Performance Analysis of FDD Massive MIMO Systems Under Channel Aging

Ribhu Chopra, Chandra R. Murthy, Himal A. Suraweera and Erik G Larsson

Chopra, R., Murthy, C. R., Suraweera, H. A., Larsson, E. G, (2018), Performance Analysis of FDD Massive MIMO Systems Under Channel Aging, *IEEE Transactions on Wireless Communications*, 17(2), 1094-1108. https://doi.org/10.1109/TWC.2017.2775629

Original publication available at:
https://doi.org/10.1109/TWC.2017.2775629

Performance Analysis of FDD Massive MIMO Systems under Channel Aging

Ribhu Chopra, *Member, IEEE*, Chandra R. Murthy, *Senior Member, IEEE*, Himal A. Suraweera, *Senior Member, IEEE*, and Erik G. Larsson, *Fellow, IEEE*

***Abstract*—In this paper, we study the effect of channel aging on the uplink and downlink performance of an FDD massive MIMO system, as the system dimension increases. Since the training duration scales linearly with the number of transmit dimensions, channel estimates become increasingly outdated in the communication phase, leading to performance degradation. To quantify this degradation, we first derive bounds on the mean squared channel estimation error. We use the bounds to derive deterministic equivalents of the receive SINRs, which yields a lower bound on the achievable uplink and downlink spectral efficiencies. For the uplink, we consider maximal ratio combining and MMSE detectors, while for the downlink, we consider matched filter and regularized zero forcing (RZF) precoders. We show that the effect of channel aging can be mitigated by optimally choosing the frame duration. It is found that using all the base station antennas can lead to negligibly small achievable rates in high user mobility scenarios. Finally, numerical results are presented to validate the accuracy of our expressions and illustrate the dependence of the performance on the system dimension and channel aging parameters.**

***Index Terms*—Massive MIMO, channel aging, channel estimation, performance analysis, achievable rate.**

I. Introduction

Multiple-input multiple-output (MIMO) with a large number of base station antennas serving multiple users, popularly known as massive MIMO, is a key enabling technology for next generation wireless communications [1]–[3]. A singular feature of these systems is the phenomenon of channel hardening due to large dimensions [4], that leads to quasi orthogonality among different channel vectors [5]–[7]. This quasi-orthogonality reduces the inter-stream interference, allowing the use of simplified transmitter and receiver architectures. Further, the array gain increases linearly with the number of base station antennas [7], [8], leading to increased spectral and energy efficiencies. However, the above advantages of massive MIMO rely heavily on the availability of accurate and up to date channel state information (CSI) at the base station and the users. In practice, the CSI at the base station is imperfect because of channel estimation errors [9]–[11], pilot contamination [10] and is also outdated due to channel aging [12].

Channel aging is caused by the time varying nature of the channel between the cellular users and the base station, which is in turn a consequence of user mobility [12], [13]. Contrary to the conventional block fading channel model, an aging channel evolves continuously with time, and is different during each transmitted symbol. This results in a mismatch between the actual channel state and the CSI acquired during training, which could degrade the performance of a massive MIMO system [13]. More importantly, since the minimum training duration for acquiring CSI scales linearly with the number of antennas, the mismatch gets exacerbated with increasing system dimensionality. As a consequence, having more antennas at the base station may in fact lead to poor performance in the presence of channel aging.

Most of the current literature [12]–[17], starting from [12], focuses on the effects of channel aging for time division duplexed (TDD) massive MIMO systems. In general, and the performance loss due to channel aging is quantified using the achievable rate of the system under study. All of the present works on channel aging consider linear receivers such as the maximal ratio combining (MRC) receiver [12] and the minimum mean squared error (MMSE) receiver [14] in the uplink, and linear precoders such as the matched filter precoder (MFP) [12] and the regularized zero forcing (RZF) [13] precoder in the downlink. Linear receivers and precoders are mainly chosen due to their simplicity and ease of analysis. The achievable rates have generally been derived using a deterministic equivalent (DE) of the signal to interference plus noise ratios (SINRs), and then using the DE-SINR to compute the achievable rates. The authors in [13] also derive the non asymptotic achievable rates in addition to the more conventional DE analysis. The results derived in these papers confirm that although the benefits of massive MIMO such as power scaling [18] are still valid in the presence of channel aging, it does indeed lead to a significant loss in the user performance, and its effect is accentuated by an increase in user mobility [19]. The effect of channel aging on TDD massive MIMO systems have been also been studied in conjunction with phase noise [13] and pilot contamination [20].

However, the extension of current massive MIMO techniques to an FDD setting has been identified as an open challenge in [18]. The study of the effect of channel aging

R. Chopra is with the Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati, Assam, India. (email: ribhufec@iitg.ernet.in).

C. R. Murthy is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India. (email: cmurthy@ece.iisc.ernet.in).

H. A. Suraweera is with the Department of Electrical and Electronic Engineering, University of Peradeniya, Peradeniya, Sri Lanka (email: himal@ee.pdn.ac.lk).

E. G. Larsson is with the Department of Electrical Engineering (ISY), Linkoping University, Sweden (email: erik.g.larsson@liu.se).

This work was financially supported in part by a research grant from the Ministry of Electronics and Information Technology, Govt. of India.

Parts of this work were supported by the Swedish Research Foundation (VR) and ELLIIT.

in an FDD setting is also important because it is easier for the current cellular systems to upgrade to massive MIMO in an FDD system, rather than switch to TDD [18], [21], [22]. Transmission schemes and channel estimation for FDD based massive MIMO have been discussed in [23] and [21], respectively. Contrary to the TDD massive MIMO model, channel reciprocity cannot be assumed in an FDD massive MIMO system. Consequently, forward link training is required for both uplink and downlink, making the problem of channel aging more pronounced when the base station is equipped with a large number of antennas. The effect of channel aging on the downlink of a beamforming based single user large MIMO system has been examined in [24]. In [24], the minimum achievable rate is optimized in terms of the number of transmit antennas and the frame duration. It is shown that, it is not always beneficial to use a larger number of base station antennas, and the overall throughput of such a system may degrade with an increase in the number of transmit antennas.

Most of the current literature on the effects of channel aging on massive MIMO systems assumes no variation in the channel during the training interval [12]–[14], [16]. While this assumption simplifies the determination of channel estimation error variances, that are needed to compute the achievable rates for different systems, it is unrealistic since the channel will continue to age during the training phase, and the quality of estimates of different users will be different due to their different training instants. Kalman filter based estimators for aging channels have been derived in [25] but the effects of estimation errors on the achievable data rates have not been discussed.

In this paper, we consider linear channel estimation, precoding, and receive combining, in a single cell multiuser FDD massive MIMO setting. We first characterize the performance of the MMSE channel estimators for both the uplink and downlink channels, and derive bounds on the quality of the channel estimates at the end of the training duration. We then use these results to obtain lower bounds on the achievable rate in both uplink and downlink. We study the behavior of the MRC and the MMSE receivers for the uplink, and the MF and the RZF precoders in the downlink. The main contributions of this work can be listed as follows:

1) We derive bounds on the mean squared channel estimation error for both the uplink and the downlink channels, accounting for the effect of channel aging during the training interval. We observe that the performance of channel estimation saturates in the presence of channel aging, and the saturated value of the MSE normalized with respect to the mean squared channel gain for downlink channel estimation can be as large as -3 dB for user velocities of the order of 100 km/h. (See Section III.)
2) We use DE analysis along with the derived bounds on the channel estimation errors to characterize the per user achievable rates in the uplink for the MRC and the MMSE receivers. Increased user mobilities are observed to reduce the achievable rate by more than a factor of two in an MMSE receiver based system with 150 cellular users. (See Section IV.)
3) We derive the DEs of the per user achievable rate in the downlink for the MFP and RZF precoders using the bounds on the mean squared channel estimation errors, and show that using a larger number of transmit antennas may not lead to improved data rates under channel aging conditions. For example, in a 100 user system with a base station equipped with 1000 antennas, using all the antennas at user velocities beyond 200 km/h may lead to a negligibly small achievable rate, even after optimizing the frame duration based on the channel aging parameters. (See Section IV.)
4) Via detailed simulations, we prescribe the optimal values of different parameters such as frame duration, number of base station antennas used, etc. under different channel aging conditions. We observe that the optimal frame duration for a system with high user mobility is a function of the number of users.

The key findings of this work are that channel aging severely affects both channel estimation and achievable rates in an FDD massive MIMO system. It is therefore not always better to use a larger number of antennas at the base station. We have thus identified new tradeoffs due to channel aging that need to be taken into account while designing massive MIMO systems.

II. Time Varying Channel Model

Our system model, as illustrated in Fig. 1, considers a single cell system with a base station having N antennas communicating with $K < N$ single antenna users indexed by $k \in \{1, \ldots, K\}$. We assume an FDD system with training for channel estimation in both the uplink and the downlink. In the uplink frame, each user first transmits orthogonal pilot symbols to the base station, where these are used to estimate the underlying channels. Following this, all the users simultaneously transmit their data to the base station. The data is detected at the base station using the CSI acquired during the pilot transmission phase. During the downlink training phase, the base station transmits orthogonal pilot symbols from its antennas which are used by the users to estimate the corresponding downlink channel. These channel estimates are then fed back to the base station by the users, and are used by the latter to appropriately precode the data symbols transmitted simultaneously to all the users during the data transmission phase. We assume an ideal delay and error free feedback channel for ease of analysis and also to isolate the effect of channel aging and estimation error on the system performance.

The uplink channel between the kth user and the ith base station antenna at the nth instant is modeled as $\sqrt{\beta_k} h_{ik}[n]$, with β_k representing the (large scale) slow fading and path loss component, and $h_{ik}[n]$ representing the fast fading component modeled as a zero mean circularly symmetric complex Gaussian (ZMCSCG) random variable (rv) with unit variance, denoted by $\mathcal{CN}(0,1)$. Similarly, the downlink channel from the ith base station antenna to the kth user can be expressed as $\sqrt{\beta_k} f_{ki}[n]$ with $f_{ki}[n]$ representing the $\mathcal{CN}(0,1)$ distributed fast fading component. Note that the slow fading and path loss component is assumed to be the same in the uplink and downlink channels, and the same across all base station antennas.

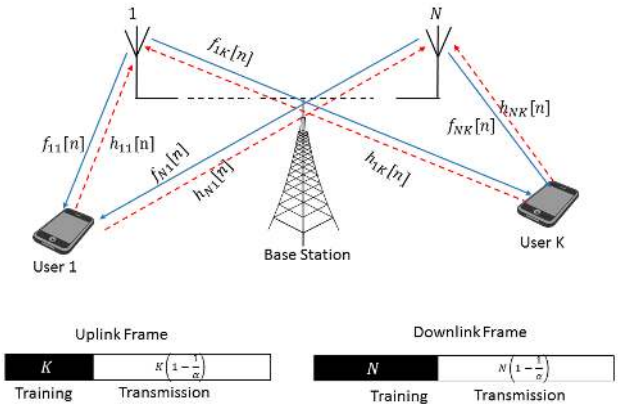


Fig. 1: The system model.

Defining $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_K]^T$, we can write the uplink channel matrix at the nth instant as $\mathbf{H}[n]\text{diag}(\sqrt{\boldsymbol{\beta}})$, with $\mathbf{H}[n] \in \mathbb{C}^{N\times K}$, with the entries of $\mathbf{H}[n]$ being $\mathcal{CN}(0,1)$, and diag$(\boldsymbol{\beta})$ representing a $K \times K$ diagonal matrix such that $(\text{diag}(\boldsymbol{\beta}))_{kk} = \beta_k$. Similarly, the downlink channel is expressed as diag$(\sqrt{\boldsymbol{\beta}})\mathbf{F}[n]$ with $\mathbf{F}[n] \in \mathbb{C}^{K\times N}$.

The temporal variations in the propagation environment, caused due to user mobility, results in the channel coefficients evolving with time. The channel evolution can be modeled as a function of its initial state and an innovation component, such that the channel between the ith base station antenna and the kth user at the nth instant is [12], [26]

$$\begin{aligned} h_{ik}[n] &= \rho[n]h_{ik}[0] + \bar{\rho}[n]z_{h,ik}[n] \\ f_{ki}[n] &= \rho[n]f_{ki}[0] + \bar{\rho}[n]z_{f,ki}[n], \end{aligned} \quad (1)$$

where $z_{h,ik}[n] \sim \mathcal{CN}(0,1)$ and $z_{f,ki}[n] \sim \mathcal{CN}(0,1)$ are respectively the innovation components for the uplink and downlink channels, $0 \leq \rho[n] \leq 1$ is the correlation between the channel instances at lags 0 and n, and $\bar{\rho}[n] = \sqrt{1-|\rho[n]|^2}$. The channel innovation processes, $z_{h,ik}[n]$ and $z_{f,ki}[n]$, need not be temporally white. Their statistics are chosen such that the second-order statistics of the model in (1) match with the temporal correlation of the channel. Specifically, if $\mathbf{C_h}$ denotes the temporal covariance matrix of the instances of one of the channel entries at times 1 through $(T-1)$, the $(T-1)\times(T-1)$ covariance $\mathbf{R}_z$ of an entry of the innovation vector sequence is given by $\mathbf{R}_z = \mathbf{D}^{-1}(\mathbf{C_h} - \underline{\rho}\,\underline{\rho}^T)\mathbf{D}^{-1}$, where $\underline{\rho} \triangleq [\rho[1], \rho[2], \ldots, \rho[T-1]]^T$ and $\mathbf{D} \triangleq \text{diag}(\bar{\rho}[1], \ldots, \bar{\rho}[T-1])$. Since channel aging is mainly caused due to relative motion between the transmitter and the receiver, the temporal channel correlation coefficient can be assumed to be the same for the uplink and the downlink channels. It is conventional to assume that the channel evolves according to the Jakes' model [12], [27], or as a first order autoregressive (AR1) process [24]. If the channel is assumed to age according to the Jakes' model [27], $\rho[n] = J_0(2\pi f_d T_s n)$, where f_d is the Doppler frequency, T_s is the sampling period, and $J_0(\cdot)$ is the Bessel function of the first kind and zeroth order [28, Eq. (9.1.18)]. Also, each entry of the channel vector is temporally as $[\mathbf{C_h}]_{mk} = J_0(2\pi(m-k)f_d T_S)$. Alternatively, if we consider the channel to evolve as an AR1 model, $\rho[n] = \rho^n$, with ρ being $J_0(2\pi f_d T_s)$, and $\mathbf{C_h}$ is a symmetric Toeplitz matrix with $[1, \rho, \rho^2, \ldots, \rho^{T-1}]$ as its first row. We assume $\rho[n]$ to be identical across all users for simplicity. Note that, the model in (1) differs from the AR1 model (or more generally, the AR(p) model) commonly used in the literature in that the channel state is determined recursively in terms of its state at the previous time instant, instead of its state at an initial time $h_{ik}[0]$. Furthermore, the coefficients of the AR(p) model are chosen such that the statistics of the channel approximately match with that of the Jakes' model. This is an advantage of the model in (1), namely, that the statistics of the innovation processes can be chosen such that the statistics of the model exactly match with that of the Jakes' model within a finite-time window of interest. Also, since the effects of phase noise discussed in [13] can be incorporated into the channel autocorrelation function, we do not discuss these explicitly, and use a generalized correlation coefficient $\rho[n]$.

III. Channel Estimation

In this section, we consider MMSE estimation of the uplink and the downlink channels, and provide expressions for the MSE for channel estimation in both the cases. We use these expressions to show that the performance of the estimator saturates due to channel aging.

A. Uplink Channels

Since there are K users transmitting to an N antenna base station in the uplink, the minimal number of uplink pilot symbols for channel estimation at the base station is K [29]. Hence, the signal received at the ith base station antenna at the nth instant can be written as,

$$y_i[n] = \sum_{k=1}^{K} \sqrt{\beta_k \mathcal{E}_{u,p,k}} h_{ik}[n]\psi_k[n] + \sqrt{N_0}w_i[n], \quad (2)$$

where $\mathcal{E}_{u,p,k}$ is the pilot energy of the kth user, N_0 is the noise variance, and $w_i[0] \sim \mathcal{CN}(0,1)$.

The quality of the channel estimate will worsen with an increase in the delay between training and transmission, hence, we consider the channel estimates at $n = K+1$, as the channel estimates at later instants will only be worse. The effective channel at the nth $(n \leq K)$ instant can be expressed in terms of the channel at the $(K+1)$th instant as

$$h_{ik}[n] = \rho^*[K+1-n]h_{ik}[K+1] + \bar{\rho}[K+1-n]\zeta_{h,ik}[n]. \quad (3)$$

Consequently, (2) can be rewritten as

$$\begin{aligned} y_i[n] &= \sum_{m=1}^{K} \sqrt{\beta_m \mathcal{E}_{u,p,m}} \rho^*[K+1-n]\psi_m[n]h_{im}[K+1] \\ &+ \sum_{m=1}^{K} \sqrt{\beta_m \mathcal{E}_{u,p,m}} \bar{\rho}[K+1-n]\psi_m[n]\zeta_{h,im}[n] + \sqrt{N_0}w_i[n], \end{aligned} \quad (4)$$

with $\psi_m[n]$ corresponding to the pilot signal transmitted by the mth user at the nth time instant. For channel estimation,

we consider a weighted combination of the received training signals with weights $c_k[n]$ chosen to match with the training sequences, $\psi_m[n]$ attenuated by the aging component, i.e.,

$$\sum_{n=1}^{K} c_k[n]\rho^*[K+1-n]\psi_m[n] = \delta[k-m], \tag{5}$$

where $\delta[\cdot]$ is the Kronecker delta function. Thus, the combined received signal can be written as

$$\begin{aligned} u_{ik} &= \sqrt{N_0}\sum_{n=1}^{K} c_k[n]w_i[n] \\ &+ \sum_{m=1}^{K}\sqrt{\beta_m\mathcal{E}_{u,p,m}}h_{im}[K+1]\sum_{n=1}^{K}\rho^*[K+1-n]c_k[n]\psi_m[n] \\ &+ \sum_{m=1}^{K}\sqrt{\beta_m\mathcal{E}_{u,p,m}}\sum_{n=1}^{K}\bar{\rho}[K+1-n]c_k[n]\psi_m[n]\zeta_{h,im}[n]. \end{aligned} \tag{6}$$

If $c_k[n]$ and $\psi_m[n]$ satisfy the orthogonality condition in (5), which is a commonly used design criterion in pilot design for channel estimation, and further letting $\psi_m[n] = \delta[n-m]$ and $c_k[n] = (\rho^*[K-n])^{-1}\delta[n-k]$ (which satisfies (5)), we get

$$\begin{aligned} u_{ik} &= \sqrt{\beta_k\mathcal{E}_{u,p,k}}h_{ik}[K+1] + \rho^{-1*}[K+1-k]\sqrt{N_0}w_i[k] \\ &+ \sqrt{\beta_k\mathcal{E}_{u,p,k}}\rho^{-1*}[K+1-k]\bar{\rho}[K+1-k]\zeta_{h,ik}[k]. \end{aligned} \tag{7}$$

In the above, the first term corresponds to the channel coefficient of interest, and the others are additive noise terms. Since the coefficients $h_{ik}[K+1]$ are i.i.d., they can be individually estimated from the corresponding u_{ik}. The MMSE estimate, $\hat{h}_{ik}$ of $h_{ik}[K+1]$ is given by

$$\hat{h}_{ik} = \rho[K+1-k]\sqrt{\frac{1}{\beta_k\mathcal{E}_{u,p,k}+N_0}}u_{ik}. \tag{8}$$

Therefore, the estimation error variance, denoted as $\sigma^2_{h,ik}$, becomes

$$\sigma^2_{h,ik} = \frac{|\bar{\rho}[K+1-k]|^2\beta_k\mathcal{E}_{u,p,k}+N_0}{\beta_k\mathcal{E}_{u,p,k}+N_0}. \tag{9}$$

Now, $\sigma^2_{h,ik}$ can be bounded as

$$\frac{|\bar{\rho}[1]|^2\beta_k\mathcal{E}_{u,p,k}+N_0}{\beta_k\mathcal{E}_{u,p,k}+N_0} \le \sigma^2_{h,ik} \le \frac{|\bar{\rho}[K]|^2\beta_k\mathcal{E}_{u,p,k}+N_0}{\beta_k\mathcal{E}_{u,p,k}+N_0}. \tag{10}$$

This holds for all sequences $c_k[n]$ and $\psi_m[n]$ satisfying the orthogonality condition (5).

Letting $\gamma_{u,p,k} \triangleq \frac{\beta_k\mathcal{E}_{u,p,k}}{N_0}$, we can write an upper bound on the MSE for channel estimation as

$$\sigma^2_{h,ik} \le \frac{|\bar{\rho}[K]|^2\gamma_{u,p,k}+1}{\gamma_{u,p,k}+1}. \tag{11}$$

Now, looking at the MSE in the absence of channel aging, $\sigma'^2_{h,ik} = \frac{1}{\gamma_{u,p,k}+1}$, we observe that in the absence of channel aging, the estimation error disappears as the pilot signal-to-noise ratio (SNR), $\gamma_{u,p,k}$ grows. However, in the presence of channel aging from (9),

$$\lim_{\gamma_{u,p,k}\to\infty}\sigma^2_{h,ik} = |\bar{\rho}[K+1-k]|^2. \tag{12}$$

That is, the channel estimation error saturates, and the limiting MSE worsens as the number of users increases. Thus, aging adversely affects the quality of the channel estimates. We next discuss the effects of aging on downlink channel estimation.

B. Downlink Channels

Similar to the uplink, the N antennas at the base station transmit a pilot sequence $\{\boldsymbol{\psi}[n]\}_{n=1}^{N}$ with energies $\mathcal{E}_{d,p}$. The signal received at the kth user at the nth instant is

$$y_k[n] = \sum_{i=1}^{N}\sqrt{\beta_k\mathcal{E}_{d,p}}f_{ki}[n]\psi_i[n] + \sqrt{N_0}w_k[n], \tag{13}$$

where $\psi_i[n]$ is the pilot symbol transmitted by the ith base station antenna at the nth instant. We are interested in estimating the channel between the kth user and the ith base station antenna, which is related to the channel at the nth instant, $\mathbf{f}_{ki}[n]$, via $f_{ki}[n] = \rho^*[N+1-n]f_{ki}[N+1] + \bar{\rho}[N+1-n]\zeta_{d,ki}[n]$. Therefore,

$$\begin{aligned} y_k[n] = \sum_{i=1}^{N}\sqrt{\beta_k\mathcal{E}_{d,p}}(&\rho^*[N+1-n]f_{ki}[N+1] \\ &+ \bar{\rho}[N+1-n]\zeta_{d,ki}[n])\psi_i[n] + \sqrt{N_0}w_k[n]. \end{aligned} \tag{14}$$

Defining $u_{ki} \triangleq \sum_{n=1}^{N} c_i[n]y_k[n]$, where $c_i[n]$ is chosen such that $\sum_{n=1}^{N} c_i[n]\rho^*[N+1-n]\psi_m[n] = \delta[i-m]$, we have,

$$\begin{aligned} u_{ki} &= \sqrt{\beta_k\mathcal{E}_{d,p}}f_{ki}[N+1] + \sqrt{N_0}\sum_{n=1}^{N} c_i[n]w_k[n] \\ &+ \sum_{m=1}^{N}\sqrt{\beta_m\mathcal{E}_{d,p}}\sum_{n=1}^{N}\bar{\rho}[N+1-n]c_i[n]\psi_m[n]\zeta_{d,km}[n]. \end{aligned} \tag{15}$$

Using[1] $\psi_m[n] = \delta[n-m]$ and $c_i[n] = \rho^{-1*}[N+1-n]\delta[n-i]$, we can calculate the MMSE estimate $\hat{f}_{ki}$ of $f_{ki}[N+1]$ as

$$\hat{f}_{ki} = \rho[N+1-i]\sqrt{\frac{1}{\beta_k\mathcal{E}_{d,p}+N_0}}u_{ki}. \tag{16}$$

The MSE for the channel estimator is then given by

$$\sigma^2_{d,ki} = \frac{\bar{\rho}^2[N+1-i]\beta_k\mathcal{E}_{d,p}+N_0}{\beta_k\mathcal{E}_{d,p}+N_0}. \tag{17}$$

Writing the pilot SNR at the kth user as $\gamma_{d,p,k} \triangleq \beta_k\mathcal{E}_{d,p}$, it can be shown that

$$\lim_{\gamma_{d,p,k}\to\infty}\sigma^2_{d,ki} = |\bar{\rho}[N+1-i]|^2. \tag{18}$$

[1]For simplicity, we assume that orthogonal training sequences of minimal length are used for estimating the uplink and downlink channels. A different choice of orthogonal pilots, where all users' pilot sequences occupy all K training instants, might result in the aging affecting all the users equally. However, designing an optimal pilot sequences for aging channels is beyond the scope of this work. Our aim here is to derive bounds on the mean squared estimation error and achievable rate, under a reasonable choice of the pilot sequence.

Therefore, the saturation in the MSE due to channel aging can be observed both in the uplink and the downlink. However, the effect will be greater in the downlink since the number of base station antennas is greater than the number of users in the cell. Also, in case the different channel coefficients are correlated, a Kalman Filter based estimator similar to the one discussed in [25] can be developed. Since the use of a Kalman filter based estimator will only improve the estimation performance, an upper bound on the mean squared estimation error will still be

$$\sigma^2_{d,ki} \leq \frac{\bar{\rho}^2[N]\beta_k\mathcal{E}_{d,p} + N_0}{\beta_k\mathcal{E}_{d,p} + N_0}. \tag{19}$$

In the following sections, use these bounds on the channel estimation error to derive the achievable uplink and downlink rates for different cases.

IV. Uplink Data Rate Analysis

In this section, we use DE analysis to obtain the achievable rates at each user for MRC and MMSE receivers, and then compute the minimum achievable rates using the upper bounds on error variance derived in the Section III. We assume that the users transmit uplink data during the instants $K+1 \leq n \leq T_c$, with T_c being the total frame duration. Let the kth user transmit the signal $s_k[n]$ with a symbol energy $\mathcal{E}_{u,s,k}$ at the nth instant. Then, the signal received at the ith base station antenna can be written as

$$y_i[n] = \sum_{k=1}^{K} \sqrt{\beta_k\mathcal{E}_{u,s,k}}h_{ik}[n]s_k[n] + w_i[n]. \tag{20}$$

However,

$$h_{im}[n] = \rho[n-K]b_k\hat{h}_{im} + \bar{b}_k\rho[n-K]\tilde{h}_{im} + \bar{\rho}[n-K]z_{h,im}[n], \tag{21}$$

where, $b_k = \sqrt{\frac{|\rho[K]|^2\beta_k\mathcal{E}_{u,p,k}}{\beta_k\mathcal{E}_{u,p,k}+N_0}}$, and $\bar{b}_k = \sqrt{1-b_k^2}$. It is to be noted that b_k and $\bar{b}_k$ correspond to the worst case channel estimation error. Substituting (21) into (20), we obtain the signal at the base station as

$$\begin{aligned}\mathbf{y}[n] = {} & \rho[n-K]\sqrt{\beta_k\mathcal{E}_{u,s,k}}b_k\hat{\mathbf{h}}_k s_k[n] \\ & + \rho[n-K]\sqrt{\beta_k\mathcal{E}_{u,s,k}}\bar{b}_k\tilde{\mathbf{h}}_k s_k[n] \\ & + \sqrt{\beta_k\mathcal{E}_{u,s,k}}\bar{\rho}[n-K]\mathbf{z}_{h,k}[n]s_k[n] \\ & + \sum_{\substack{m=1\\m\neq k}}^{K}\sqrt{\beta_m\mathcal{E}_{u,s,m}}\mathbf{h}_m[n]s_m[n] + \sqrt{N_0}\mathbf{w}[n],\end{aligned} \tag{22}$$

where $\mathbf{h}_k[n]$, $\hat{\mathbf{h}}_k$, $\tilde{\mathbf{h}}_k$, $\mathbf{z}_{h,k}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$, and represent the vector channel from the base station to the kth user at the nth instant, the channel estimate available at the base station, the estimation error, and the innovation component in the channel at the nth instant, respectively.

Multiplying $\mathbf{y}[n]$ with the receiver matrix $\mathbf{V}^H[n] \in \mathbb{C}^{K\times N}$, we can define the processed receive signal vector $\mathbf{r}[n] \in \mathbb{C}^{K\times 1}$ as

$$\mathbf{r}[n] = \mathbf{V}^H[n]\mathbf{y}[n]. \tag{23}$$

Now, the kth component of $\mathbf{r}[n]$ corresponds to the symbol transmitted by the kth user, and can be expanded as

$$\begin{aligned}r_k[n] = {} & \rho[n-K]\sqrt{\beta_k\mathcal{E}_{u,s,k}}b_k\mathbf{v}_k^H[n]\hat{\mathbf{h}}_k s_k[n] \\ & + \sum_{\substack{m=1\\m\neq k}}^{K}\rho[n-K]\sqrt{\beta_m\mathcal{E}_{u,s,m}}\mathbf{v}_k^H[n]\hat{\mathbf{h}}_m s_m[n] \\ & + \sum_{m=1}^{K}\rho[n-K]\bar{b}_m\sqrt{\beta_m\mathcal{E}_{u,s,m}}\mathbf{v}_k^H[n]\tilde{\mathbf{h}}_m s_m[n] \\ & + \sum_{m=1}^{K}\sqrt{\beta_m\mathcal{E}_{u,s,m}}\bar{\rho}[n-K]\mathbf{v}_k^H[n]\mathbf{z}_{h,m}[n]s_m[n] + \sqrt{N_0}\mathbf{v}_k^H[n]\mathbf{w}[n] \\ = {} & \rho[n-K]\sqrt{\beta_k\mathcal{E}_{u,s,k}}b_k\mathbf{v}_k^H[n]\hat{\mathbf{h}}_k s_k[n] \\ & + \sum_{\substack{m=1\\m\neq k}}^{K}\rho[n-K]\sqrt{\beta_m\mathcal{E}_{u,s,m}}\mathbf{v}_k^H[n]\mathbf{h}_m s_m[n] \\ & + \rho[n-K]\bar{b}_k\sqrt{\beta_k\mathcal{E}_{u,s,k}}\mathbf{v}_k^H[n]\tilde{\mathbf{h}}_k s_k[n] \\ & + \sqrt{\beta_k\mathcal{E}_{u,s,k}}\bar{\rho}[n-K]\mathbf{v}_k^H[n]\mathbf{z}_{h,k}[n]s_k[n] + \sqrt{N_0}\mathbf{v}_k^H[n]\mathbf{w}[n].\end{aligned}$$

In both versions of the above equation, the first term corresponds to the desired signal component, and all other terms correspond to noise and interference caused by channel estimation errors, channel aging and the data transmission by other users. We will use both these versions of the above equation for the analysis to follow. In the following two subsections, we use the above expressions, along with DE analysis to derive the SINRs for $\mathbf{V}[n]$ corresponding to the MRC and the MMSE receivers.

A. The MRC Receiver

For the MRC receiver, $\mathbf{V}[n] = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \ldots, \hat{\mathbf{h}}_K]$, therefore $r_k[n]$ can be re-arranged as

$$\begin{aligned}r_k[n] = {} & \rho[n-K]\sqrt{\beta_k\mathcal{E}_{u,s,k}}b_k\hat{\mathbf{h}}_k^H\hat{\mathbf{h}}_k s_k[n] \\ & + \sum_{\substack{m=1\\m\neq k}}^{K}\rho[n-K]\sqrt{\beta_m\mathcal{E}_{u,s,m}}\hat{\mathbf{h}}_k^H\mathbf{h}_m[n]s_m[n] \\ & + \rho[n-K]\sqrt{\beta_k\mathcal{E}_{u,s,k}}\bar{b}_m\hat{\mathbf{h}}_k^H\tilde{\mathbf{h}}_k s_m[n] \\ & + \sqrt{\beta_k\mathcal{E}_{u,s,k}}(\bar{\rho}[n-K])\hat{\mathbf{h}}_k^H\mathbf{z}_{h,k}[n]s_k[n] + \sqrt{N_0}\hat{\mathbf{h}}_k^H\mathbf{w}[n].\end{aligned} \tag{24}$$

Therefore, the SINR for the kth stream can be expressed as

$$\begin{aligned}\eta_{u,k}^{\text{MRC}}[n] = {} & \left(|\rho[n-K]|^2\beta_k b_k^2 E\left[|\hat{\mathbf{h}}_k^H\hat{\mathbf{h}}_k|^2\right]\mathcal{E}_{u,s,k}\right) \\ & \times\Bigg(|\rho[n-K]|^2\beta_k\bar{b}_k^2 E\left[|\hat{\mathbf{h}}_k^H\tilde{\mathbf{h}}_k|^2\right]\mathcal{E}_{u,s,k} \\ & + \beta_k(|\bar{\rho}[n-K]|^2)E\left[|\hat{\mathbf{h}}_k^H\mathbf{z}_{h,k}[n]|^2\right]\mathcal{E}_{u,s,k} \\ & + \sum_{\substack{m=1\\m\neq k}}^{K}\beta_m E\left[|\hat{\mathbf{h}}_k^H\mathbf{h}_m[n]|^2\right]\mathcal{E}_{u,s,m} + N_0 E\left[|\hat{\mathbf{h}}_k^H\mathbf{w}[n]|^2\right]\Bigg)^{-1}.\end{aligned} \tag{25}$$

Now, each of $\hat{\mathbf{h}}_k$, $\tilde{\mathbf{h}}_k$, $\mathbf{h}_m$, and $\mathbf{w}[n]$ are independent with i.i.d. $\mathcal{CN}(0,1)$ entries. Therefore, for large system dimensions, we can use results from random matrix theory to approximate the expected values in (25) using their corresponding DEs [30]. Note that, in the large antenna regime, the instantaneous SINR and the average SINR are the same with high probability, and can be well approximated using DEs. We can therefore use the results from [12, Lemma 1] to simplify (25) as shown in equation (26)

We now assume that the users employ path loss based uplink power control, that is, each user scales its data power in inverse proportionality to the long term path loss β_k. Consequently, the pilot SNR at the base station becomes $\gamma_{u,p} = \frac{\beta_k \mathcal{E}_{u,p,k}}{N_0}$, and

$$b_k = b = \sqrt{\frac{|\rho[K]|^2}{1+\gamma_{u,p}^{-1}}}\ k = 1, 2 \ldots, K. \tag{27}$$

Defining $\gamma_{u,s} \triangleq \frac{\beta_k \mathcal{E}_{u,s,k}}{N_0}$, the SINR for each user at the nth instant can be expressed as

$$\eta_u^{\text{MRC}}[n] - \frac{|\rho[n-K]|^2 b^2 N}{|\rho[n-K]|^2 \bar{b}^2 + (|\bar{\rho}[n-K]|^2) + (K-1) + \gamma_{u,s}^{-1}} \xrightarrow{a.s.} 0. \tag{28}$$

It is important to note that the user SINR after accounting for channel aging is a decreasing function of the symbol index, and therefore, the spectral efficiencies achievable by different symbols in a frame are no longer constant. We use a different codebook for each symbol in the frame (and each codeword spans across a large number of frames) [31]. Therefore, for a frame duration T_u^{MRC}, each user transmits from $T_u^{\text{MRC}} - K$, one for each time index, with each codebook spanning over data/pilot blocks. In this case, the average spectral efficiency per user can be expressed as

$$R_{u,\text{var}}^{\text{MRC}} = \frac{1}{T_u^{\text{MRC}}} \sum_{n=K+1}^{T_u^{\text{MRC}}} \log_2\left(1+\eta_u^{\text{MRC}}[n]\right). \tag{29}$$

Defining $\alpha_u^{\text{MRC}} \triangleq \frac{K}{T_u^{\text{MRC}}}$, the above equation becomes

$$R_{u,\text{var}}^{\text{MRC}} = \frac{\alpha_u^{\text{MRC}}}{K} \sum_{n=K+1}^{\frac{K}{\alpha_u^{\text{MRC}}}} \log_2\left(1+\eta_u^{\text{MRC}}[n]\right). \tag{30}$$

In (29), the parameter α_u^{MRC} is the ratio of training duration to the total usable duration of a channel. Large values of α_u^{MRC} will result in higher values of SINRs increasing the argument of the log term in (30) . At the same time, these will result in more frequent training, thus reducing the number of summation terms in (30). On the other hand, smaller values of α_u^{MRC} will reduce the argument of the log term, but will allow for longer transmission durations. Consequently, α_u^{MRC} has to be optimized based on the number of users, and the properties of the channel. Since $0 < \alpha_u^{\text{MRC}} \leq 1$, the per user achievable rate can be numerically optimized in terms of α_u^{MRC}, by searching over the interval $(0, 1]$. We discuss the choice of α_u^{MRC} in detail in Section IV-C.

B. The MMSE Receiver

For an MMSE receiver, the matrix $\mathbf{V}[n]$ becomes $\mathbf{V}[n] = \mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}[n]\hat{\mathbf{H}}$, where $\mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}[n]$ is the covariance matrix of the received signal at the nth instant for the given channel estimate $\hat{\mathbf{H}}$. Therefore, $\mathbf{V}[n]$ has to be recomputed at each instant. We first use the knowledge of the statistics of the received signal to derive $\mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}[n]$ and $\mathbf{V}[n]$, and then use these to characterize the SINR performance of the MMSE receiver.

Defining $\mathbf{b} \triangleq [b_1, \ldots, b_K]^T$, $\bar{\mathbf{b}} \triangleq [\bar{b}_1, \ldots, \bar{b}_K]^T$ where b_k and $\bar{b}_k$ are defined after (21), and writing $\dot{\mathbf{D}} \triangleq \sqrt{\text{diag}(\boldsymbol{\beta})\text{diag}(\boldsymbol{\mathcal{E}}_{u,s})}$, $\mathbf{D} \triangleq \text{diag}(\mathbf{b})\dot{\mathbf{D}}$, and $\bar{\mathbf{D}} \triangleq \text{diag}(\bar{\mathbf{b}})\dot{\mathbf{D}}$, the vector equivalent of the received signal can be written as

$$\mathbf{y}[n] = \rho[n-K]\hat{\mathbf{H}}\mathbf{D}\mathbf{s}[n] + \rho[n-K]\tilde{\mathbf{H}}\bar{\mathbf{D}}\mathbf{s}[n] + \bar{\rho}[n-K]\mathbf{Z}_h\dot{\mathbf{D}}\mathbf{s}[n] + \sqrt{N_0}\mathbf{w}[n]. \tag{31}$$

Since the data transmitted by the different users are independent, $E[\mathbf{s}[n]\mathbf{s}^H[n]] = \mathbf{I}_K$. Also, the channel estimation error and the transmitted data are independent [7], [29], consequently,

$$\mathbf{R}_{yy|\hat{\mathbf{H}}}[n] = |\rho[n-K]|^2\hat{\mathbf{H}}\mathbf{D}^2\hat{\mathbf{H}}^H + |\rho[n-K]|^2 E[\tilde{\mathbf{H}}\bar{\mathbf{D}}^2\tilde{\mathbf{H}}^H] + |\bar{\rho}[n-K]|^2 E[\mathbf{Z}_h\dot{\mathbf{D}}^2\mathbf{Z}_h^H] + N_0\mathbf{I}_N. \tag{32}$$

Since the entries of $\tilde{\mathbf{H}}$ and $\mathbf{Z}_h$ are i.i.d. zero mean complex Gaussian with unit variance, $E[\tilde{\mathbf{H}}\bar{\mathbf{D}}^2\tilde{\mathbf{H}}^H] = \mathbf{I}_N \sum_{k=1}^K \bar{b}_k^2\beta_k\mathcal{E}_{u,s,k}$, $E[\mathbf{Z}_h\dot{\mathbf{D}}^2{\mathbf{Z}_h}^H] = \mathbf{I}_N \sum_{k=1}^K \beta_k\mathcal{E}_{u,s,k}$, therefore (32) becomes

$$\mathbf{R}_{yy|\hat{\mathbf{H}}}[n] = |\rho[n-K]|^2\hat{\mathbf{H}}\mathbf{D}^2\hat{\mathbf{H}}^H + \epsilon[n]\mathbf{I}_N, \tag{33}$$

where

$$\epsilon[n] = |\rho[n-K]|^2 \sum_{k=1}^K \bar{b}_k^2\beta_k\mathcal{E}_{u,s,k} + |\bar{\rho}[n-K]|^2 \sum_{k=1}^K \beta_k\mathcal{E}_{u,s,k} + N_0. \tag{34}$$

The MMSE detector for the kth stream at the nth instant therefore becomes $\mathbf{v}_k[n] = \mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}\hat{\mathbf{h}}_k$, and hence

$$\begin{aligned} r_k[n] = {} & \rho[n-K]\sqrt{\beta_k\mathcal{E}_{u,s,k}}b_k\hat{\mathbf{h}}_k^H\mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}\hat{\mathbf{h}}_k s_k[n] \\ & + \sum_{\substack{m=1 \\ m\neq k}}^K \rho[n-K]b_m\sqrt{\beta_m\mathcal{E}_{u,s,m}}\hat{\mathbf{h}}_k^H\mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}\hat{\mathbf{h}}_m s_m[n] \\ & + \sum_{m=1}^K \rho[n-K]\bar{b}_m\sqrt{\beta_m\mathcal{E}_{u,s,m}}\hat{\mathbf{h}}_k^H\mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}\tilde{\mathbf{h}}_m s_m[n] \\ & + \sum_{m=1}^K \sqrt{\beta_m\mathcal{E}_{u,s,m}}(\bar{\rho}[n-K])\hat{\mathbf{h}}_k^H\mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}\mathbf{z}_{h,m}[n]s_m[n] \\ & + \sqrt{N_0}\mathbf{v}_k^H[n]\mathbf{w}[n]. \end{aligned} \tag{35}$$

In order to calculate the SINR, we need the variances of the individual terms in (35). Denoting the variance of the first term

$$\eta_{u,k}^{\text{MRC}}[n] - \frac{|\rho[n-K]|^2 b_k^2 \beta_k N \mathcal{E}_{u,s,k}}{|\rho[n-K]|^2 \beta_k \bar{b}_k^2 \mathcal{E}_{u,s,k} + \beta_k |\bar{\rho}[n-K]|^2 \mathcal{E}_{u,s,k} + \sum_{\substack{m=1 \\ m\neq k}}^{K} \beta_m \mathcal{E}_{u,s,m} + N_0} \xrightarrow{a.s.} 0. \tag{26}$$

of $r_k[n]$ as $\sigma_{1,k}^2[n]$, we get,

$$\begin{aligned} &\sigma_{1,k}^2[n] \\ &= b_k^2 \beta_k \mathcal{E}_{u,s,k} |\rho[n-K]|^2 E\left[\left|\hat{\mathbf{h}}_k^H \mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}[n] \hat{\mathbf{h}}_k\right|^2\right] E[|s_k[n]|^2]. \end{aligned} \tag{36}$$

Since the entries of $\hat{\mathbf{H}}$ are $\mathcal{CN}(0,1)$, therefore for large system dimensions, the expectation operation in (36) can be replaced with the DE. It is shown in Appendix A that using results from random matrix theory the DE of $\sigma_{1,k}^2[n]$ can be simplified as

$$\sigma_{1,k}^2[n] - \frac{N^2 b_k^2 \beta_k \mathcal{E}_{u,s,k} |\rho[n-K]|^2 \varphi_k^2[n]}{|1 + b_k^2 \beta_k \mathcal{E}_{u,s,k} |\rho[n-K]|^2 N \varphi_k[n]|^2} \xrightarrow{a.s} 0, \tag{37}$$

where

$$\varphi_k[n] \triangleq \left(|\rho[n-K]|^2 \sum_{\substack{m=1 \\ m\neq k}}^{K} \frac{b_m^2 \beta_m \mathcal{E}_{u,s,k}}{1 + e_{k,m}[n]} + \epsilon[n] \right)^{-1}, \tag{38}$$

and $e_{k,m}[n]$ is iteratively computed as

$$e_{k,m}^{(t)}[n] = \frac{|\rho[n-K]|^2 b_m^2 \beta_m \mathcal{E}_{u,s,m}}{|\rho[n-K]|^2 \sum_{i=1;i\neq k}^{K} \frac{b_i^2 \beta_i \mathcal{E}_{u,s,i}}{1+e_{k,i}^{(t-1)}[n]} + \epsilon[n]}, \tag{39}$$

with the initialization $e_{k,m}^{(0)}[n] = \frac{1}{\epsilon[n]}$.

Letting $\sigma_{2,k}^2[n]$ denote the variance of the second term of (35), we show in Appendix B that its DE can be calculated as

$$\sigma_{2,k}^2[n] - \frac{N \sum_{m=1 m\neq k}^{K} \mu_{k,m}[n]}{|1 + b_k^2 \beta_k \mathcal{E}_{u,s,k} |\rho[n-K]|^2 N \varphi_k[n]|^2} \xrightarrow{a.s.} 0, \tag{40}$$

where

$$\begin{aligned} \mu_{k,m} = \dot{\varphi}_{k,m}^2[n] + &\frac{|b_m^2 \beta_m \mathcal{E}_{u,s,k}|^2 |\rho[n-K]|^4 N^2 \dot{\varphi}_{k,m}^4[n]}{|1 + b_m^2 \beta_m \mathcal{E}_{u,s,m} |\rho[n-K]|^2 N \dot{\varphi}_{k,m}[n]|^2} \\ &- 2\Re\left\{ \frac{|b_m^2 \beta_m \mathcal{E}_{u,s,k}|^2 |\rho[n-K]|^2 N \dot{\varphi}_{k,m}^3}{1 + b_m \beta_m \mathcal{E}_{u,s,m} |\rho[n-K]|^2 N \dot{\varphi}_{k,m}[n]} \right\} \Bigg), \end{aligned} \tag{41}$$

$\Re\{.\}$ denotes the real part of a complex number,

$$\dot{\varphi}_{k,m}[n] = \left(|\rho[n-K]|^2 \sum_{l=1;l\neq m,k}^{K} \frac{b_l^2 \beta_l \mathcal{E}_{u,s,m}}{1 + \dot{e}_{k,m,l}[n]} + \epsilon[n] \right)^{-1}, \tag{42}$$

and $\dot{e}_{k,m,l}[n]$ is iteratively computed as

$$\dot{e}_{k,m,l}^{(t)}[n] = \frac{|\rho[n-K]|^2 b_l^2 \beta_l \mathcal{E}_{u,s,l}}{|\rho[n-K]|^2 \sum_{i=1;i\neq m,k}^{K} \frac{b_i^2 \beta_i \mathcal{E}_{u,s,i}}{1+\dot{e}_{k,m,i}^{(t-1)}[n]} + \epsilon[n]}, \tag{43}$$

such that $\dot{e}_{k,m,l}^{(0)}[n] = \frac{1}{\epsilon[n]}$.

Also, using [12, Lemma 1] and (73), it is easy to show that the DEs of the variances of the third, fourth, and fifth terms are respectively given as

$$\begin{aligned} &\sigma_{3,k}^2[n] \\ &-N|\rho[n-K]|^2 \frac{\sum_{m=1}^{K} \beta_m \mathcal{E}_{u,s,m} \bar{b}_m^2 \varphi_k^2[n]}{|1 + b_k \beta_k \mathcal{E}_{u,s,k} \rho[n-K]|^2 N \varphi_k[n]|^2} \xrightarrow{a.s.} 0, \end{aligned} \tag{44}$$

$$\begin{aligned} &\sigma_{4,k}^2[n] \\ &-N|\bar{\rho}[n-K]|^2 \frac{\sum_{m=1}^{K} \beta_m \mathcal{E}_{u,s,m} \varphi_k^2[n]}{|1 + b_k^2 \beta_k \mathcal{E}_{u,s,k} |\rho[n-K]|^2 N \varphi_k[n]|^2} \xrightarrow{a.s.} 0, \end{aligned} \tag{45}$$

and

$$\sigma_{5,k}^2[n] - N \frac{N_0}{|1 + b_k^2 \beta_k \mathcal{E}_{u,s,k} |\rho[n-K]|^2 N \varphi_k[n]|^2} \xrightarrow{a.s.} 0. \tag{46}$$

It can be observed that $\sigma_{1,k}^2[n]$, $\sigma_{2,k}^2[n]$ and $\sigma_{3,k}^2[n]$ are all decreasing functions of n, which implies that the signal power, the residual interference due to the canceled streams, and the interference due to channel estimation errors, will all reduce as the channel ages. On the other hand, $\sigma_{4,k}^2[n]$ is an increasing function of the time index, and corresponds to the cumulative interference from all the data streams due to channel aging. Therefore, the advantage offered by the MMSE receiver over the MRC receiver diminishes due to channel aging, and the effect of channel aging is more pronounced for the MMSE receiver compared to the MRC receiver.

Considering simple uplink power control such that $\beta_k \mathcal{E}_{u,s,k} = \mathcal{E}_{u,s}$, the matrix $\mathbf{D}$ becomes a scaled version of the identity matrix. It can then be shown via simple algebraic manipulation that $\varphi_k[n]$ and $\mu_k[n]$ are independent of the user indices, and can be denoted by $\varphi[n]$ and $\mu[n]$ respectively. This results in a per user SINR given by (47)

The per user SINR is again a decreasing function of time. Similar to the MRC case, we can use a different codebook for each symbol. Assuming the frame duration to be T_u^{MMSE}, and letting $\alpha_u^{\text{MMSE}} = \frac{K}{T_u^{\text{MMSE}}}$, the achievable spectral efficiency can be written as

$$R_{u,\text{var}}^{\text{MMSE}} = \frac{\alpha_u^{\text{MMSE}}}{K} \sum_{n=K+1}^{\frac{K}{\alpha_u^{\text{MMSE}}}} \log_2\left(1 + \eta_u^{\text{MMSE}}[n]\right). \tag{48}$$

The optimal value of the parameter α_u^{MMSE} can again be obtained by searching over the set $(0,1]$. This is discussed in greater detail, along with the choice of α_u^{MRC}, in the following subsection.

C. Numerical Results

Here, we present numerical results to demonstrate the effects of channel aging on the uplink of an FDD massive MIMO system. We consider a 1000 antenna base station with a carrier frequency (f_c) of 2 GHz. The signal bandwidth is

$$\eta_u^{\text{MMSE}}[n] - \frac{|\rho[n-K]|^2 N b^2}{(K-1)|\rho[n-K]|^2 b^2 \frac{\mu[n]}{\varphi^2[n]} + K\bar{b}^2|\rho[n-K]|^2 + K(1-|\rho[n-K]|^2) + \gamma_{u,s}^{-1}} \xrightarrow{a.s.} 0. \tag{47}$$

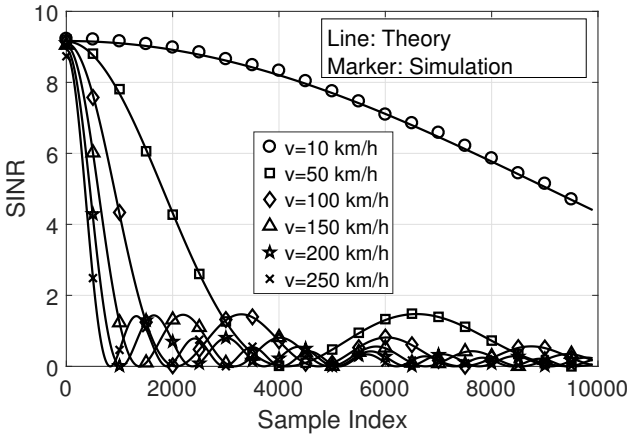
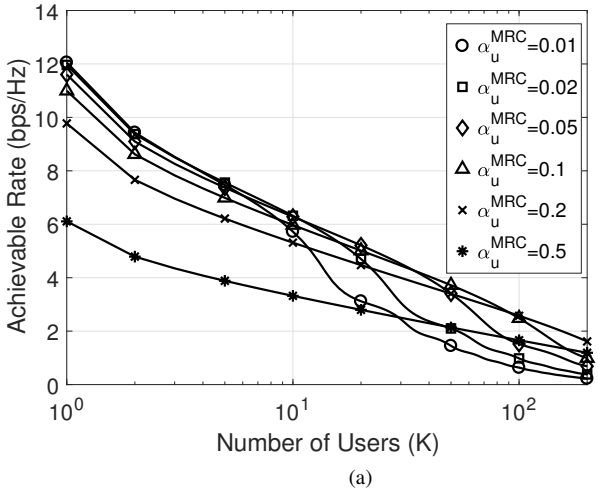
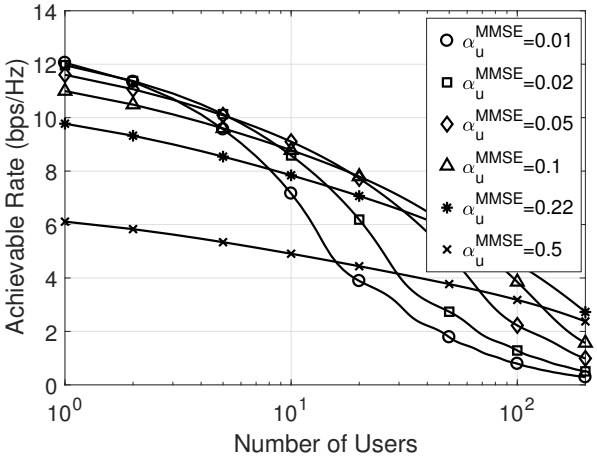


Fig. 2: Per user SINRs at different time instants for different user mobilities, and data and pilot SNR of 10 dB.



(a)



(b)

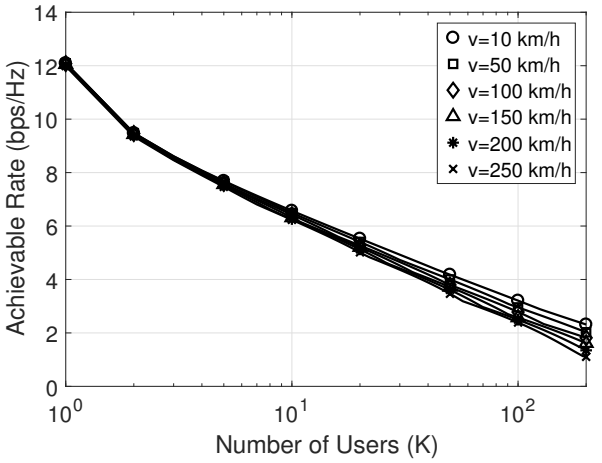
Fig. 3: Achievable uplink rate as a function of the number of users for a user speed of 150 km/h and data and pilot SNR of 10 dB for (a) MRC and (b) MMSE receivers.

assumed to be 1 MHz, and the base station is assumed to sample at the Nyquist rate of the complex baseband signal, rate, i.e., at 1 MHz. The channel is assumed to age according to the Jakes model, i.e. $\rho[n] = J_0(2\pi f_d T_s n)$.

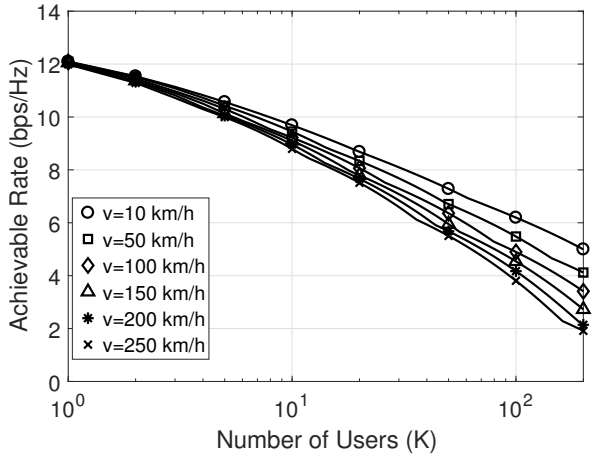
In Fig. 2, we compare the per user SINRs in (28) against its simulated values for a 100 user system at different user velocities. The data and pilot SNR in this case is assumed to be 10 dB. It can be observed that the achievable SINR at a given sample index, reduces with the user mobility, which is as expected. Also, the ripples observed in the SINR are caused due to its dependence on the correlation coefficient, which in this case is assumed to take the form of the Bessel function of first order. It is also to be noted that since there is a close match between the derived and simulated values, the former can be used to accurately study the behavior of the latter.

In Figs. 3a and 3b, we plot the per user achievable rates for MRC and MMSE based decoders as a function of the number of users for different values of α_u^{MRC}. These plots assume data and pilot SNRs of 10 dB and a user velocity of 150 km/h. The per user achievable rates drop with an increasing number of users due to the increased interference. The increased number of users also results in more severe aging of the available estimates which further adds to the interference issue. The optimal value of α_u^{MRC} depends on the number of users and the pilot and data SNRs. Its value is determined using a line search over the interval $[0, 0.5]$. The slight irregularities in the otherwise monotonic decrease with the number of users is due to the fact that the SINR at any given instant is dependent on $J_0(2\pi f_d T_s n)$. The ripples in the Bessel function cause the SINR, and hence the achievable rate, to become a non-monotonic function of the sample index. Thus, the total achievable rate per frame is not a monotonic function of the frame duration. Since we optimize the overall frame duration depending on the number of antennas/users, the achievable rate per frame becomes a non-monotonic function of the number of antennas/users as well. However, the per user achievable rate for the MMSE receiver is more sensitive to the value of α_u^{MMSE} as compared to the MRC receiver. Therefore, a suitable choice of the frame duration is necessary in order to fully exploit the interference canceling abilities of the MMSE receiver.

In Figs. 4a and 4b, we plot the per user maximum achievable rate (maximized over α_u^{MRC}) against the number of users for different user velocities at a data and pilot SNR of 10 dB. Here, even though there is a dramatic influence of the velocity on the SINR, as observed in Fig. 2, its effect on the optimized rate is limited. This is because the system adapts to



(a)



(b)

Fig. 4: Achievable uplink rate optimized over (a) α_u^{MRC} and (b) α_u^{MMSE} as a function of the number of users for different user speeds, and data and pilot SNR of 10 dB.

higher mobility by reducing the frame duration and optimally choosing the frame duration to achieve nearly the same rate. It is observed that the MMSE decoder results in larger per user rates as compared to the MRC decoder for slow moving users, but the advantage offered by MMSE decoder over an MRC decoder (see Fig. 4a) reduces with an increase in the user mobility. This is in accordance with the observations made from (47).

V. Downlink Data Rates

In this section, we derive the DEs of the downlink SINRs as seen by the cellular users, with the matched filtering (MF) and the RZF precoders being used at the base stations [12], [14]. The base station transmits data symbols during the time instants $N+1 \leq n \leq T_d$. It is also assumed the noisy and aged CSI available at the users is made available to the base station at time $(N+1)$ through an ideal feedback channel. Letting $s_k[n]$ be the symbol transmitted by the base station to the kth user at the nth time instant with an energy $\mathcal{E}_{d,s,k}$ and using a precoder vector $\mathbf{p}_k[n] \in \mathbb{C}^{N\times 1}$, $\mathbf{s}[n] = [s_1[n], s_2[n], \ldots, s_K[n]]^T$, $\mathbf{P}[n] \triangleq [\mathbf{p}_1[n], \ldots, \mathbf{p}_K[n]]$ and $\boldsymbol{\mathcal{E}}_{d,s} = [\mathcal{E}_{d,s,1}, \ldots, \mathcal{E}_{d,s,K}]^T$, the symbol transmitted by the base station becomes

$$\mathbf{x}[n] = \mathbf{P}[n]\text{diag}\left(\sqrt{\frac{\boldsymbol{\mathcal{E}}_{d,s}}{N}}\right)\mathbf{s}[n]. \tag{49}$$

Consequently, the signal received by the kth user at the nth instant can be expressed as

$$y_k[n] = \sqrt{\beta_k}\mathbf{f}^{(k)}[n]\mathbf{P}[n]\text{diag}\left(\sqrt{\frac{\boldsymbol{\mathcal{E}}_{d,s}}{N}}\right)\mathbf{s}[n] + w_k[n]. \tag{50}$$

Again, letting $b_{d,k} = \sqrt{\frac{|\rho[N]|^2\beta_k\mathcal{E}_{d,p}}{\beta_k\mathcal{E}_{d,p}+N_0}}$, and $\bar{b}_{d,k} = \sqrt{1-b_{d,k}^2}$, we can describe the channel from the kth user to the base station as the row vector $\mathbf{f}^{(k)}$,

$$\mathbf{f}^{(k)}[n] = \rho[n-N]b_{d,k}\hat{\mathbf{f}}^{(k)} + \rho[n-N]\bar{b}_{d,k}\tilde{\mathbf{f}}^{(k)} + \bar{\rho}[n-N]\mathbf{z}_f^{(k)}[n], \tag{51}$$

where $\hat{\mathbf{f}}^{(k)}$, $\tilde{\mathbf{f}}$, and $\mathbf{z}_f$ are defined respectively as the channel estimate, the channel estimation error, and the innovation component. Substituting (51) into (50) and simplifying, we get

$$\begin{aligned} y_k[n] &= \rho[n-N]b_{d,k}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,k}}{N}}\hat{\mathbf{f}}^{(k)}\mathbf{p}_k[n]s_k[n] \\ &\quad + \sum_{\substack{m=1\\m\neq k}}^{K}\rho[n-N]b_{d,k}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,m}}{N}}\mathbf{f}^{(k)}\mathbf{p}_m[n]s_m[n] \\ &\quad + \rho[n-N]\bar{b}_{d,k}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,k}}{N}}\tilde{\mathbf{f}}^{(k)}\mathbf{p}_k[n]s_k[n] \\ &\quad + \rho[n-N]\sqrt{\frac{\beta_k\mathcal{E}_{d,s,k}}{N}}\mathbf{z}_f^{(k)}[n]\mathbf{p}_k[n]s_k[n] + \sqrt{N_0}w_k[n]. \end{aligned} \tag{52}$$

Again, the first term is the desired signal power, and all the other terms contribute to noise and interference. In the following subsections, we use techniques similar to those in Section IV to derive the DEs of the SINRs for different designs of the matrix $\mathbf{P}[n]$.

A. The MF Precoder

The matched filter precoder takes the form, $\mathbf{P}[n] = \hat{\mathbf{F}}^H$, such that $\mathbf{p}_k[n] = \hat{\mathbf{f}}^{(k)H}$, consequently, the received signal can be expressed as

$$\begin{aligned} y_k[n] &= \rho[n-N]b_{d,k}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,k}}{N}}\hat{\mathbf{f}}^{(k)}\hat{\mathbf{f}}^{(k)H}s_k[n] \\ &\quad + \sum_{\substack{m=1\\m\neq k}}^{K}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,m}}{N}}\mathbf{f}^{(m)}\hat{\mathbf{f}}^{(k)H}s_m[n] \\ &\quad + \rho[n-N]\bar{b}_{d,k}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,k}}{N}}\tilde{\mathbf{f}}^{(k)}\hat{\mathbf{f}}^{(k)H}s_k[n] \\ &\quad + \bar{\rho}[n-N]\sqrt{\frac{\beta_k\mathcal{E}_{d,s,k}}{N}}\mathbf{z}_k^{(k)}[n]\hat{\mathbf{f}}^{(k)H}s_k[n] + \sqrt{N_0}w_k[n]. \end{aligned} \tag{53}$$

Similar to the uplink case discussed in Section IV, the variances of the individual terms of $y_k[n]$ can be approximated using their DEs. Therefore, the DE of the SINR at the kth user can be computed using [12, Lemma 1], and can be expressed as (54).

Note that the pilot signals are common across users while the data signals are transmitted with power control. In this work, we consider two types of downlink data power control, viz, equal power allocation, and channel inversion power control. In the first scheme, the signals of all the users are transmitted with equal powers such that $\mathcal{E}_{d,s,k} = \mathcal{E}_{d,s}$, and consequently, the downlink SINR can be expressed as It is to be noted that both β_k and b_k are decreasing functions of distance of the user from the base station, and therefore the achievable SINR will be the minimum for a user at the cell edge. For channel inversion based power control, each user is allocated a data power in inverse proportionality to its long term fading coefficient β_k. Hence, the SINR is given by

$$\eta_{d,k}^{\text{MFP}}[n] - \frac{N|\rho[n-N]|^2 b_{d,k}^2}{\sum_{\substack{m=1\\m\neq k}}^{K} \beta_m^{-1} + |\rho[n-N]|^2 \bar{b}_{d,k}^2 + |\bar{\rho}[n-N]|^2 + \gamma_{d,s}^{-1}} \xrightarrow{a.s.} 0, \tag{56}$$

with $\gamma_{d,s} = \frac{\beta_k \mathcal{E}_{d,s}}{N_0}$. In this case, the users closer to the base station will see increased interference from the signals meant for the user located at a greater distance from the base station. Therefore, power control for downlink massive MIMO channels in the presence of channel aging is a nontrivial problem, and is an interesting direction for future work.

Another important observation from (54)-(56) is that with forward link training, the minimum achievable SINR at a user is no longer a monotonically increasing function of N, the number of base station antennas. In addition to this, a larger number of base station antennas also increases the training duration, and reduces the effective usable time of the channel for a given frame duration. As demonstrated in the previous section, the frame duration that maximizes the achivable rate is a function of the number of base station antennas, the channel aging characteristics, and the number of users. Consequently, we can optimize the minimum achievable rate by each user by selecting a subset of the antennas from the total number of available antennas and appropriately choosing the frame duration.

Considering equal power allocation for simplicity, we can write the achievable throughput with the base station using a different codebook at each instant as

$$R_{d,\min}^{\text{MFP}} = \min_k \frac{1}{T_d^{\text{MFP}}} \sum_{n=N+1}^{T_d^{\text{MFP}}} \log_2\left(1 + \eta_{d,k}^{\text{MFP}}[T_d^{\text{MFP}}]\right). \tag{57}$$

Now, since $\eta_{d,k}$ is minimized for the farthest user, whose distance from the base station can be approximated as the cell radius. Defining $\alpha_d^{\text{MFP}} \triangleq \frac{M}{T_d^{\text{MFP}}}$ as the ratio of the number of antennas used (M) to the total frame duration we can write the problem of selecting the optimal number of base station antennas and frame duration as (58), where $\beta_{\min} = \min_k \beta_k$, $0 < \alpha_d^{\text{MFP}} \leq 1$, $M \leq N$, $b_{d,min} = \sqrt{\frac{|\rho[N-1]|^2 \beta_{\min}\mathcal{E}_{d,p}}{\beta_{\min}\mathcal{E}_{d,p}+N_0}}$ and $\gamma_{d,\min} = \frac{\beta_{\min}\mathcal{E}_{d,s}}{N_0}$.

Here, the optimal values of M and α_d^{MFP} can be obtained by searching over the intervals $M \in 1, \ldots, N$ and $\alpha_d^{\text{MFP}} \in (0, 1)$. This is discussed in greater detail in the section on numerical results. We next derive the DEs for SINR in case of an RZF precoder.

B. The RZF Precoder

Here, the precoding matrix can be expressed as

$$\mathbf{P}[n] = \mathbf{Q}_{\hat{\mathbf{F}}}^{-1}[n]\mathbf{F}^H = (|\rho[n-K]|^2 \hat{\mathbf{F}}\mathbf{D}_d^2\hat{\mathbf{F}}^H + d[n]\mathbf{I}_N)^{-1}\hat{\mathbf{F}}^H, \tag{59}$$

with $d[n]$ being the regularization parameter at the nth instant, and $\mathbf{D}_d = \text{diag}(\sqrt{\boldsymbol{\mathcal{E}}_{d,s}})\text{diag}(\mathbf{b})\text{diag}(\sqrt{\boldsymbol{\beta}})$, and $\mathbf{b} \triangleq [b_1, \ldots, b_K]^T$. Hence,

$$\mathbf{p}_k[n] = \mathbf{Q}_{\hat{\mathbf{F}}}^{-1}[n]\mathbf{f}^{(k)H}. \tag{60}$$

Using this, we can write (52) as

$$\begin{aligned} y_k[n] &= \rho[n-N]b_{d,k}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,k}}{N}}\hat{\mathbf{f}}^{(k)}\mathbf{Q}_{\hat{\mathbf{F}}}^{-1}[n]\hat{\mathbf{f}}^{(k)H}s_k[n] \\ &+ \sum_{\substack{m=1\\m\neq k}}^{K} b_m\sqrt{\frac{\beta_k\mathcal{E}_{d,s,m}}{N}}\hat{\mathbf{f}}^{(m)}\mathbf{Q}_{\hat{\mathbf{F}}}^{-1}[n]\hat{\mathbf{f}}^{(k)H}s_m[n] \\ &+ \sum_{m=1}^{K}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,m}}{N}}\rho[n-N]\bar{b}_{d,m}\tilde{\mathbf{f}}^{(m)}\mathbf{Q}_{\hat{\mathbf{F}}}^{-1}[n]\hat{\mathbf{f}}^{(k)H}s_m[n] \\ &+ \bar{\rho}[n-N]\sum_{m=1}^{K}\sqrt{\frac{\beta_k\mathcal{E}_{d,s,m}}{N}}\mathbf{z}^{(m)}\mathbf{Q}_{\hat{\mathbf{F}}}^{-1}[n]\hat{\mathbf{f}}^{(k)H}s_m[n] + \sqrt{N_0}w_k[n]. \end{aligned} \tag{61}$$

Following steps similar to those in Appendix A, we can now show that the DE of the variance of the desired signal can be written as

$$\sigma_{1,k}^2[n] - \frac{|\rho[n-N]|^2 b_{d,k}^2\beta_k\mathcal{E}_{d,s,k}N\varphi_k^2[n]}{|1 + b_k^2\mathcal{E}_{d,s,k}\beta_k|\rho[n-N]|^2 N\varphi_k[n]|^2} \xrightarrow{a.s.} 0, \tag{62}$$

where

$$\varphi_k[n] = \left(\frac{|\rho[n-N]|^2}{N}\sum_{\substack{m=1\\m\neq k}}^{K}\frac{b_m^2\beta_m\mathcal{E}_{d,s,m}}{1+e_m[n]} + d[n]\right)^{-1}, \tag{63}$$

and $e_{k,m}^{(t)}[n]$ is computed using iterative equation below,

$$e_{k,m}^{(t)}[n] = \left(\frac{|\rho[n-N]|^2 b_m\beta_k\mathcal{E}_{d,s,m}}{\frac{|\rho[n-K]|^2}{N}\sum_{i=1;i\neq k}^{K}\frac{b_i\beta_i\mathcal{E}_{d,s,i}}{1+e_{k,i}^{(t-1)}[n]} + d[n]}\right), \tag{64}$$

and is initialized as $e_{k,m}^{(0)}[n] = \frac{1}{d[n]}$. Similarly, the residual interference power after RZF precoding becomes

$$\sigma_{2,k}^2[n] - |\rho[n-N]|^2\frac{\beta_k\sum_{m\neq k}b_m^2\mu_{k,m}[n]\mathcal{E}_{d,s,m}}{|1 + b_k^2\beta_k\mathcal{E}_{d,s,k}|\rho[n-N]|^2\varphi_k[n]|^2} \xrightarrow{a.s.} 0. \tag{65}$$

$$\eta_{d,k}^{\text{MFP}}[n] - \frac{N|\rho[n-N]|^2 b_{d,k}^2 \beta_k \mathcal{E}_{d,s,k}}{\beta_k \sum_{m=1;m\neq k}^{K} \mathcal{E}_{d,s,m} + |\rho[n-N]|^2 \bar{b}_{d,k}^2 \beta_k \mathcal{E}_{d,s,k} + |\bar{\rho}[n-N]|^2 \beta_k \mathcal{E}_{d,s,k} + N_0} \xrightarrow{a.s.} 0 \tag{54}$$

$$\eta_{d,k}^{\text{MFP}}[n] - \frac{N|\rho[n-N]|^2 b_{d,k}^2 \beta_k \mathcal{E}_{d,s}}{(K-1)\beta_k \mathcal{E}_{d,s} + |\rho[n-N]|^2 \bar{b}_{d,k}^2 \beta_k \mathcal{E}_{d,s} + |\bar{\rho}[n-N]|^2 \beta_k \mathcal{E}_{d,s} + N_0} \xrightarrow{a.s.} 0. \tag{55}$$

$$\max_{M,\alpha_d^{\text{MFP}}} R_{d,\min}^{\text{MFP}} = \frac{M}{\alpha_d^{\text{MFP}}} \sum_{n=N+1}^{\frac{\alpha_d^{\text{MFP}} M}{M}} \log_2 \left(1 + \frac{M|\rho[n-M]|^2 b_{d,\min}^2}{(K-1) + |\rho[n-M]|^2 \bar{b}_{d,\min}^2 + |\bar{\rho}[n-M]|^2 + \gamma_{d,\min}^{-1}}\right) \tag{58}$$

where

$$\mu_{k,m}[n] = \dot{\varphi}_{k,m}^2[n] + \frac{|b_m^2 \beta_k \mathcal{E}_{d,s,m}|^2 |\rho[n-N]|^4 \dot{\varphi}_{k,m}^4[n]}{|1+\beta_k \mathcal{E}_{d,s,m} b_m^2 |\rho[n-N]|^2 \dot{\varphi}_{k,m}[n]|^2} - 2\Re\left\{\frac{|b_m^2 \beta_k \mathcal{E}_{d,s,m}||\rho[n-N]|^2 \dot{\varphi}_{k,m}^3}{1+|b_m^2 \beta_k \mathcal{E}_{d,s,m} \rho[n-N]|\dot{\varphi}_{k,m}[n]}\right\}\Bigg), \tag{66}$$

$$\dot{\varphi}_{k,m}[n] = \left(\frac{|\rho[n-N]|^2}{N} \sum_{l=1;l\neq m,k}^{K} \frac{b_l^2 \beta_l \mathcal{E}_{d,s,m}}{1+\dot{e}_{k,m,i}[n]} + d[n]\right)^{-1}, \tag{67}$$

and $\dot{e}_{k,m,l}[n]$ is iteratively computed as

$$\dot{e}_{k,m,l}^{(t)}[n] = \left(\frac{\frac{|\rho[n-N]|^2}{N} b_l^2 \beta_l \mathcal{E}_{d,s,l}}{\frac{|\rho[n-N]|^2}{N} \sum_{i=1;i\neq m,k}^{K} \frac{b_i^2 \beta_i \mathcal{E}_{d,s,i}}{1+\dot{e}_{k,m,i}[n]} + d[n]}\right), \tag{68}$$

and is initialized as $\dot{e}_{k,m,l}^{(t)}[n] = \frac{1}{d[n]}$. Also, the variances of the interference components due to channel estimation errors and channel aging can respectively be computed as

$$\sigma_{3,k}^2[n] - |\rho[n-N]|^2 \frac{\beta_k \varphi_k^2[n] \sum_{m=1}^{K} \bar{b}_m^2 \mathcal{E}_{d,s,m}}{|1+|\rho[n-N]|^2 b_{d,k}^2 \beta_k \mathcal{E}_{d,s,k} \varphi_k[n]|^2} \xrightarrow{a.s.} 0, \tag{69}$$

and

$$\sigma_{4,k}^2[n] - \bar{\rho}[n-N]|^2 \frac{\beta_k \varphi_k^2[n] \sum_{m=1}^{K} \mathcal{E}_{d,s,m}}{|1+b_{d,k}^2 |\rho[n-N]|^2 \beta_k \mathcal{E}_{d,s,k} \varphi_k[n]|^2} \xrightarrow{a.s.} 0. \tag{70}$$

Considering equal downlink power being allocated to all the users, the SINR at the kth user for N base station antennas can be expressed as

It can be observed that the SINR is no longer a nondecreasing function of the number of base station antennas, therefore, similar to the MF case, we need to optimize the throughput in terms of the number of antennas and the frame duration. Therefore, the optimization problem in terms of the number of antennas and the frame duration can be written as

$$\max_{M,\alpha_d^{\text{RZF}}} R_{d,\min}^{\text{RZF}} = \frac{M}{\alpha_d^{\text{RZF}}} \sum_{n=N+1}^{\frac{\alpha_d^{\text{RZF}} M}{M}} \log_2 \left(1 + \bar{\eta}_1(M,n)\right), \tag{72}$$

subject to the constraints, $\beta_{min} = \min_k \beta_k$, $0 < \alpha_d^{\text{RZF}} \leq 1$, $M \leq N$, $b_{d,k} = \sqrt{\frac{|\rho[N-1]|^2 \beta_k \mathcal{E}_{d,p}}{\beta_k \mathcal{E}_{d,p} + N_0}}$, and $\bar{\eta}_1(M,n) = \min_k \eta_k^{\text{RZF}}[M,n]$.

The optimal values of M and α_d^{RZF} can be obtained by line search. We next present numerical results to illustrate the system tradeoff revealed by the above expressions.

C. Numerical Results

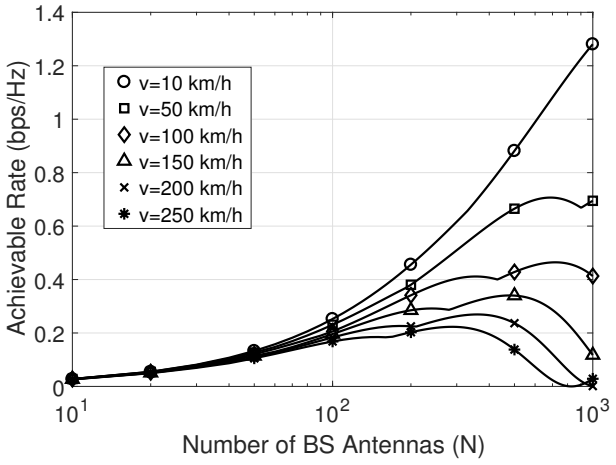
Here, we present numerical results to demonstrate the effects of channel aging on the performance. The simulation setup is same as described in IV-C.

In Figs. 5a, and Fig 5b we plot the per user achievable rates for the MFP and the RZF precoders as a function of the number of base station antennas, optimized over the frame duration for different user velocities. These plot corresponds to data and pilot SNRs of 10 dB. It is observed that at high user velocities, the use of a larger number of base station antennas results in a significant deterioration in the achievable rates. Consequently, the number of base station antennas used for communication should be determined using (58). It is also observed that the RZF precoder offers a significant advantage over the MFP for low user mobility. However, this advantage, arising due to the cancellation of interfering streams also disappears with an increase in user velocities. The non-monotonicity of these plots arises due to the oscillatory nature of the Bessel function of the first order.

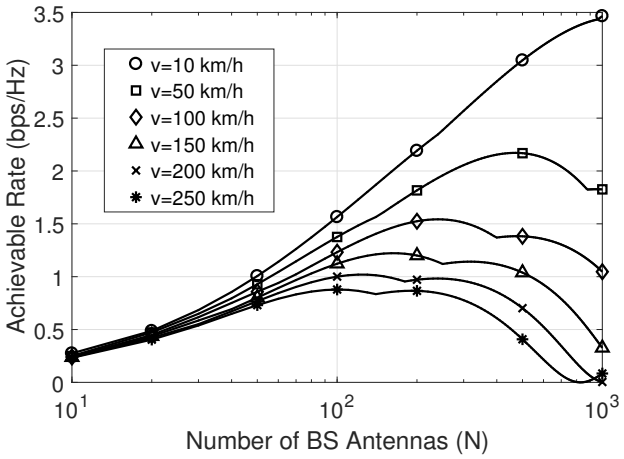
VI. Conclusions

In this paper, we considered the performance of an FDD massive MIMO system under channel aging and derived the dependence of the per user achievable rate on the user mobility and the number of base station antennas. We first derived bounds on the channel estimation error in the presence of channel aging. Following this, we used these bounds along with DE analysis to derive an expression for the per user achievable rate in both uplink and downlink. We considered the MRC and the MMSE receiver in the uplink, and the MFP and the RZF precoders in the downlink. The analysis revealed that under high user mobility, the number of base station antennas maximizing the per user data rate for a given number of users may be less than the number of available base station antennas. We also optimized the frame duration to maximize the per user achivable rate. We showed that optimally choosing the frame duration and the number of base station antennas is an important design issue for massive MIMO systems. Interesting directions for future work include

$$\eta_k^{\text{RZF}}[N,n] - \frac{b_{d,k}^2 \beta_k \mathcal{E}_{d,s} N}{\beta_k \mathcal{E}_{d,s} \sum_{\substack{m=1 \\ m\neq k}}^{K} b_{d,m}^2 \frac{\mu_{k,m}[n]}{\varphi_k^2[n]} + \beta_k \mathcal{E}_{d,s} \sum_{m=1}^{K} \bar{b}_m^2 + (K-1)\frac{|\bar{\rho}[n-N]|^2}{|\rho[n-N]|^2}\beta_k \mathcal{E}_{d,s} + \frac{N_0}{|\rho[n-N]|^2}} \xrightarrow{a.s.} 0. \quad (71)$$



(a)



(b)

Fig. 5: Achievable downlink rate for (a) MFP and (b) RZF precoders based system as a function of the number of base station antennas for different user velocities for a 100 user system optimized over the frame duration.

the optimal power allocation for uplink and downlink channels and the study of wideband massive MIMO systems under channel aging.

APPENDIX A
DERIVATION OF $\sigma_{1,k}^2[n]$

From [32], we know that

$$\mathbf{x}^H(\mathbf{A}+\tau\mathbf{x}\mathbf{x}^H)^{-1} = \frac{\mathbf{x}^H\mathbf{A}^{-1}}{1+\tau\mathbf{x}^H\mathbf{A}^{-1}\mathbf{x}}. \quad (73)$$

Defining

$$\bar{\mathbf{R}}_{n,k} \triangleq |\rho[n-K]|^2 \sum_{\substack{m=1 \\ m\neq k}}^{K} b_m^2 \beta_m \mathcal{E}_{u,s,m} \hat{\mathbf{h}}_m \hat{\mathbf{h}}_m^H + \epsilon[n]\mathbf{I}_k, \quad (74)$$

such that $\mathbf{R}_{yy|\hat{\mathbf{H}}}[n] = \bar{\mathbf{R}}_{n,k} + |\rho[n-K]|^2 b_k^2 \beta_k \mathcal{E}_{u,s,k} \hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H$. Therefore, we can write

$$|\hat{\mathbf{h}}_k^H \mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}[n]\hat{\mathbf{h}}_k|^2 = \frac{|\hat{\mathbf{h}}_k^H \bar{\mathbf{R}}_{n,k}^{-1}\hat{\mathbf{h}}_k|^2}{|1+|\rho[n-K]|^2 b_k^2 \beta_k \mathcal{E}_{u,s,k} \hat{\mathbf{h}}_k^H \bar{\mathbf{R}}_{n,k}^{-1}\hat{\mathbf{h}}_k|^2}, \quad (75)$$

Also, since $\hat{\mathbf{h}}_k$ is independent of $\bar{\mathbf{R}}_{n,k}$, using [12, Lemma 1] and use (73), we get

$$\lim_{N\to\infty} \frac{1}{N^2}|\hat{\mathbf{h}}_k^H \bar{\mathbf{R}}_{n,k}^{-1}[n]\hat{\mathbf{h}}_k|^2 - \frac{1}{N^2}|\text{Tr}(\bar{\mathbf{R}}_{n,k}^{-1})|^2 \xrightarrow{a.s.} 0. \quad (76)$$

Now, the following result holds for a random matrix $\mathbf{X} \in \mathbb{C}^{N\times K}$ with columns $\mathbf{x}_k \sim \mathcal{CN}(0, \frac{1}{N}\mathbf{R}_k)$ such that $\mathbf{R}_k$ has a uniformly bounded spectral norm with respect to M for any $\rho > 0$ [8],

$$\lim_{N\to\infty} \frac{1}{N}\text{Tr}(\mathbf{X}\mathbf{X}^H + \rho\mathbf{I}_N)^{-1} - \frac{1}{N}\text{Tr}(\mathbf{T}(\rho)) \xrightarrow{a.s.} 0, \quad (77)$$

where

$$\mathbf{T}(\rho) = \left(\sum_{k=1}^{K} \frac{\frac{1}{N}\mathbf{R}_k}{1+e_k(\rho)} + \rho\mathbf{I}_N\right)^{-1}, \quad (78)$$

and $e_k(\rho)$ is iteratively computed as

$$e_k^{(t)}(\rho) = \frac{1}{N}\text{Tr}\left(\mathbf{R}_k\left(\sum_{j=1}^{K}\frac{\frac{1}{N}\mathbf{R}_j}{1+e_k^{(t-1)}(\rho)} + \rho\mathbf{I}_N\right)^{-1}\right), \quad (79)$$

with the initialization $e_k^{(t)}(\rho) = \frac{1}{\rho}$. Substituting, $\mathbf{R}_k = N b_k^2 \beta_k \mathcal{E}_{u,s,k}|\rho[n-K]|^2 \mathbf{I}_N \ \forall k$, $\rho = \epsilon[n]$, and defining $e_k[n] \triangleq e_k(\epsilon[n])$ for simplicity we get

$$\mathbf{T}_k(\rho) = \left(|\rho[n-K]|^2 \sum_{\substack{m=1 \\ m\neq k}}^{K} \frac{\beta_m \mathcal{E}_{u,s,k} b_m^2}{1+e_m[n]} + \epsilon[n]\right)^{-1} \mathbf{I}_N. \quad (80)$$

Therefore,

$$|\mathbf{h}_k^H \bar{\mathbf{R}}_{n,k}^{-1}[n]\hat{\mathbf{h}}_k|^2 - N\left(|\rho[n-K]|^2 \sum_{\substack{m=1 \\ m\neq k}}^{K} \frac{\beta_m \mathcal{E}_{u,s,m} b_m^2}{1+e_m[n]} + \epsilon[n]\right)^{-1} \xrightarrow{a.s.} 0. \quad (81)$$

Defining $\varphi_k[n]$ as in (38), it can be shown that

$$|\hat{\mathbf{h}}_k^H \mathbf{R}_{yy|\hat{\mathbf{H}}}^{-1}[n]\hat{\mathbf{h}}_k|^2 - \frac{N^2\varphi_k^2[n]}{|1+\beta_k\mathcal{E}_{u,s,k}b_k^2|\rho[n-K]|N\varphi_k[n]|^2} \xrightarrow{a.s.} 0 \quad (82)$$

and consequently, $\sigma_{1,k}^2[n]$ can be expressed as (37).

APPENDIX B
DERIVATION OF $\sigma_{2,k}^2[n]$

Defining

$$\dot{\mathbf{R}}_{n,k,m} \triangleq \bar{\mathbf{R}}_{n,k} - b_m^2\beta_m\mathcal{E}_{u,s,m}|\rho[n-K]|^2\hat{\mathbf{h}}_m\hat{\mathbf{h}}_m^H, \quad (83)$$

and using the matrix inversion lemma [8]

$$(\mathbf{A}+\tau\mathbf{x}\mathbf{x}^H)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\tau\mathbf{x}\mathbf{x}^H\mathbf{A}^{-1}}{1+\tau\mathbf{x}^H\mathbf{A}^{-1}\mathbf{x}}, \quad (84)$$

we can write

$$\hat{\mathbf{h}}_k^H\bar{\mathbf{R}}_{n,k}^{-1}[n]\hat{\mathbf{h}}_m = \hat{\mathbf{h}}_k^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m - b_m^2\beta_m\mathcal{E}_{u,s,m}|\rho[n-K]|^2 \times \frac{\hat{\mathbf{h}}_k^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m\hat{\mathbf{h}}_m^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m}{1+\beta_m\mathcal{E}_{u,s,m}b_m^2|\rho[n-K]|^2\hat{\mathbf{h}}_m^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m}. \quad (85)$$

Consequently,

$$|\hat{\mathbf{h}}_k^H\bar{\mathbf{R}}_{n,k}^{-1}[n]\hat{\mathbf{h}}_m|^2 = |\hat{\mathbf{h}}_k^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2 + \frac{|b_m^2\beta_m\mathcal{E}_{u,s,m}|^2|\rho[n-K]|^4|\hat{\mathbf{h}}_k^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2|\hat{\mathbf{h}}_m^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2}{|1+b_m^2\beta_m\mathcal{E}_{u,s,m}|\rho[n-K]|^2\hat{\mathbf{h}}_m^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2} - 2\Re\left\{\frac{|b_m^2\beta_m\mathcal{E}_{u,s,m}||\rho[n-K]|^2|\hat{\mathbf{h}}_k^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2\hat{\mathbf{h}}_m^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m}{1+\beta_m\mathcal{E}_{u,s,m}b_m^2|\rho[n-K]|^2\hat{\mathbf{h}}_m^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m}\right\}.$$

Since the columns of $\dot{\mathbf{R}}_{n,k,m}^{-1}$ are independent w.r.t both $\hat{\mathbf{h}_k}$ and $\hat{\mathbf{h}}_\mathbf{m}$, we can use [12, Lemma 1] to show that

$$\frac{1}{N}|\hat{\mathbf{h}}_k^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2 - \frac{1}{N}\mathrm{Tr}(\dot{\mathbf{R}}_{n,k,m}^{-2}) \xrightarrow{a.s.} 0. \quad (86)$$

Letting $\dot{\varphi}_{k,m}[n]$ be defined as in (42), we can write

$$|\hat{\mathbf{h}}_k^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2 - N\dot{\varphi}_{k,m}^2[n] \xrightarrow{a.s.} 0. \quad (87)$$

Also from [12, Lemma 1],

$$|\hat{\mathbf{h}}_m^H\dot{\mathbf{R}}_{n,k,m}^{-1}\hat{\mathbf{h}}_m|^2 \xrightarrow{a.s.} \mathrm{Tr}^2(\dot{\mathbf{R}}_{n,k,m}^{-1}) \xrightarrow{a.s.} N^2\dot{\varphi}_{k,m}^2[n]. \quad (88)$$

Using these, $\sigma_{2,k}^2$ can be expressed as (40).

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, "Survey of large-scale MIMO systems," *IEEE Commn. Surv. Tut.*, vol. 17, no. 3, pp. 1738–1760, Third quarter 2015.

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

[4] T. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 847–860, Oct. 2014.

[5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[6] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[7] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, Cambridge, UK, 2016.

[8] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[9] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2911–2934, May 2012.

[10] 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.

[11] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[12] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338–351, Aug. 2013.

[13] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1428–1444, Feb. 2016.

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

[15] G. Amarasuriya and H. V. Poor, "Impact of channel aging in multi-way relay networks with massive MIMO," in *Proc. IEEE Intl. Conf. Commun. (ICC 2015)*, London, UK, Jun. 2015, pp. 1951–1957.

[16] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, and Z. Zhang, "Sum-rate and power scaling of massive MIMO systems with channel aging," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4879–4893, Dec. 2015.

[17] Q. Bao, H. Wang, Y. Chen, and C. Liu, "Downlink sum-rate and energy efficiency of massive mimo systems with channel aging," in *Proc. 8th Int. Conf. on Wireless Commun. and Signal Process. (WCSP)*, Yangzhou, China, Oct. 2016, pp. 1–5.

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

[19] H. Rydén, "Massive MIMO in LTE with MRT precoder: Channel aging and throughput analysis in a single-cell deployment," Master's thesis, Linköping University, Sweden, 2014.

[20] A. Papazafeiropoulos, H. Ngo, and T. Ratnarajah, "Performance of massive MIMO uplink with zero-forcing receivers under delayed channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3158–3169, Apr. 2017.

[21] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.

[22] A. Decurninge, M. Guillaud, and D. T. M. Slock, "Channel covariance estimation in massive MIMO frequency division duplex systems," in *Proc. IEEE Globecom 2015 Workshops*, San Diego, CA, Dec. 2015, pp. 1–6.

[23] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2868–2882, May 2015.

[24] R. Chopra, C. R. Murthy, and H. A. Suraweera, "On the throughput of large MIMO beamforming systems with channel aging," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1523–1527, Nov. 2016.

[25] S. Kashyap, C. Mollén, E. Björnson, and E. G. Larsson, "Performance analysis of TDD massive MIMO with Kalman channel predication," in *Proc. Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2017)*, New Orleans, LA, March 2017, pp. 3554–3558.

[26] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[27] W. C. Jakes and D. C. Cox, Eds., *Microwave Mobile Communications*, 2nd ed. IEEE Press, New York: IEEE Press, 1994.

[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. New York: Dover, 1964.

[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, Apr. 2003.

[30] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*, 1st ed. Cambridge University Press, Cambridge, UK, 2011.
[31] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal mrc in massive MIMO systems subject to phase noise," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 711–723, Feb. 2015.
[32] J. Silverstein and Z. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175 – 192, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0047259X85710512

Ribhu Chopra (S'11–M'17) received the B.E. degree in Electronics and Communication Engineering from Panjab University, Chandigarh, India in 2009, and the M. Tech. and Ph. D. Degrees in Electronics and Communication Engineering from the Indian Institute of Technology Roorkee, India in 2011 and 2016 respectively. He worked as a project associate at Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore from Aug. 2015, till May 2016. From May 2016 to March 2017 he worked as an institute research associate at the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India. In April 2017, he joined the department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati, Assam, India. His research interests include statistical and adaptive signal processing, massive MIMO communications, and cognitive communications.

Chandra R. Murthy (S'03–M'06–SM'11) received the B. Tech. degree in Electrical Engineering from the Indian Institute of Technology Madras, India, in 1998, and the M. S. and Ph. D. degrees in Electrical and Computer Engineering from Purdue University and the University of California, San Diego, USA, in 2000 and 2006, respectively. From 2000 to 2002, he worked as an engineer for Qualcomm Inc., where he worked on WCDMA baseband transceiver design and 802.11b baseband receivers. From Aug. 2006 to Aug. 2007, he worked as a staff engineer at Beceem Communications Inc. on advanced receiver architectures for the 802.16e Mobile WiMAX standard. In Sept. 2007, he joined the Department of Electrical Communication Engineering at the Indian Institute of Science, Bangalore, India, where he is currently working as an Associate Professor.

His research interests are in the areas of energy harvesting communications, multiuser MIMO systems, and sparse signal recovery techniques applied to wireless communications. He has coauthored 45+ journal papers and 80+ conference papers. His paper won the best paper award in the Communications Track in the National Conference on Communications 2014. He was an associate editor for the IEEE Signal Processing Letters during 2012-16. He is an elected member of the IEEE SPCOM Technical Committee for the years 2014-16. He is a past Chair of the IEEE Signal Processing Society, Bangalore Chapter, and is currently serving as an associate editor for the IEEE Transactions on Signal Processing, IEEE Transactions on Communications, and Sadhana Journal.

Himal A. Suraweera (S'04–M'07–SM'15) received the B.Sc.Eng. degree (Hons.) from the University of Peradeniya, Sri Lanka, in 2001, and the Ph.D. degree from Monash University, Australia, in 2007. He is currently a Senior Lecturer with the Department of Electrical and Electronic Engineering, University of Peradeniya. His research interests include relay networks, energy harvesting communications, full-duplex, physical layer security, and multiple input-multiple-output systems.

Dr. Suraweera was a recipient of the 2017 IEEE ComSoc Leonard G. Abraham Prize, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2011, the WCSP Best Paper Award in 2013, and the SigTelCom Best Paper Award in 2017. He is serving as a Co-Chair of the Emerging Technologies, Architectures and Services Track of the IEEE WCNC 2018. He was an Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Series on Green Communications and Networking) from 2015 to 2016 and the IEEE COMMUNICATIONS LETTERS from 2010 to 2015. He is currently serving as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING.

Erik G. Larsson (S'99-M'03–SM'10–F'16) received the Ph.D. degree from Uppsala University, Uppsala, Sweden, in 2002.

He is currently Professor of Communication Systems at Linköping University (LiU) in Linköping, Sweden. He was with the Royal Institute of Technology (KTH) in Stockholm, Sweden, the University of Florida, USA, the George Washington University, USA, and Ericsson Research, Sweden. In 2015 he was a Visiting Fellow at Princeton University, USA, for four months. His main professional interests are within the areas of wireless communications and signal processing. He has co-authored some 130 journal papers on these topics, he is co-author of the two Cambridge University Press textbooks *Space-Time Block Coding for Wireless Communications* (2003) and *Fundamentals of Massive MIMO* (2016). He is co-inventor on 16 issued and many pending patents on wireless technology.

He was Associate Editor for, among others, the *IEEE Transactions on Communications* (2010-2014) and the *IEEE Transactions on Signal Processing* (2006-2010). From 2015 to 2016 he served as chair of the IEEE Signal Processing Society SPCOM technical committee, and in 2017 he is the past chair of this committee. From 2014 to 2015 he served as chair of the steering committee for the *IEEE Wireless Communications Letters*. He was the General Chair of the Asilomar Conference on Signals, Systems and Computers in 2015, and its Technical Chair in 2012. He is a member of the IEEE Signal Processing Society Awards Board during 2017–2019.

He received the IEEE Signal Processing Magazine Best Column Award twice, in 2012 and 2014, the IEEE ComSoc Stephen O. Rice Prize in Communications Theory in 2015, and the IEEE ComSoc Leonard G. Abraham Prize in 2017. He is a Fellow of the IEEE.