. Here, we assume that when the k^{th} user receives the DL signal, user q (i.e., the next scheduled user) is transmitting its UL pilot to the BS at the same time, thus, creating inter-node interference at the k^{th} user. The channel between the q^{th} user and the k^{th} user is denoted by $h_{q,k}^{\text{ul}}$, where $h_{q,k}^{\text{ul}} \sim \mathcal{CN}(0, 1)$. s_q is the pilot symbol sent by the q^{th} user to the BS, where $|s_q|^2 = 1/\beta_{\text{ul}}$. The transmitted signal is subject to the average power constraint of $\mathbb{E}[|\mathbf{x}|^2] \leq P$. The transmitted power of the q^{th} user is given by p_q . We denote the imperfect CSI at the BS in the cycle c , by $\hat{\mathbf{H}}_c = [\hat{\mathbf{h}}_1, \dots, \hat{\mathbf{h}}_j]$, where $\hat{\mathbf{h}}_k$ denotes the imperfect CSI of the k^{th} user and $j = c - 1$. As discussed earlier, in this study, we investigate the performance of both MRT and ZF beamforming schemes in the FD mode. The ZF beamforming vector of the k^{th} user in cycle c is denoted by $\hat{\mathbf{v}}_{k,c}^{\text{ZF}}$, which is the k^{th} normalized column of the matrix $\hat{\mathbf{V}}_c$, where $\hat{\mathbf{V}}_c = \hat{\mathbf{H}}_c^H (\hat{\mathbf{H}}_c \hat{\mathbf{H}}_c^H)^{-1}$. The MRT beamforming vector of the k^{th} user is given by $\hat{\mathbf{v}}_k^{\text{MRT}} = \hat{\mathbf{h}}_k / \|\hat{\mathbf{h}}_k\|$. Note that the MRT beamforming vector of the k^{th} user remains the same in each cycle. We assume that the k^{th} user only has the knowledge of the expected effective channel gains $\mathbb{E}[a_{k,k}^{\text{ZF}}]$ or $\mathbb{E}[a_{k,k}^{\text{MRT}}]$ of ZF or MRT beamforming schemes, respectively, where $a_{k,k}^{\text{ZF}} = \mathbf{h}_k^H \hat{\mathbf{v}}_{k,c}^{\text{ZF}}$ and $a_{k,k}^{\text{MRT}} = \mathbf{h}_k^H \hat{\mathbf{v}}_k^{\text{MRT}}$. These expected channel gains are obtained at the users using the approximations derived in Section III, and are utilized by the users for the coherent detection of transmitted signals [42], [48].

B. Dedicated Training

Once the BS receives the CSI of all the K users, we propose to use the ZF beamforming scheme for DL data transmission, that completely eliminates the interference in the system, in the presence of perfect CSI at the BS. It is also shown to be an effective technique under imperfect CSI in MIMO broadcast channels [49]. The advantage of the ZF beamforming scheme over the MRT beamforming scheme is that it yields a superior performance at moderate to high SNRs. Typically, $T_d \gg T_t$ holds, therefore it is beneficial for users to know the instantaneous effective channel gain at this stage for coherent detection. For this purpose, additional training is required before invoking the ZF beamforming scheme, which is termed as dedicated training [28]. We denote the duration of the DL pilot sequence transmitted from the BS by β_{dt} , which is expressed in terms of the number of channel uses. Following a similar approach as in [42], the BS transmits the pilots using the ZF beamforming matrix, $\hat{\mathbf{V}}$. Then user k applies the orthogonal complement of the pilot matrix to the received pilot matrix in order to obtain the received pilot vector, given by

$$\bar{\mathbf{y}}_k = \sqrt{\beta_{\text{dt}} P} \mathbf{a}_k + \mathbf{z}_k, \quad (4)$$

²The SI cancelling or mitigating techniques are out of the scope of this paper, and therefore, in this paper, we only focus on the residual SI.

where $\mathbf{a}_k = [a_{k,1}^{\text{ZF}} \ a_{k,2}^{\text{ZF}} \ \cdots \ a_{k,K}^{\text{ZF}}]$ with $a_{k,i}^{\text{ZF}} = \mathbf{h}_k^H \mathbf{v}_i^{\text{ZF}}$, $i = 1, \dots, K$ and \mathbf{z}_k has independent and identically distributed (i.i.d.) $\mathcal{CN}(0, N_0)$ elements. Owing to lack of space, we refer the reader to [42], which provides the details of the presented pilot-based training approach. Then user k estimates its channel coefficient using the MMSE estimate on the k^{th} entry of $\bar{\mathbf{y}}_k$, given by $\bar{y}_{k,k} = \sqrt{\beta_{\text{dt}} P} a_{k,k}^{\text{ZF}} + z_{k,k}$, yielding [28]

$$\hat{a}_{k,k}^{\text{ZF}} = \frac{\sqrt{\beta_{\text{dt}} P}}{\beta_{\text{dt}} P + N_0} \bar{y}_{k,k}. \quad (5)$$

Then user k utilizes these estimated values of $\hat{a}_{k,i}^{\text{ZF}}$, $\forall i$ to decode the data [28]. We can also express $a_{k,k}^{\text{ZF}}$ as $a_{k,k}^{\text{ZF}} = \hat{a}_{k,k}^{\text{ZF}} + \lambda_k$, where λ_k is the error and its variance of λ_k is given by $\sigma_{\lambda_k}^2 = N_0 / (\beta_{\text{dt}} P)$. Similarly, the variance of $\hat{a}_{k,k}$ is given by $1 - \sigma_{\lambda_k}^2$. Note that there is no SI at this stage, since the BS is only transmitting and not receiving any signal.

C. HD based DL Data Transmission

Following the DL dedicated training phase, the BS transmits the coded data symbols denoted by u_1, \dots, u_K for all the K users in the system using the ZF beamforming scheme until the end of the frame duration, T . In this phase, we consider the conventional HD transmission mode (i.e., no SI). The signal received by the k^{th} user can be expressed as $y_k^{\text{HD}} = \mathbf{h}_k^H \mathbf{x} + n_k$ and it utilizes the estimated instantaneous effective channel gain in (5) for the coherent detection of the transmitted signal.

III. ACHIEVABLE ERGODIC RATE IN FD MODE

In this section, we derive ergodic rates attained by the proposed DL transmission and UL CSI acquisition model presented in Section II. Here, we assume that the proposed strategy is implemented over a long sequence of frames where each frame has a similar structure as shown in Fig. 1a. Therefore, the ergodic achievable rate (in bits per channel use) of the proposed model consisting of both FD and HD modes can be written as

$$R_k = \underbrace{\frac{\beta_{\text{ul}}}{T} \sum_{c=2}^K \frac{1}{K} \sum_{k=1}^{c-1} R_{k,c}^{\text{FD}}}_{R_k^{\text{FD}}} + \underbrace{\left(1 - \frac{T_t}{T}\right) R_k^{\text{ZF-HD}}}_{R_k^{\text{HD}}}, \quad (6)$$

where $R_{k,c}^{\text{FD}}$ denotes the rate of the k^{th} user in the FD mode when $j = c - 1$ users are served by the BS in the c^{th} cycle, while the BS is waiting for the UL training symbols arriving from the remaining $K - j$ users in the system. $R_k^{\text{ZF-HD}}$ represents the rate when the channels of all the K users are estimated at the BS and it serves the users employing ZF beamforming (after the dedicated training phase) in the HD mode. For fairness among users, the user indices $1, 2, \dots, K$ are randomly assigned so that the probability of any user to become the k^{th} user is $1/K$. The analysis of the ergodic achievable rate in the FD mode for a user is provided in the following two subsections for both the MRT and ZF beamforming techniques.

A. MRT Beamforming in the FD Mode

In this subsection, we derive the lower bounded achievable ergodic rate approximation in the FD mode (i.e., the first term in (6)) for the proposed model using MRT beamforming. The residual SI is also present at this stage. Similar to [28], [42], [47], [50], the BS estimates \mathbf{h}_k by applying the MMSE criterion to the observation \mathbf{r}_k , given as

$$\hat{\mathbf{h}}_k = \mathbb{E} [\mathbf{h}_k \mathbf{r}_k^H] \mathbb{E} [\mathbf{r}_k \mathbf{r}_k^H]^{-1} \mathbf{r}_k \quad (7)$$

$$= \frac{\sqrt{\beta_{\text{ul}} p_k}}{(\beta_{\text{ul}} p_k + N_0 + P\nu^2)} \mathbf{r}_k. \quad (8)$$

The details of the above derivation are provided in Appendix I. We can write the actual channel of the k^{th} user as

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{e}_k, \quad (9)$$

where the estimation error vector, \mathbf{e}_k , is independent of $\hat{\mathbf{h}}_k$ and has a covariance of $\sigma_{\mathbf{e}_k}^2 \mathbf{I} = \mathbb{E} [\mathbf{e}_k \mathbf{e}_k^H]$, where

$$\sigma_{\mathbf{e}_k}^2 = 1 - \frac{\beta_{\text{ul}} p_k}{(\beta_{\text{ul}} p_k + N_0 + P\nu^2)}. \quad (10)$$

We use a lower bound on the average achievable rate derived in [39], [48], where the expected effective channel gain is known, but no instantaneous channel knowledge is available at the user. Using the results of [39], [48], we can write the lower bounded achievable rate for the k^{th} user given the knowledge of the expected effective channel gain $\mathbb{E} [a_{k,k}^{\text{MRT}}] = \mathbb{E} [\mathbf{h}_k^H \hat{\mathbf{v}}_k^{\text{MRT}}]$, where j users are served by the BS, as

$$R_{k,j}^{\text{MRT-FD}} \geq R_{k,j}^{\text{MRT-FD-LB}} = \quad (11)$$

$$\log_2 \left(1 + \frac{\frac{P}{jN_0} |\mathbb{E} [a_{k,k}^{\text{MRT}}]|^2}{1 + \frac{P}{jN_0} \left(\text{var} \{a_{k,k}^{\text{MRT}}\} + \sum_{i \neq k} \mathbb{E} [|a_{k,i}^{\text{MRT}}|^2] \right) + \frac{P\gamma_q}{\beta_{\text{ul}} N_0}} \right),$$

where $\gamma_q = \mathbb{E} [|h_{q,k}^{\text{ul}}|^2]$ and $a_{k,i}^{\text{MRT}} = \mathbf{h}_k^H \hat{\mathbf{v}}_i^{\text{MRT}}$. The expectation of the effective channel gain $a_{k,k}^{\text{MRT}}$ can be approximated by

$$\mathbb{E} [a_{k,k}^{\text{MRT}}] \approx \sqrt{(\beta_{\text{ul}} p_k M) (\beta_{\text{ul}} p_k + N_0 + P\nu^2)^{-1}}, \quad (12)$$

and its proof is given in Appendix II. We can approximate the variance of $a_{k,k}^{\text{MRT}}$ as

$$\text{var} \{a_{k,k}^{\text{MRT}}\} \approx \delta + \sigma_{\mathbf{h}_k}^2 M \delta \left(\delta - \frac{M+1}{M} \right), \quad (13)$$

where $\delta = \frac{M\Gamma^2(M)}{\Gamma^2(M+(1/2))}$ and $\sigma_{\mathbf{h}_k}^2 = (\beta_{\text{ul}} p_k) (\beta_{\text{ul}} p_k + N_0 + P\nu^2)^{-1}$. The proof of (13) is given in Appendix III. Next, we compute the quantity $\mathbb{E} [|\mathbf{h}_k^H \hat{\mathbf{v}}_i^{\text{MRT}}|^2]$, using (9), we have

$$\begin{aligned} \mathbb{E} [|\hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_i^{\text{MRT}} + \mathbf{e}_k^H \hat{\mathbf{v}}_i^{\text{MRT}}|^2] &\stackrel{(a)}{\leq} \mathbb{E} [|\hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_i^{\text{MRT}}|^2] + \mathbb{E} [|\mathbf{e}_k^H \hat{\mathbf{v}}_i^{\text{MRT}}|^2] \\ &= \mathbb{E} [\|\hat{\mathbf{h}}_k^H\|^2] \mathbb{E} \left[\left| \frac{\hat{\mathbf{h}}_k^H}{\|\hat{\mathbf{h}}_k^H\|} \hat{\mathbf{v}}_i^{\text{MRT}} \right|^2 \right] + \mathbb{E} [\hat{\mathbf{v}}_i^{\text{MRT}H} \mathbb{E} [\mathbf{e}_k \mathbf{e}_k^H] \hat{\mathbf{v}}_i^{\text{MRT}}] \\ &\stackrel{(b)}{=} \frac{\beta_{\text{ul}} p_k}{(\beta_{\text{ul}} p_k + N_0 + P\nu^2)} + \sigma_{\mathbf{e}_k}^2 = 1, \end{aligned} \quad (14)$$

where (a) follows by ignoring the expectation of cross terms and (b) comes from the fact that $(\hat{\mathbf{h}}_k^H / \|\hat{\mathbf{h}}_k^H\|)$ and $\hat{\mathbf{v}}_i^{\text{MRT}}$ are independent unit vectors, yielding $\mathbb{E}[\|(\hat{\mathbf{h}}_k^H / \|\hat{\mathbf{h}}_k^H\|)\hat{\mathbf{v}}_i^{\text{MRT}}\|^2] = 1$. Substituting (12), (13) and (14) in (11), as well as using $\gamma_q = \mathbb{E}[|h_{q,k}^{\text{ul}}|^2] = 1$, gives the approximation of the ergodic achievable rate of MRT beamforming relying on analog feedback, which can be expressed as

$$R_{k,j}^{\text{MRT-FD-LB}} \approx \log_2 \left(1 + \frac{\frac{P}{jN_0} (\sigma_{\mathbf{h}_k}^2 M)}{1 + \frac{P}{jN_0} \left[\left(\delta + \sigma_{\mathbf{h}_k}^2 M \delta \left(\delta - \frac{M+1}{M} \right) + j - 1 \right) + \frac{P_q}{\beta_{\text{ul}} N_0} \right]} \right). \quad (15)$$

We can write (15) in terms of the number of cycles, $c = j + 1$, given in (6), after some re-arrangement as

$$R_{k,c}^{\text{MRT-FD-LB}} \approx \log_2 \left(1 + \frac{\frac{P}{N_0} (\sigma_{\mathbf{h}_k}^2 M)}{(c-1)\eta + \frac{P}{N_0} \left[\left(\delta + \sigma_{\mathbf{h}_k}^2 M \delta \left(\delta - \frac{M+1}{M} \right) - 1 \right) \right]} \right), \quad (16)$$

where $\eta = (1 + \frac{P}{N_0} + \frac{P_q}{\beta_{\text{ul}} N_0})$. Now we can write the rates of the different cycles from (6) as

$$\sum_{c=2}^K \sum_{k=1}^{c-1} R_{k,c}^{\text{MRT-FD-LB}} = \underbrace{R_{1,2}^{\text{MRT-FD-LB}}}_{\text{cycle 2}} + \underbrace{R_{1,3}^{\text{MRT-FD-LB}} + R_{2,3}^{\text{MRT-FD-LB}}}_{\text{cycle 3}} + \dots + \underbrace{R_{1,K}^{\text{MRT-FD-LB}} + \dots + R_{K-1,K}^{\text{MRT-FD-LB}}}_{\text{cycle K}}. \quad (17)$$

Using (16), we can express (17) as

$$\sum_{c=2}^K \sum_{k=1}^{c-1} R_{k,c}^{\text{MRT-FD-LB}} \approx R_{1,2}^{\text{MRT-FD-LB}} + \sum_{c=3}^K \log_2 \left(1 + \frac{\frac{P}{N_0} (\sigma_{\mathbf{h}_k}^2 M)}{(c-1)\eta + \frac{P}{N_0} \left[\left(\delta + \sigma_{\mathbf{h}_k}^2 M \delta \left(\delta - \frac{M+1}{M} \right) - 1 \right) \right]} \right), \quad (18)$$

where the BS experiences no SI for the first scheduled user in cycle 2. Therefore, the approximated lower-bounded rate of this user is denoted by $R_{1,2}^{\text{MRT-FD-LB}}$, which is given by (16) with $P\nu^2 = 0$. Note that the lower-bounded ergodic achievable rate of the proposed model in the FD mode with MRT beamforming is given by

$$\bar{R}_k^{\text{MRT-FD-LB}} = \frac{\beta_{\text{ul}}}{T} \sum_{c=2}^K \frac{1}{K} \sum_{k=1}^{c-1} R_{k,c}^{\text{MRT-FD-LB}}. \quad (19)$$

Substituting (18) in (19), the lower-bounded ergodic achievable rate approximation for the proposed model in the FD mode (i.e., the first term in (6)), is given by

$$\bar{R}_k^{\text{MRT-FD-LB}} = \frac{\beta_{\text{ul}}}{KT} \left[R_{1,2}^{\text{MRT-FD-LB}} + \sum_{c=3}^K \log_2 \left(1 + \frac{\frac{P}{N_0} (\sigma_{\mathbf{h}_k}^2 M)}{(c-1)\eta + \frac{P}{N_0} \left[\left(\delta + \sigma_{\mathbf{h}_k}^2 M \delta \left(\delta - \frac{M+1}{M} \right) - 1 \right) \right]} \right) \right]. \quad (20)$$

From (20), it becomes evident that for the given value of T , if β_{ul} is increased, the ergodic achievable rate of the proposed

model in the FD mode also increases.

B. ZF Beamforming in FD Mode

In this subsection, we derive the lower bounded ergodic achievable rate approximation for the proposed model in the FD mode (i.e., the first term in (6)) with ZF beamforming. As described in the previous subsection, here we assume that the users have the knowledge of the expected effective channel gain, $\mathbb{E}[a_{k,k}^{\text{ZF}}]$, where $a_{k,k}^{\text{ZF}} = \mathbf{h}_k^H \hat{\mathbf{v}}_{k,c}^{\text{ZF}}$. Therefore, we can write the lower bounded achievable ergodic rate for the k^{th} user relying on ZF beamforming as [39], [48]

$$R_{k,j}^{\text{ZF-FD}} \geq R_{k,j}^{\text{ZF-FD-LB}} = \log_2 \left(1 + \frac{\frac{P}{jN_0} |\mathbb{E}[a_{k,k}^{\text{ZF}}]|^2}{1 + \frac{P}{jN_0} \left(\text{var}\{a_{k,k}^{\text{ZF}}\} + \sum_{i \neq k}^K \mathbb{E}[|a_{k,i}^{\text{ZF}}|^2] \right) + \frac{P_q \gamma_q}{\beta_{\text{ul}} N_0}} \right), \quad (21)$$

From [41], we have $\mathbb{E}[a_{k,k}^{\text{ZF}}] = \mathbb{E}[1/\|\hat{\mathbf{V}}_c(:,k)\|]$, where $\hat{\mathbf{V}}_c(:,k)$ is the unnormalized k^{th} column of the matrix $\hat{\mathbf{V}}_c$. Using the results in [41, Lemma 4], we have

$$\mathbb{E}[a_{k,k}^{\text{ZF}}] = \mathbb{E} \left[\frac{1}{\|\hat{\mathbf{V}}_c(:,k)\|} \right] \quad (22)$$

$$\approx \sqrt{M-j+1} \sqrt{\sigma_{\mathbf{h}_k}^2}, \quad (23)$$

To compute the variance $\text{var}\{a_{k,k}^{\text{ZF}}\}$, we again refer to the results in [41, Lemma 4], where

$$\text{var}\{a_{k,k}^{\text{ZF}}\} = \text{var} \left\{ \frac{1}{\|\hat{\mathbf{V}}_c(:,k)\|} \right\} \quad (24)$$

$$\approx \frac{1}{4} \sigma_{\mathbf{h}_k}^2 \quad (25)$$

$$\approx \frac{\beta_{\text{ul}} p_k}{4(\beta_{\text{ul}} p_k + N_0 + P\nu^2)}. \quad (26)$$

Now, we compute the expected interference term in the denominator of (21), which is given by

$$\mathbb{E}[\|\mathbf{h}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}}\|^2] = \mathbb{E}[\|\hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}} + \mathbf{e}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}}\|^2] \quad (27)$$

$$\stackrel{(a)}{\leq} \mathbb{E}[\|\hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}}\|^2] + \mathbb{E}[\|\mathbf{e}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}}\|^2] \quad (28)$$

$$\stackrel{(b)}{=} \mathbb{E}[\|\mathbf{e}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}}\|^2] \quad (29)$$

$$\stackrel{(c)}{\leq} \mathbb{E}[\hat{\mathbf{v}}_{i,c}^{\text{ZF}H} \mathbb{E}[\mathbf{e}_k \mathbf{e}_k^H] \hat{\mathbf{v}}_{i,c}^{\text{ZF}}] \quad (30)$$

$$= \sigma_{\mathbf{e}_k}^2. \quad (31)$$

where (a) is obtained from the fact that $\mathbb{E}[\hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}} \hat{\mathbf{v}}_{i,c}^{\text{ZF}H} \mathbf{e}_k] = 0$ and that \mathbf{e}_k has a zero mean and it is independent of $\hat{\mathbf{h}}_k$ and $\hat{\mathbf{v}}_{i,c}^{\text{ZF}}$. Furthermore, (b) follows from the fact that $\hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_{i,c}^{\text{ZF}} = 0$ and (c) arises from the independence of \mathbf{e}_k and $\hat{\mathbf{v}}_{i,c}^{\text{ZF}}$. Substituting (23), (26) and (31) into (21) and using $\mathbb{E}[|h_{q,k}^{\text{ul}}|^2] = 1$, give the approximation of the lower-bounded

ergodic achievable rate for the k^{th} user with ZF beamforming in the FD mode, given by

$$R_{k,j}^{\text{ZF-FD-LB}} \approx \log_2 \left(1 + \frac{\frac{P}{jN_0} \sigma_{\mathbf{h}_k}^2 (M-j+1)}{1 + \frac{P}{jN_0} \left[\frac{\sigma_{\mathbf{h}_k}^2}{4} + (j-1) \sigma_{\mathbf{e}_k}^2 \right] + \frac{p_q}{\beta_{\text{ul}} N_0}} \right). \quad (32)$$

In terms of the number of cycles, $c = j + 1$, as given in (6), we can transform (32) into

$$R_{k,c}^{\text{ZF-FD-LB}} \approx \log_2 \left(1 + \frac{\frac{P}{N_0} \sigma_{\mathbf{h}_k}^2 (M-c+2)}{(c-1) \left(1 + \frac{P}{N_0} \sigma_{\mathbf{e}_k}^2 + \frac{p_q}{\beta_{\text{ul}} N_0} \right) + \frac{P}{N_0} \left[\frac{\sigma_{\mathbf{h}_k}^2}{4} - \sigma_{\mathbf{e}_k}^2 \right]} \right). \quad (33)$$

For the proposed model, the lower-bounded achievable ergodic rate for the ZF beamforming scheme in the FD mode can be expressed as

$$R_k^{\text{ZF-FD-LB}} = \frac{\beta_{\text{ul}}}{T} \sum_{c=2}^K \frac{1}{K} \sum_{k=1}^{c-1} R_{k,c}^{\text{ZF-FD-LB}}. \quad (34)$$

Substituting (33) in (34), we can write the approximated lower bounded ergodic achievable rate for the proposed model in the FD mode using ZF beamforming as

$$\tilde{R}_k^{\text{ZF-FD-LB}} = \frac{\beta_{\text{ul}}}{KT} \left[R_{1,2}^{\text{ZF-FD-LB}} + \sum_{c=3}^K \log_2 \left(1 + \frac{\frac{P}{N_0} \sigma_{\mathbf{h}_k}^2 (M-c+2)}{(c-1) \left(1 + \frac{P}{N_0} \sigma_{\mathbf{e}_k}^2 + \frac{p_q}{\beta_{\text{ul}} N_0} \right) + \frac{P}{N_0} \left[\frac{\sigma_{\mathbf{h}_k}^2}{4} - \sigma_{\mathbf{e}_k}^2 \right]} \right) \right], \quad (35)$$

where the BS experiences no SI for the first scheduled user in cycle 2. Therefore, the lower-bounded approximated rate of this user is denoted by $R_{1,2}^{\text{ZF-FD-LB}}$, which is given by (33) with $P\nu^2 = 0$.

IV. ACHIEVABLE ERGODIC RATE IN THE HD MODE

In the previous section, we computed the approximated lower-bounded ergodic achievable rate of the proposed model in the FD mode relying on both the MRT and ZF beamforming schemes. In the FD mode, the BS becomes fully aware of the imperfect CSI of all the users. Therefore, for the rest of the frame length, ZF beamforming is utilized for the DL data transmission, as detailed in Section II. In this section, we provide the analysis of the achievable ergodic rate of the HD mode. Dedicated training is performed before the beginning of the HD mode, hence the users have the instantaneous effective channel gain, $\hat{a}_{k,k}^{\text{ZF}}$.

For the achievable rate analysis associated with instantaneous effective channel gain knowledge at the user, $\hat{a}_{k,k}^{\text{ZF}}$, we use the lower-bounded rate expression derived in [28], where the achievable ergodic rate of the ZF beamforming associated with Gaussian inputs, DL training and UL channel feedback has been shown to be lower-bounded by [28]

$$R_k^{\text{ZF-HD}} \geq R_k^{\text{ZF-HD-LB}} = \mathbb{E} \left[\log_2 \left(1 + \frac{\frac{P}{jN_0} |\hat{a}_{k,k}^{\text{ZF}}|^2}{1 + \frac{P}{jN_0} \left(\sigma_{\lambda_k}^2 + \mathbb{E} \left[|I_k|^2 | \hat{a}_{k,k}^{\text{ZF}} \right] \right)} \right) \right], \quad (36)$$

where the interference term is given by $I_k = \sum_{i \neq k}^K \mathbf{h}_k^H \hat{\mathbf{v}}_i^{\text{ZF}}$. It is analytically challenging to compute the expectation in (36). Therefore, similar to [51]–[53], we approximate (36) by taking the expectation of the ratio, giving

$$R_k^{\text{ZF-HD-LB}} \approx \log_2 \left(1 + \frac{\frac{P}{jN_0} \mathbb{E} \left[|\hat{a}_{k,k}^{\text{ZF}}|^2 \right]}{1 + \frac{P}{jN_0} \sigma_{\lambda_k}^2 + \frac{P}{jN_0} \sum_{i \neq k}^K \mathbb{E} \left[\mathbb{E} \left[|I_k|^2 | \hat{a}_{k,k}^{\text{ZF}} \right] \right]} \right) \stackrel{(a)}{=} \log_2 \left(1 + \frac{\frac{P}{jN_0} \mathbb{E} \left[|\hat{a}_{k,k}^{\text{ZF}}|^2 \right]}{1 + \frac{P}{jN_0} \sigma_{\lambda_k}^2 + \frac{P}{jN_0} \sum_{i \neq k}^K \mathbb{E} \left[|I_k|^2 \right]} \right), \quad (37)$$

where (a) follows from the law of total expectation, so that $\mathbb{E} \left[\mathbb{E} \left[|I_k|^2 | \hat{a}_{k,k}^{\text{ZF}} \right] \right] = \mathbb{E} \left[|I_k|^2 \right]$. We compute $\mathbb{E} \left[|\hat{a}_{k,k}^{\text{ZF}}|^2 \right]$ by using (5), to have

$$\begin{aligned} \mathbb{E} \left[|\hat{a}_{k,k}^{\text{ZF}}|^2 \right] &= \mathbb{E} \left[\left| \frac{\sqrt{\beta_{\text{dt}} P}}{N_0 + \beta_{\text{dt}} P} \left(\sqrt{\beta_{\text{dt}} P} a_{k,k}^{\text{ZF}} + z_{k,k} \right) \right|^2 \right] \\ &= \frac{\beta_{\text{dt}} P}{(N_0 + \beta_{\text{dt}} P)^2} \mathbb{E} \left[\left(\sqrt{\beta_{\text{dt}} P} \hat{\mathbf{v}}_k^{\text{ZF}H} \mathbf{h}_k + z_{k,k} \right) \left(\sqrt{\beta_{\text{dt}} P} \mathbf{h}_k^H \hat{\mathbf{v}}_k^{\text{ZF}} + z_{k,k} \right) \right] \\ &= \frac{\beta_{\text{dt}} P}{(N_0 + \beta_{\text{dt}} P)^2} \left(\mathbb{E} \left[\beta_{\text{dt}} P \left(\hat{\mathbf{v}}_k^{\text{ZF}H} \mathbf{h}_k \mathbf{h}_k^H \hat{\mathbf{v}}_k^{\text{ZF}} \right) \right] + \mathbb{E} \left[z_{k,k}^H z_{k,k} \right] \right) \\ &\stackrel{(a)}{\approx} \frac{\beta_{\text{dt}} P \left[\beta_{\text{dt}} P (M-K+1) + N_0 \right]}{(N_0 + \beta_{\text{dt}} P)^2}, \end{aligned} \quad (38)$$

where (a) is obtained by using $\mathbb{E} \left[\hat{\mathbf{v}}_k^{\text{ZF}H} \mathbf{h}_k \mathbf{h}_k^H \hat{\mathbf{v}}_k^{\text{ZF}} \right] \approx M-K+1$ (assuming that the residual SI is small) and $\mathbb{E} \left[z_{k,k}^H z_{k,k} \right] = N_0$. Now, we compute the expectation of the interference term in (37) using the lower bound given in [28, Appendix III], yielding:

$$\mathbb{E} \left[|I_k|^2 \right] = \mathbb{E} \left[\left| \mathbf{h}_k^H \hat{\mathbf{v}}_i^{\text{ZF}} \right|^2 \right] \quad (39)$$

$$\stackrel{(a)}{\leq} \mathbb{E} \left[\left| \hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_i^{\text{ZF}} + \mathbf{e}_k^H \hat{\mathbf{v}}_i^{\text{ZF}} \right|^2 \right] \quad (40)$$

$$\stackrel{(b)}{\leq} \mathbb{E} \left[\left| \mathbf{e}_k^H \hat{\mathbf{v}}_i^{\text{ZF}} \right|^2 \right] \quad (41)$$

$$\stackrel{(c)}{\leq} \mathbb{E} \left[\hat{\mathbf{v}}_i^{\text{ZF}H} \mathbb{E} \left[\mathbf{e}_k \mathbf{e}_k^H \right] \hat{\mathbf{v}}_i^{\text{ZF}} \right] \quad (42)$$

$$\stackrel{(d)}{\leq} \sigma_{\mathbf{e}_k}^2, \quad (43)$$

where (a) follows from (9), (b) comes from the fact that $\hat{\mathbf{h}}_k^H \hat{\mathbf{v}}_i^{\text{ZF}} = 0$, (c) is due to the fact that \mathbf{e}_k and $\hat{\mathbf{v}}_i^{\text{ZF}}$ are independent vectors, while (d) is obtained by using $\mathbb{E} \left[\mathbf{e}_k \mathbf{e}_k^H \right] = \sigma_{\mathbf{e}_k}^2 \mathbf{I}$. The lower bounded ergodic achievable rate of the k^{th} user in the HD mode using ZF beamforming, can be approximated

by substituting (38) and (43) into (37), giving

$$\begin{aligned} \tilde{R}_k^{\text{HD-LB}} &= \left(1 - \frac{T_t}{T}\right) R_k^{\text{ZF-HD-LB}} \approx \\ &\left(1 - \frac{T_t}{T}\right) \log_2 \left(1 + \frac{\frac{P}{KN_0} \frac{\beta_{\text{dt}} P (\beta_{\text{dt}} P (M-K+1) + N_0)}{(N_0 + \beta_{\text{dt}} P)^2}}{1 + \frac{P}{KN_0} \sigma_{\lambda_k}^2 + \frac{P}{KN_0} (K-1) \sigma_{e_k}^2}\right). \end{aligned} \quad (44)$$

Note that in (44), we have assumed that the UL CSI of the first scheduled user at the BS is also suffering from the effects of the residual SI. Note that the effect of residual SI is included in $\sigma_{e_k}^2$ (10) for $\tilde{R}_k^{\text{HD-LB}}$. For the traditional HD scheme, there is no residual SI, therefore, we have $P\nu^2 = 0$ in $\sigma_{e_k}^2$. The approximation of the lower-bounded ergodic achievable rate of the proposed model relying on MRT beamforming in the FD mode is obtained by summing up (20) and (44). On the other hand, when ZF beamforming is employed in the FD mode, the lower-bounded ergodic achievable rate of the proposed model can be approximated by summing (35) and (44). It is important to note that in the proposed model a total of $K-1$ users are served by the BS in the FD mode, whereas, all the K users are supported in the HD mode.

V. NUMERICAL RESULTS

In this section, we provide numerical results for the proposed DL data transmission relying on UL CSI acquisition model. We compare the performance of ZF beamforming and MRT beamforming schemes in the FD mode while ZF beamforming is employed in the HD mode for both schemes. We also compare the combined FD and HD schemes to the traditional HD based ZF beamforming scheme. Here, we also validate the approximated ergodic achievable rate analysis carried out in Section III and Section IV. Throughout this section, we assume that the noise power is normalized to one, so that $N_0 = 1$ and. The SNR is given by P/N_0 . Similar to [28], in this section, we also assume that the DL and UL SNR values are the same, by using $p_k = P$. The variance of the residual SI, ν^2 , is set to -10 dB, unless stated otherwise. The results presented in this section are averaged over a total of 5000 channel realizations. In this section, the lower bounded rate approximations derived in Section III and Section IV are also confirmed by the Monte Carlo simulations.

The carrier frequency of the system is given by $f_c = 5$ GHz. Denoting λ_0 as the wavelength, we compute the coherence time as $T_{\text{coh}} \approx \lambda_0/2v = 15.43$ ms (at $v = 7$ km/h). Using delay spread of $5 \mu\text{s}$ [54], we can compute the coherence bandwidth as $B_{\text{coh}} \approx 1/8(5\mu\text{s}) \approx 25$ KHz. The total number of channel uses can be calculated by using $T = T_{\text{coh}} B_{\text{coh}} \approx 385$. Therefore, throughout the simulations, we use $T = 400$ channel uses, unless stated otherwise.

In the figures, we refer to the ‘FD (ZF) - HD’ scheme, where ZF beamforming is used in both the FD and HD modes, whereas, we use ‘FD (MRT) - HD’ to represent the scheme, where MRT beamforming is used in the FD mode and ZF beamforming is utilized in the HD mode. For the traditional HD based ZF beamforming method, we refer to as ‘HD’ in the figures. Note that for a fair comparison, the traditional HD rates are scaled by the factor $(1 - T_t/T)$. Note

that the conventional HD system characterized in this section is illustrated in Fig. 1b, where only a HD BS is assumed, hence no SI is imposed on the system.

In this section, we plot the lower bounds on sum-rates, using the lower-bounded rates given by (11) and (36) for the ‘FD (MRT) - HD’ scheme, and (21) as well as (36) for the ‘FD (ZF) - HD’ scheme. On the other hand, the lower-bounded achievable rate of the traditional HD scheme is computed by using only the second term in (6), given by $R_k^{\text{ZF-HD-LB}}$ in (37). Note that there is no residual SI (i.e., $P\nu^2 = 0$) for the conventional HD scheme. Denoting $C^{\text{sum-MRT}}$ and $C^{\text{sum-ZF}}$ as the lower-bounded sum-rates of the MRT beamforming and ZF beamforming schemes, respectively, we define

$$C^{\text{sum-MRT}} \triangleq (K-1) R_k^{\text{MRT-FD-LB}} + K R_k^{\text{ZF-HD-LB}} \quad (45)$$

and

$$C^{\text{sum-ZF}} \triangleq (K-1) R_k^{\text{ZF-FD-LB}} + K R_k^{\text{ZF-HD-LB}}, \quad (46)$$

where $R_k^{\text{MRT-FD-LB}}$, $R_k^{\text{ZF-FD-LB}}$ and $R_k^{\text{ZF-HD-LB}}$ are given by (19), (34) and (37), respectively.

The approximated lower-bounded ergodic achievable rate of ‘FD (MRT) - HD’ is obtained by summing up (35) and (44). Similarly, the approximated lower bounded achievable rate of ‘FD (ZF) - HD’ is calculated by summing up (20) and (44). To obtain the sum-rate results, the FD rates are scaled by $K-1$, whereas, the HD rate is multiplied by K . Denoting $C_{\text{approx}}^{\text{sum-MRT}}$ and $C_{\text{approx}}^{\text{sum-ZF}}$ as the approximated lower-bounded sum-rates of the MRT beamforming and ZF beamforming schemes, respectively, we define

$$C_{\text{approx}}^{\text{sum-MRT}} \triangleq (K-1) \tilde{R}_k^{\text{MRT-FD-LB}} + K \tilde{R}_k^{\text{HD-LB}} \quad (47)$$

and

$$C_{\text{approx}}^{\text{sum-ZF}} \triangleq (K-1) \tilde{R}_k^{\text{ZF-FD-LB}} + K \tilde{R}_k^{\text{HD-LB}}, \quad (48)$$

where $\tilde{R}_k^{\text{MRT-FD-LB}}$, $\tilde{R}_k^{\text{ZF-FD-LB}}$ and $\tilde{R}_k^{\text{HD-LB}}$ are given by (20), (35) and (44), respectively.

The sum-rate performance is plotted against the SNR in Fig. 2 with $M = 16$, $\beta_{\text{ul}} = 12$ and $\beta_{\text{dt}} = 8$. We characterize two cases in this figure associated with $K = 8$ and $K = 16$. In both cases, we note that the sum-rate performance of the ‘FD (ZF) - HD’ scheme is higher than that of the ‘FD (MRT) - HD’ scheme, which becomes more dominant at high SNR values. The improvement in the sum-rate of ‘FD (ZF) - HD’ is roughly about 1 bit/channel use in the $K = 8$ case, but much higher gains are reported in the case of $K = 16$. This performance trend is due to the reason that as the SNR increases, the inter-user interference also increases. Hence in contrast to the MRT beamforming scheme, in the FD mode the ZF beamforming scheme mitigates the interference more efficiently. The sum-rate gain attained by utilizing the FD-HD modes for the proposed DL data transmission and UL CSI acquisition is clearly evident from Fig. 2, when compared to the traditional HD based ZF beamforming scheme. However, for the $K = 8$ case, at higher SNR values (i.e., SNR > 18 dB), the traditional HD scheme outperforms the proposed FD-HD model. The reason for this trend is that as the SNR increases, the residual SI power (i.e., $P\nu^2$ in (8)) also increases for the proposed FD-HD schemes, hence reducing the quality

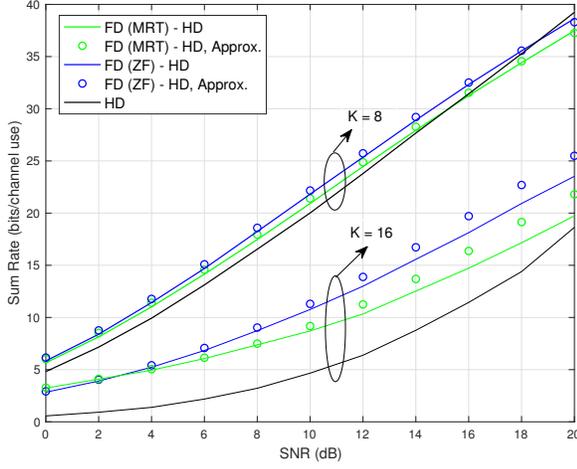


Fig. 2: The sum-rate performance (45), (46), (47) and (48) against SNR values with $M = 16$, $K = \{8, 16\}$, $\beta_{ul} = 12$ and $\beta_{dt} = 8$.

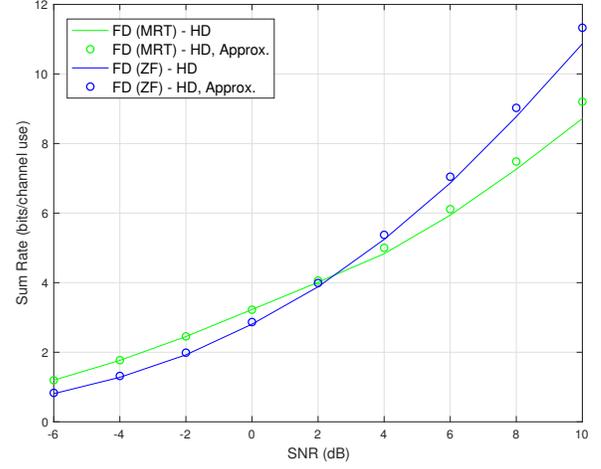


Fig. 4: The sum-rate performance (45), (46), (47) and (48) against SNR values with $M = K = 16$, $\beta_{ul} = 12$ and $\beta_{dt} = 4$.

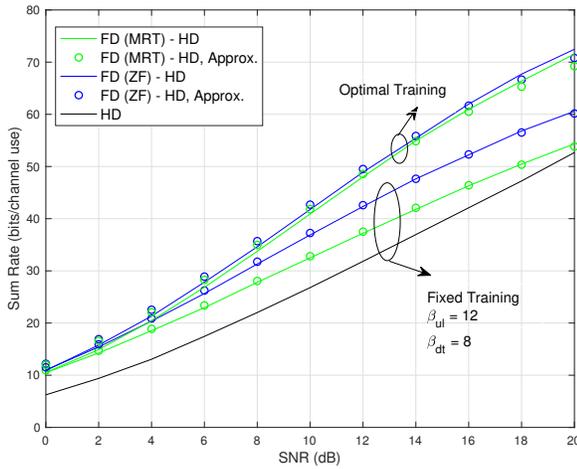


Fig. 3: The sum rate performance (45), (46), (47) and (48) against SNR values with $M = 32$, $K = 16$, $\beta_{ul} = \{12, \beta_{ul}^*\}$ and $\beta_{dt} = \{8, \beta_{dt}^*\}$.

of channel estimation at the BS. On the other hand, for the $K = 16$ case, the HD scheme has a significantly lower performance than the proposed schemes. It is also observed that the approximations of the lower bounded ergodic rates obtained in Section III (for the FD mode) and Section IV (for the HD mode) closely match the average lower bounded rates attained via Monte Carlo simulations.

In Fig. 3, we plot the sum-rate versus SNR for $M = 32$ and $K = 16$ for both fixed and optimal training parameters. The fixed training durations are $\beta_{ul} = 12$ and $\beta_{dt} = 8$. We denote the optimal training duration by β_{ul}^* and β_{dt}^* , which are obtained by performing the exhaustive search over the optimization problems given by $\max(\tilde{R}_k^{\text{MRT-FD-LB}} + \tilde{R}_k^{\text{HD-LB}})$ for the ‘FD (MRT) - HD’ scheme and by $\max(\tilde{R}_k^{\text{ZF-FD-LB}} + \tilde{R}_k^{\text{HD-LB}})$ for the ‘FD (ZF) - HD’ scheme. The exhaustive search is

performed over the set of integer numbers for both β_{ul} and β_{dt} . The ‘FD (ZF) - HD’ scheme outperforms all the other schemes characterized in Fig. 3. For example, at SNR = 10 dB, the ‘FD (ZF) - HD’ scheme provides a gain of roughly 5 bits/channel use and 12 bits/channel use compared to the ‘FD (MRT) - HD’ and HD schemes, respectively. Again, we note that the lower-bounded ergodic rate approximations derived in Section III and Section IV are tight and match the numerically computed results. The optimal training (i.e., β_{ul}^* and β_{dt}^*) results give a significant performance gain compared to the fixed training parameters. This underlines the importance of the optimal training in FD transmission schemes to improve the spectral efficiency of the system. We also compare the performance of the proposed schemes at low SNR values in Fig. 4. We note that the MRT beamforming scheme in FD mode yields a better performance than the ZF beamforming scheme for SNR < 2 dB. Thereafter, the ZF beamforming scheme starts dominating since it becomes capable of mitigating the increasing inter-user interference more effectively.

To investigate the performance of the proposed transmission scheme upon varying the frame length, T , we plot the sum-rate results against the T values in Fig. 5. Intuitively, as the number of channel uses T increases, the training duration T_t should also increase. Therefore, in Fig. 5, we use the results derived in [34] to vary T_t and to compute the β_{ul} values according to T_t . As reported in [34], the values of β_{dt} have a marginal effect on the performance. Therefore, we keep the β_{dt} value fix to $\beta_{dt} = 4$. According to [34], $T_t \propto \sqrt{(M-1)T}$, and therefore, we set $\beta_{ul} = (\lfloor T_t \rfloor - \beta_{dt}) / K$, where $\lfloor \cdot \rfloor$ represents the floor function and $T_t \triangleq \sqrt{(M-1)T}$. These values are not optimal values for this study and the main reason to use them is to scale the duration of the training period, T_t , upon changing the T values. In Fig. 5, the ‘FD (ZF) - HD’ scheme outperforms the ‘FD (MRT) - HD’ and HD based imperfect CSI schemes as T increases. More specifically, the proposed FD-HD based model yields a better sum-rate performance

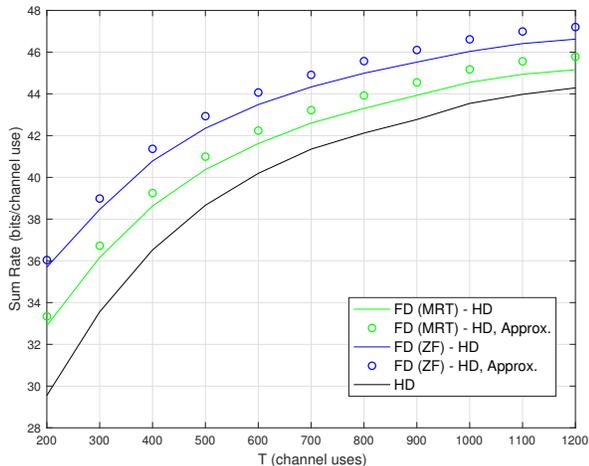


Fig. 5: The sum-rate performance (45), (46), (47) and (48) against different number of channel uses with SNR = 10 dB, $M = 32$, $K = 16$ and $\beta_{dt} = 4$.

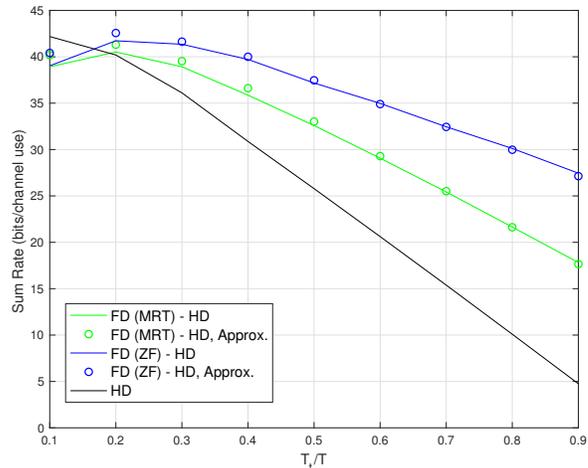


Fig. 7: The sum-rate performance (45), (46), (47) and (48) versus T_t/T values for SNR = 10 dB, $T = 400$, $M = 32$, $K = 16$ and $\beta_{dt} = 4$.

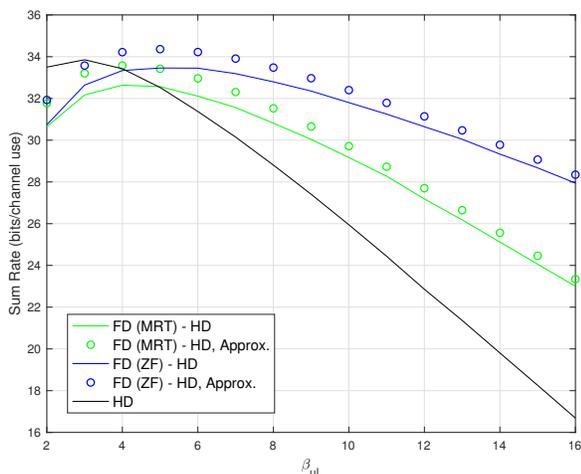


Fig. 6: The sum rate performance (45), (46), (47) and (48) against various β_{ul} values with SNR = 10 dB, $T = 400$, $M = 25$, $K = 15$ and $\beta_{dt} = 4$.

compared to its HD counterpart. Here, the ‘HD Perfect CSI’ scheme yields the best performance. All the schemes plotted in Fig. 5 show improvement, as T increases. This improvement is due to the fact that as T increases, the training duration T_t also increases, but the ratio T_t/T reduces. This is the reason that the performance gap between the FD-HD based proposed schemes and the conventional HD scheme reduces as T increases. In other words as T increases, the rate gain of the FD mode decreases, because the factor $(1 - T_t/T)$ approaches 1.

Fig. 6 shows the sum-rate performance against different β_{ul} values associated with $M = 25$, $K = 15$, SNR = 10 dB and $\beta_{dt} = 4$. Here, we notice that the sum-rate performance of all the schemes degrades, as β_{ul} increases. This is because as β_{ul} increases, the scaling factor $(1 - T_t/T)$ reduces (given

the fixed value of $T = 400$), resulting in smaller sum-rates in the HD mode for the proposed schemes. We also note that increasing β_{ul} improves the performance gap between the proposed FD-HD schemes and the conventional HD scheme. The performance of the conventional HD scheme with perfect CSI diminishes rapidly, as β_{ul} increases. At smaller values of β_{ul} , the performance of the traditional HD scheme becomes better than that of the proposed model. The reason for this trend is that the ratio T_t/T becomes small for smaller β_{ul} values and the performance is mostly dominated by the sum-rate of the HD mode. Therefore, the presence of residual SI during the UL CSI acquisition degrades the performance of the proposed model compared to the traditional HD scheme.

In Fig. 7 we portray the sum-rate performance upon increasing the training duration, T_t , for the fixed value of T . More specifically, we define a ratio T_t/T where T_t varies while T is kept fixed to 400. Initially, the proposed schemes show performance improvement, but as the ratio T_t/T increases, the performance gradually decreases. However, the performance of the conventional HD scheme degrades severely as the ratio increases, which is due to the fact that the scaling factor $(1 - T_t/T)$ decreases as T_t increases. The gradual decrease in the performance of the proposed FD-HD schemes is also due to the reduction in the achievable rate of the HD mode, as the ratio T_t/T increases. Figures 6 and 7 clearly indicate that when the training period is long, the FD assisted BS can help in accomplishing higher sum-rates. In Fig. 8, we plot the sum-rate results against different values of the residual SI power, ν^2 . It is seen that the sum-rate performance reduces, as the residual SI power increases for the proposed FD-HD model. This performance reduction is due to the fact that when the residual SI power increases, it reduces the quality of the channel estimation at the BS by degrading the UL CSI acquisition. On the other hand, the performance of the traditional HD scheme remains constant, since it is not affected by the residual SI.

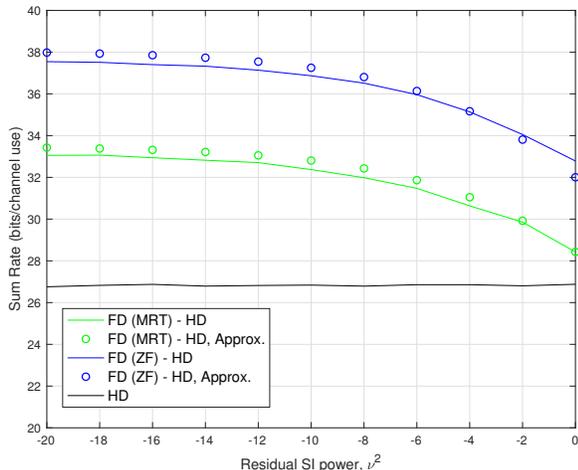


Fig. 8: The sum-rate performance (45), (46), (47) and (48) for various ν^2 values with SNR = 10 dB, $M=32$, $K=16$, $\beta_{ul}=12$ and $\beta_{dt}=8$.

TABLE I: Complexity Comparison with $M = K$.

	$M = 4$ (flops)	$M = 8$ (flops)	$M = 12$ (flops)	$M = 16$ (flops)
FD (MRT) - HD	1.9×10^3	2.3×10^4	9.2×10^4	2.3×10^5
FD (ZF) - HD	2.4×10^4	8.6×10^5	6.8×10^6	2.9×10^7
HD	1.1×10^3	1.5×10^4	6.4×10^4	1.6×10^5

To quantify the implementation complexity of the proposed scheme, we rely on the number of floating-point operations per second (FLOPS) [55]. The complexity of computing the MRT and ZF beamforming vectors at the BS can be represented by $\xi^{\text{MRT}} = 3M - 1$ and $\xi^{\text{ZF}} = 4MK^2 + K^3 + K(1 - M) + 3M - 1$, respectively. Denoting \mathcal{T}^{MRT} and \mathcal{T}^{ZF} as the complexity of 'FD (MRT) - HD' and 'FD (ZF) - HD' schemes, we have

$$\mathcal{T}^{\text{MRT}} = \left(\sum_{c=2}^K (c-1) \beta_{ul} \xi^{\text{MRT}} \right) + KM + (T - T_t) K \xi^{\text{ZF}} \quad (49)$$

and

$$\mathcal{T}^{\text{ZF}} = \left(\sum_{c=2}^K (c-1) \beta_{ul} \xi^{\text{ZF}} \right) + KM + (T - T_t) K \xi^{\text{ZF}}, \quad (50)$$

respectively, where the additive term KM comes from the vector scaling operation required for the MMSE operation and $(T - T_t) K \xi^{\text{ZF}}$ is the complexity of the HD scheme. The computational complexity of the proposed schemes is compared to that of the traditional HD scheme in Table I. The ZF beamforming scheme imposes a higher complexity than the MRT beamforming and the conventional HD schemes, because the ZF beamformer requires matrix multiplication and inversion operations.

Some key points observed in this section are that the FD-assisted BS improves the sum-rate of the MIMO system. When the UL training period is long, the FD mode can be utilized to achieve a better performance than the HD scheme. Although

the residual SI degrades the rate of the FD mode, it still achieves a better sum-rate than the conventional HD scheme, provided that the residual SI power is not too high.

VI. CONCLUSION

In this paper, we have proposed MU MIMO DL transmission and UL CSI acquisition based on combined FD mode and HD mode operations. The FD mode is utilized simultaneously for supporting both the DL data transmission and the UL CSI pilot reception. We derived the lower-bounded ergodic achievable rates for the proposed scheme, in the presence of residual SI. Through simulation results, it is observed that using the FD mode for DL data transmission during the UL CSI acquisition stage provides an achievable rate improvement when compared to the conventional HD transmission strategy. Furthermore, it is also observed that when the residual SI power becomes high, it degrades the performance of the proposed scheme by reducing the rate advantage attained during the FD mode. Moreover, we observed that the lower-bounded ergodic achievable rate approximations derived in this paper closely match the numerical results. For future work, it will be worthwhile investigating the derived lower-bounded ergodic achievable rate approximations to determine the optimal training duration under different system parameters.

APPENDIX I

PROOF OF (8)

From (7), we have

$$\hat{\mathbf{h}}_k = \mathbb{E} [\mathbf{h}_k \mathbf{r}_k^H] \mathbb{E} [\mathbf{r}_k \mathbf{r}_k^H]^{-1} \mathbf{r}_k. \quad (51)$$

In the above expression, the first expectation can be written as

$$\mathbb{E} [\mathbf{h}_k \mathbf{r}_k^H] = \mathbb{E} \left[\mathbf{h}_k \left(\sqrt{\beta_{ul} p_k} \mathbf{h}_k + \tilde{\boldsymbol{\psi}} + \tilde{\mathbf{n}}_k \right)^H \right]. \quad (52)$$

We can write $\tilde{\boldsymbol{\psi}} = \mathbf{H}_{\text{SI}} \sum_{i=1}^U \hat{\mathbf{v}}_i^{\text{MRT}} q_i$, where $q_i = \tilde{u}_{i,1} + \dots + \tilde{u}_{i,\beta_{ul}}$ and $\tilde{u}_{i,n} = u_{i,n} s_n$, such that $\mathbb{E} [|q_i|^2] = P/U$. Therefore, we can write (52) as

$$\begin{aligned} \mathbb{E} [\mathbf{h}_k \mathbf{r}_k^H] &= \mathbb{E} \left[\mathbf{h}_k \left(\sqrt{\beta_{ul} p_k} \mathbf{h}_k + \sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i + \tilde{\mathbf{n}}_k \right)^H \right] \\ &= \mathbb{E} \left[\sqrt{\beta_{ul} p_k} \mathbf{h}_k \mathbf{h}_k^H + \sum_{i=1}^U \mathbf{h}_k q_i^H \hat{\mathbf{v}}_i^{\text{MRT}H} \mathbf{H}_{\text{SI}}^H + \mathbf{h}_k \tilde{\mathbf{n}}_k^H \right] \\ &\stackrel{(a)}{=} \mathbb{E} \left[\sqrt{\beta_{ul} p_k} \mathbf{h}_k \mathbf{h}_k^H \right] \\ &= \sqrt{\beta_{ul} p_k} \mathbf{I}_M, \end{aligned} \quad (53)$$

where (a) comes from the fact that due to independence, we have $\mathbb{E} [\mathbf{h}_k \tilde{\mathbf{n}}_k^H] = \mathbf{0}$ and $\mathbb{E} [\mathbf{h}_k \left(\sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i \right)^H] = \mathbf{0}$. Now, we find the second expectation in (51), which can be explicitly expressed as

$$\begin{aligned}
\mathbb{E} \left[\mathbf{r}_k \mathbf{r}_k^H \right]^{-1} &= \mathbb{E} \left[\left(\sqrt{\beta_{\text{ul}} p_k} \mathbf{h}_k + \sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i + \tilde{\mathbf{n}}_k \right) \times \right. \\
&\quad \left. \left(\sqrt{\beta_{\text{ul}} p_k} \mathbf{h}_k + \sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i + \tilde{\mathbf{n}}_k \right)^H \right]^{-1} \\
&\stackrel{(a)}{=} \left(\mathbb{E} \left[\beta_{\text{ul}} p_k \mathbf{h}_k \mathbf{h}_k^H \right] + U \mathbb{E} \left[\mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i q_i^H \hat{\mathbf{v}}_i^{\text{MRT}H} \mathbf{H}_{\text{SI}}^H \right] + \mathbb{E} \left[\tilde{\mathbf{n}}_k \tilde{\mathbf{n}}_k^H \right] \right)^{-1} \\
&\stackrel{(b)}{=} \left(\beta_{\text{ul}} p_k \mathbf{I}_M + \left(U \frac{P}{U} \frac{1}{M} M \nu^2 \right) \mathbf{I}_M + N_0 \mathbf{I}_M \right)^{-1} \\
&= (\beta_{\text{ul}} p_k + N_0 + P \nu^2)^{-1} \mathbf{I}_M, \tag{54}
\end{aligned}$$

where (a) follows from the fact that due to independence, we have $\mathbb{E}[\mathbf{h}_k \tilde{\mathbf{n}}_k^H] = \mathbf{0}$, $\mathbb{E}[\tilde{\mathbf{n}}_k \mathbf{h}_k^H] = \mathbf{0}$, $\mathbb{E}[\mathbf{h}_k (\sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i)^H] = \mathbf{0}$, $\mathbb{E}[(\sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i) \mathbf{h}_k^H] = \mathbf{0}$ and $\mathbb{E}[(\sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i) \tilde{\mathbf{n}}_k^H] = \mathbf{0}$, and (b) is obtained using $\mathbb{E}[q_i q_i^H] = (P/U) \mathbf{I}_M$, $\mathbb{E}[\hat{\mathbf{v}}_i^{\text{MRT}} \hat{\mathbf{v}}_i^{\text{MRT}H}] = (1/M) \mathbf{I}_M$ and $\mathbb{E}[\mathbf{H}_{\text{SI}} \mathbf{H}_{\text{SI}}^H] = (M \nu^2) \mathbf{I}_M$. Substituting (53) and (54) into (51) yields (8) and also completes the proof. ■

APPENDIX II

PROOF OF (12)

We can write the expectation of $a_{k,k}^{\text{MRT}} = \mathbf{h}_k^H \hat{\mathbf{v}}_k^{\text{MRT}}$ as

$$\mathbb{E} [a_{k,k}^{\text{MRT}}] = \mathbb{E} [\mathbf{h}_k^H \hat{\mathbf{v}}_k^{\text{MRT}}] = \mathbb{E} \left[\mathbf{h}_k^H \frac{\hat{\mathbf{h}}_k}{\|\hat{\mathbf{h}}_k\|} \right] \tag{55}$$

$$= \mathbb{E} \left[\mathbf{h}_k^H \frac{\mathbf{r}_k}{\|\mathbf{r}_k\|} \right]. \tag{56}$$

We can express (56) as a ratio of two random variables R and S , where $R = \mathbf{h}_k^H \mathbf{r}_k$ and $S = \|\mathbf{r}_k\|$. The mean values of R and S are denoted by μ_R and μ_S , respectively. The variances of R and S are represented by σ_R^2 and σ_S^2 , respectively. The computation of the expectation, $\mathbb{E}[R/S]$, is often not straightforward. However, it can be solved using the Taylor-series approximation [56], [57]. We use the approach presented in [57] for approximating $\mathbb{E}[R/S]$, where second-order Taylor expansion is used to obtain

$$\mathbb{E} \left[\frac{R}{S} \right] \approx \frac{\mu_R}{\mu_S} - \frac{\text{cov}\{R, S\}}{\mu_S^2} + \frac{\sigma_S^2 \mu_R}{\mu_S^3}, \tag{57}$$

where $\text{cov}\{\cdot, \cdot\}$ denotes the covariance. We further simplify the approximation in (57) by ignoring $\text{cov}\{R, S\}/\mu_S^2$ as $\mu_S^2 \gg \text{cov}\{R, S\}$, yielding

$$\mathbb{E} \left[\frac{R}{S} \right] \approx \frac{\mu_R}{\mu_S} + \frac{\sigma_S^2 \mu_R}{\mu_S^3}. \tag{58}$$

We can find the value of μ_R as

$$\begin{aligned}
\mu_R &= \mathbb{E} [\mathbf{h}_k^H \mathbf{r}_k] \\
&= \mathbb{E} \left[\mathbf{h}_k^H \left(\sqrt{\beta_{\text{ul}} p_k} \mathbf{h}_k + \sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i + \tilde{\mathbf{n}}_k \right) \right] \\
&\stackrel{(a)}{=} \mathbb{E} \left[\sqrt{\beta_{\text{ul}} p_k} \|\mathbf{h}_k\|^2 \right] \\
&= \sqrt{\beta_{\text{ul}} p_k} M, \tag{59}
\end{aligned}$$

where (a) follows the fact that due to independence $\mathbb{E}[\mathbf{h}_k^H \tilde{\mathbf{n}}_k] = 0$ and $\mathbb{E}[\mathbf{h}_k^H (\sum_{i=1}^U \mathbf{H}_{\text{SI}} \hat{\mathbf{v}}_i^{\text{MRT}} q_i)] = 0$. The

variance of the elements in the vector \mathbf{r}_k is equal to $\sigma_r^2 = \beta_{\text{ul}} p_k + N_0 + P \nu^2$, which is equal to the sum of the variance of each term in \mathbf{r}_k . The elements in \mathbf{h}_k have a unit variance, and therefore, we can find the variance of R becomes:

$$\begin{aligned}
\sigma_R^2 &= \text{var} \{ \mathbf{h}_k^H \mathbf{r}_k \} \\
&= \text{var} \{ h_1^H r_1 + h_2^H r_2 + \dots + h_M^H r_M \} \\
&= M (\beta_{\text{ul}} p_k + N_0 + P \nu^2). \tag{60}
\end{aligned}$$

The value of μ_S is calculated as

$$\mu_S = \mathbb{E} [\|\mathbf{r}_k\|]. \tag{61}$$

Although the residual SI term in \mathbf{r}_k in (2) is not distributed according to the complex Gaussian distribution, it is observed that the entries of \mathbf{r}_k follow the normal distribution with variance equal to $\sigma_r^2 \mathbf{I}$. Therefore, we can approximate $\mathbf{r}_k \approx \mathbf{t}_k$, where $\mathbf{t}_k \sim \mathcal{CN}(0, \sigma_r^2 \mathbf{I})$. Using this assumption, we can approximate (61) as

$$\mu_S \approx \sqrt{\beta_{\text{ul}} p_k + N_0 + P \nu^2} \mathbb{E} [\|\bar{\mathbf{t}}_k\|], \tag{62}$$

where $\bar{\mathbf{t}}_k = \mathbf{t}_k / \|\mathbf{t}_k\|$ and the entries of $\bar{\mathbf{t}}_k$ are distributed according to the complex normal distribution with zero mean and unit variance. Thus according to [48], we have $\mathbb{E}[\|\bar{\mathbf{t}}_k\|] = \Gamma(M + 1/2) / \Gamma(M)$, giving

$$\mu_S \approx \sqrt{\beta_{\text{ul}} p_k + N_0 + P \nu^2} \frac{\Gamma(M + \frac{1}{2})}{\Gamma(M)}. \tag{63}$$

The variance of S can be expressed as

$$\sigma_S^2 = \mathbb{E} [S^2] - \mathbb{E} [S]^2 \tag{64}$$

$$= \mathbb{E} [\|\mathbf{r}_k\|^2] - \mathbb{E} [\|\mathbf{r}_k\|]^2. \tag{65}$$

Using a similar approach to that in Appendix I, we have $\mathbb{E} [\|\mathbf{r}_k\|^2] = M(\beta_{\text{ul}} p_k + N_0 + P \nu^2)$. The value of $\mathbb{E}[\|\mathbf{r}_k\|]$ is approximated in (63). Therefore, we can write (65) as

$$\sigma_S^2 \approx (\beta_{\text{ul}} p_k + N_0 + P \nu^2) \left[M - \frac{\Gamma^2(M + \frac{1}{2})}{\Gamma^2(M)} \right]. \tag{66}$$

Substituting the values of (59), (63) and (66) in (58) and rearranging gives

$$\begin{aligned}
\mathbb{E} \left[\frac{R}{S} \right] &= \frac{\sqrt{\beta_{\text{ul}} p_k}}{\sqrt{\beta_{\text{ul}} p_k + N_0 + P \nu^2}} \left[\frac{M \Gamma(M)}{\Gamma(M + \frac{1}{2})} + \right. \\
&\quad \left. \frac{\Gamma(M) (M \Gamma^2(M) - \Gamma^2(M + \frac{1}{2}))}{\Gamma^3(M + \frac{1}{2})} \right] \\
&\stackrel{(a)}{\approx} \frac{\sqrt{\beta_{\text{ul}} p_k} M}{\sqrt{\beta_{\text{ul}} p_k + N_0 + P \nu^2}}, \tag{67}
\end{aligned}$$

where (a) comes from $\frac{M \Gamma(M)}{\Gamma(M + (1/2))} \approx \sqrt{M}$ and $\frac{\Gamma(M)}{\Gamma(M + (1/2))} \approx 1/\sqrt{M}$. This completes the proof. ■

APPENDIX III

PROOF OF (13)

To compute the variance in (13), we rely on the Taylor series approximation of the variance of the ratio of the two random variables R and S , given by $\text{var}\{R/S\}$. As shown in Appendix

II, in this study, $R = \mathbf{h}_k^H \mathbf{r}_k$ and $S = \|\mathbf{r}_k\|$. According to [57], the variance of the ratio R/S can be approximated using the first-order Taylor series expansion, as

$$\text{var} \left\{ \frac{R}{S} \right\} \approx \frac{\mu_R^2}{\mu_S^2} \left[\frac{\sigma_R^2}{\mu_R^2} - 2 \frac{\text{cov} \{R, S\}}{\mu_R \mu_S} + \frac{\sigma_S^2}{\mu_S^2} \right]. \quad (68)$$

Using the Cauchy-Schwarz inequality, we can write $\text{cov} \{R, S\} \leq \sqrt{\sigma_R^2 \sigma_S^2}$, and therefore, we can approximate the quantity $(2 \text{cov} \{R, S\} / (\mu_R \mu_S)) \approx 1/M$. Now substituting the values of (59), (60), (63) and (66) in (68) and re-arranging it gives

$$\begin{aligned} \text{var} \left\{ \frac{R}{S} \right\} &\approx \frac{M\Gamma^2(M)}{\Gamma^2(M + \frac{1}{2})} + \left[\frac{\beta_{\text{ul}} p_k M^2 \Gamma^2(M)}{(\beta_{\text{ul}} p_k + N_0 + P\nu^2) \Gamma^2(M + \frac{1}{2})} \times \right. \\ &\quad \left. \left(\frac{M\Gamma^2(M)}{\Gamma^2(M + \frac{1}{2})} - \frac{1}{M} - 1 \right) \right] \\ &= \delta + \sigma_{\mathbf{h}_k}^2 M \delta \left(\delta - \frac{M+1}{M} \right), \end{aligned} \quad (69)$$

where $\delta = \frac{M\Gamma^2(M)}{\Gamma^2(M + \frac{1}{2})}$ and $\sigma_{\mathbf{h}_k}^2 = (\beta_{\text{ul}} p_k)(\beta_{\text{ul}} p_k + N_0 + P\nu^2)^{-1}$. This completes the proof. ■

REFERENCES

- [1] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 128–137, May 2015.
- [2] D. Nguyen, L. N. Tran, P. Pirinen, and M. Latva-aho, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896–4910, Sep. 2014.
- [3] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.
- [4] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.
- [5] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4381–4393, Dec. 2012.
- [6] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Sep. 2012.
- [7] Y. Deng, K. J. Kim, T. Q. Duong, M. Elkashlan, G. K. Karagiannis, and A. Nallanathan, "Full-duplex spectrum sharing in cooperative single carrier systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 2, no. 1, pp. 68–82, March 2016.
- [8] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.
- [9] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.
- [10] P. Lioliou, M. Viberg, M. Coldrey, and F. Athley, "Self-interference suppression in full-duplex MIMO relays," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2010, pp. 658–662.
- [11] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.
- [12] Z. Zhang, K. Long, A. V. Vasilakos, and L. Hanzo, "Full-duplex wireless communications: Challenges, solutions, and future research directions," *Proc. IEEE*, vol. 104, no. 7, pp. 1369–1409, Feb. 2016.
- [13] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3933–3938, Aug. 2009.
- [14] Y. Y. Kang and J. H. Cho, "Capacity of MIMO wireless channel with full-duplex amplify-and-forward relay," in *Proc. IEEE PIMRC*, Tokyo, Japan, Sep. 2009, pp. 117–121.
- [15] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.
- [16] —, "Residual self-interference in full-duplex MIMO relays after null-space projection and cancellation," in *Proc. ASILOMAR*, Pacific Grove, CA, USA, Nov 2010, pp. 653–657.
- [17] D. Nguyen, L. N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038–4050, Aug. 2013.
- [18] H. H. M. Tam, H. D. Tuan, and D. T. Ngo, "Successive convex quadratic programming for quality-of-service management in full-duplex MU-MIMO multicell networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2340–2353, June 2016.
- [19] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2205–2217, April 2015.
- [20] A. C. Cirik, Y. Rong, and Y. Hua, "Achievable rates of full-duplex MIMO radios in fast fading channels with imperfect channel estimation," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3874–3886, Aug. 2014.
- [21] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Trans. Wireless Communications*, vol. 12, no. 5, pp. 2498–2511, May 2013.
- [22] A. C. Cirik, R. Wang, Y. Rong, and Y. Hua, "Mse-based transceiver designs for full-duplex MIMO cognitive radios," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2056–2070, June 2015.
- [23] Y. Liao, L. Song, Z. Han, and Y. Li, "Full duplex cognitive radio: a new design paradigm for enhancing spectrum usage," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 138–145, May 2015.
- [24] A. Shojaeifard, K. K. Wong, M. D. Renzo, K. A. Hamdi, and J. Tang, "Design and analysis of full-duplex massive MIMO cellular networks," in *Proc. IEEE Globecom*, Washington, DC, Dec 2016, pp. 1–6.
- [25] S. Akbar, Y. Deng, A. Nallanathan, M. Elkashlan, and G. K. Karagiannis, "Massive multiuser MIMO in heterogeneous cellular networks with full duplex small cells," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4704–4719, Nov 2017.
- [26] H. H. M. Tam, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "MIMO energy harvesting in full-duplex multi-user networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3282–3297, May 2017.
- [27] V. D. Nguyen, T. Q. Duong, H. D. Tuan, O. S. Shin, and H. V. Poor, "Spectral and energy efficiencies in full-duplex wireless information and power transfer," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2220–2233, May 2017.
- [28] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.
- [29] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, April 2003.
- [30] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct 1998.
- [31] X. Du, J. Tadrous, and A. Sabharwal, "Sequential beamforming for multiuser MIMO with full-duplex training," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8551–8564, Dec. 2016.
- [32] X. Du, J. Tadrous, C. Dick, and A. Sabharwal, "MIMO broadcast channel with continuous feedback using full-duplex radios," in *Proc. Asilomar Conf. Signals, Sys. and Comp.*, Pacific Grove, CA, USA, Nov. 2014, pp. 1701–1705.
- [33] —, "MU-MIMO beamforming with full-duplex open-loop training," in *Proc. SPAWC*, Stockholm, Sweden, June 2015, pp. 301–305.
- [34] M. Kobayashi, N. Jindal, and G. Caire, "Training and feedback optimization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.
- [35] Y. Zhang, S. Han, C. Yang, and G. Wang, "Energy efficient optimization for full-duplex assisted closed-loop MISO downlink transmission," in *Proc. PIMRC*, Valencia, Spain, Sep. 2016, pp. 1–6.
- [36] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y. S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: pushing the limits," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 80–87, Sep. 2016.
- [37] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference

cancellation algorithms for in-band full duplex relays,” *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 91–101, May 2015.

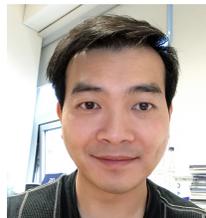
- [38] J. Arnau and M. Kountouris, “Timely CSI acquisition exploiting full duplex,” in *Proc. IEEE WCNC*, San Francisco, CA, USA, March 2017, pp. 1–6.
- [39] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, “Channel estimation and linear precoding in multiuser multiple-antenna TDD systems,” *IEEE Trans. Veh. Tech.*, vol. 60, no. 5, pp. 2102–2116, Jun 2011.
- [40] H. Yang and T. L. Marzetta, “Performance of conjugate and zero-forcing beamforming in large-scale antenna systems,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
- [41] A. Khansefid and H. Minn, “Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination,” *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4849–4864, Dec. 2015.
- [42] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots,” in *Proc. Annual Allerton Conf. Commun., Control, and Computing*, Monticello, IL, USA, Oct. 2013, pp. 293–298.
- [43] M. Chung, M. S. Sim, J. Kim, D. K. Kim, and C. b. Chae, “Prototyping real-time full duplex radios,” *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 56–63, Sep. 2015.
- [44] S. Hong, J. Brand, J. I. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, “Applications of self-interference cancellation in 5G and beyond,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 114–121, Feb. 2014.
- [45] G. Noh, H. Wang, C. Shin, S. Kim, Y. Jeon, H. Shin, J. Kim, and I. Kim, “Enabling technologies toward fully LTE-compatible full-duplex radio,” *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 188–195, March 2017.
- [46] D. Bharadia, E. McMillin, and S. Katti, “Full duplex radios,” in *Proc. ACM SIGCOMM*, vol. 43, no. 4, Hong Kong, China, Aug. 2013, pp. 375–386.
- [47] T. T. Do, H. Q. Ngo, T. Q. Duong, T. J. Oechtering, and M. Skoglund, “Massive MIMO pilot retransmission strategies for robustification against jamming,” *IEEE Wireless Commun. Lett.*, vol. 6, no. 1, pp. 58–61, Feb 2017.
- [48] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [49] N. Jindal, “MIMO broadcast channels with finite-rate feedback,” *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [50] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Springer Science & Business Media, 2013.
- [51] D. A. Basnayaka, P. J. Smith, and P. A. Martin, “Performance analysis of macrodiversity MIMO systems with MMSE and ZF receivers in flat Rayleigh fading,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2240–2251, May 2013.
- [52] L. Yu, W. Liu, and R. Langley, “SINR analysis of the subtraction-based SMI beamformer,” *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5926–5932, Nov. 2010.
- [53] Q. Zhang, S. Jin, K. K. Wong, H. Zhu, and M. Matthaiou, “Power scaling of uplink massive MIMO systems with arbitrary-rank channel means,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct 2014.
- [54] P. Jung, P. W. Baier, and A. Steil, “Advantages of CDMA and spread spectrum techniques over FDMA and TDMA in cellular mobile radio applications,” *IEEE Trans. Veh. Tech.*, vol. 42, no. 3, pp. 357–364, Aug 1993.
- [55] R. Hunger, “Floating point operations in matrix-vector calculus,” *Munich University of Technology, Inst. for Circuit Theory and Signal Processing*, Tech. Rep. Ver. 1.3, 2007.
- [56] G. Van Kempen and L. Van Vliet, “Mean and variance of ratio estimators used in fluorescence ratio imaging,” *Cytometry Part A*, vol. 39, no. 4, pp. 300–305, 2000.
- [57] H. Seltman, “Approximations for mean and variance of a ratio,” 2012. [Online]. Available: <http://www.stat.cmu.edu/~hseltman/files/ratio.pdf>



Jawad Mirza (S’12-M’16) received the B.Sc. degree in electrical (telecommunications) engineering from COMSATS Institute of Information Technology, Islamabad, Pakistan, in 2007; the M.Sc. degree in communications engineering from the University of Manchester, Manchester, U.K., in 2009; and the Ph.D. degree from Victoria University of Wellington, Wellington, New Zealand, in 2016. He has worked as a Postdoctoral Research Associate in the Wolfson School Mechanical, Electrical and Manufacturing Engineering at Loughborough University, UK. He is currently an Assistant Professor in the Department of Electrical Engineering, COMSATS Institute of Information Technology, Islamabad, Pakistan. His research interests include massive multiple-input multiple-output and full-duplex communications.



Gan Zheng (S’05-M’09-SM’12) received the BEng and the MEng from Tianjin University, Tianjin, China, in 2002 and 2004, respectively, both in Electronic and Information Engineering, and the PhD degree in Electrical and Electronic Engineering from The University of Hong Kong in 2008. He is currently a Senior Lecturer in the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, UK. His research interests include UAV communications, edge caching, full-duplex radio, wireless power transfer, cooperative communications, cognitive radio and physical-layer security. He is the first recipient for the 2013 IEEE Signal Processing Letters Best Paper Award, and he also received 2015 GLOBECOM Best Paper Award. He currently serves as an Associate Editor for IEEE Communications Letters.



Kai-Kit Wong (M’01-SM’08-F’16) received the BEng, the MPhil, and the PhD degrees, all in Electrical and Electronic Engineering, from the Hong Kong University of Science and Technology, Hong Kong, in 1996, 1998, and 2001, respectively. After graduation, he took up academic and research positions at the University of Hong Kong, Lucent Technologies, Bell-Labs, Holmdel, the Smart Antennas Research Group of Stanford University, and the University of Hull, UK. He is Chair in Wireless Communications at the Department of Electronic and Electrical Engineering, University College London, UK.

His current research centers around 5G and beyond mobile communications, including topics such as massive MIMO, full-duplex communications, millimetre-wave communications, edge caching and fog networking, physical layer security, wireless power transfer and mobile computing, V2X communications, and of course cognitive radios. There are also a few other unconventional research topics that he has set his heart on, including for example, fluid antenna communications systems, remote ECG detection and etc. He is a co-recipient of the 2013 IEEE Signal Processing Letters Best Paper Award and the 2000 IEEE VTS Japan Chapter Award at the IEEE Vehicular Technology Conference in Japan in 2000, and a few other international best paper awards.

He is Fellow of IEEE and IET and is also on the editorial board of several international journals. He has served as Senior Editor for IEEE Communications Letters since 2012 and also for IEEE Wireless Communications Letters since 2016. He had also previously served as Associate Editor for IEEE Signal Processing Letters from 2009 to 2012 and Editor for IEEE Transactions on Wireless Communications from 2005 to 2011. He was also Guest Editor for IEEE JSAC SI on virtual MIMO in 2013 and currently Guest Editor for IEEE JSAC SI on physical layer security for 5G.



Sangarapillai Lambotharan is Professor of Digital Communications and the Head of Signal Processing and Networks Research Group in the Wolfson School Mechanical, Electrical and Manufacturing Engineering at Loughborough University, UK. He received the PhD degree in Signal Processing from Imperial College London, UK in 1997, where he remained until 1999 as a postdoctoral research associate. He was a visiting scientist at the Engineering and Theory Centre of Cornell University, USA in 1996. Between 1999 and 2002, he was with

Motorola Applied Research Group, UK and investigated various projects including physical link layer modelling and performance characterization of GPRS, EGPRS and UTRAN. He was with Kings College London and Cardiff University as a lecturer and a senior lecturer respectively from 2002 to 2007. His current research interests include 5G networks, MIMO, radars, smart grids, machine learning, network security and convex optimizations and game theory. He has published approximately 200 technical journal and conference articles in these areas.



Lajos Hanzo (<http://www-mobile.ecs.soton.ac.uk>) FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of

Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 112 PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1751 research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Council's Advanced Fellow Grant and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE ComSoc and VTS. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to <http://www-mobile.ecs.soton.ac.uk>