Training for Massive MIMO Systems with Non-Identically Aging User Channels

Soumendu Ghosh, Ribhu Chopra

Abstract

In this paper, we consider the problem of channel state information (CSI) acquisition in massive MIMO systems with the users exhibiting non-identical channel aging profiles. For this purpose, we derive the per user achievable rates, and the corresponding CSI outage times for different users. We then use these derived results to argue that scheduling the training of all the users, with respect to the outage time of the fastest moving user is sub optimal, and results in unnecessary training overhead. Therefore, we present and evaluate the performance of two simple schemes for scheduling the training of users with different mobilities. Via numerical results, we show that training a subset of users based on the ages of their channel estimates can significantly increase the spectral efficiency of a massive MIMO system.

Keywords: Massive MIMO, channel aging, channel estimation, performance analysis, achievable rate.

1. Introduction

Multiuser multiple input multiple output (MIMO) communication systems serving a large number of users, and having a larger number of base station (BS) antennas, have been recognized as a key enabler for next generation wireless communication systems [1]. This is due to the fact that the large system dimensions lead to quasi orthogonality among the channels to the different users [2], facilitating the use of simple linear precoders and detectors, and reducing the system complexity. Also, the large system dimensions produce an increased array gain, further resulting in high spectral and energy efficiencies. However, these advantages are contingent on the availability of accurate channel state information (CSI) at the BS, and have been shown to deteri-
orate significantly in the presence of CSI inaccuracies. The major causes for CSI imperfections in massive MIMO systems are additive noise, pilot contamination [3], calibration imperfections [4], and channel aging [5].

Channel aging is a consequence of the changing propagation environment of fast moving users, and manifests as a mismatch between the channel state estimate available at the BS and the actual channel state at the time of transmission [5–10]. Recent studies in this direction have revealed that the correlation between the true channel and the channel estimate at the BS worsens as the user mobility, and the delay between CSI acquisition and data transmission increase [5]. Some inherent advantages of massive MIMO, such as power scaling, have been shown to remain unaffected by channel aging [11]. However, channel aging and increased user mobilities place a limit on the number of users being served in a massive MIMO setting [9], as well as the system dimensions [12]. Further, increased user mobilities result in quicker aging of channel estimates, thus leading to the requirement of frequent retraining, and shorter data transmission durations. Hence, more recent studies argue that the frame structures should be designed to compensate for the effects of channel aging [12, 13]. Some recent works have also explored the possibility of Kalman filter based channel estimation for tracking an aging channel [13–15].

However, the works listed above, except [15], assume that the channels to all the users in the system age identically. Consequently, it is common to model the process of channel aging with respect to the fastest moving user [5–10, 13]. While this assumption is acceptable for analyzing the worst case performance of massive MIMO systems under aging, it is unduly pessimistic, and can lead to suboptimal design of the system dimension and frame structure. Therefore, we address the problem of training in a massive MIMO uplink system with the channels to different users aging at non-identical rates.

In this paper, we consider a single cell uplink massive MIMO system with a frame structure designed to guarantee a specific quality of service to the fastest moving user in the cell, but with the users having non-identical velocities. This results in different users aging at different rates, leading to their respective CSIs being outdated at different points of time. To capture this effect, we define the outage time for a user as the time after which the channel to that user becomes unusable, and needs to be retrained. We then show that the outage times for users moving at different velocities
are unequal and retraining the channels to all the users with respect to the shortest outage time is wasteful. We then present two greedy algorithms to schedule the training of the users and discuss their performance. Our main contributions are:

1. We analyze the achievable rates and outage times for a massive MIMO system with independently aging user channels. (See Section 3).
2. We present two algorithms, based on the received signal strength at the BS, to schedule the training of different users. (See Section 4).
3. Via extensive numerical simulations, we analyze and compare the behavior of the two scheduling algorithms for different user velocity profiles. (See Section 5).

The key finding of this work is that for a frame structure designed to ensure a pre-determined quality of service to the user worst affected by aging, the frequency of training for users with different mobilities is different. Also, that the optimal training schedule needs to be chosen based on the overall user mobility profile for the system. We note that selective training of users based on their individual rate of channel variation has been considered for inclusion in the 5G standard [16]. In the next section, we discuss the system model for this work.

2. System Model

We consider the uplink of a single cell TDD massive MIMO system with an $N$ antenna (indexed as $i \in \{1, \ldots, N\}$) BS serving $K$ users (indexed as $k \in \{1, \ldots, K\}$). The $k$th user is assumed to move at a velocity $v_k$. The channel coefficient between the $k$th user and the $i$th BS antenna at the $n$th instant is given as $\sqrt{\beta_k} h_{ki}[n]$, with $\beta_k$ being the macroscopic fading coefficient between the BS and the $k$th user, assumed to be constant across all the BS antennas due to the small inter-antenna spacing. Also, the user velocities considered by us are limited to tens of meters per second, and the time between two consecutive CSI acquisitions is of the order of milliseconds, and the displacement is small enough to be ignored for large scale fading coefficients, and therefore $\beta_k$ is assumed to be constant in time. Also, $h_{ki}[n]$ is the fast fading coefficient between the $k$th user and the $i$th antenna at the $n$th instant, distributed as a zero mean circularly symmetric complex Gaussian random variable with unit variance (denoted as $h_{ki}[n] \sim CN(0, 1)$). We also assume that
$h_{ki}[n]$s are independent across $k$ and $i$ but not across $n$, and the evolution of the fast fading channel coefficient vector from the $k$th user to the BS $\mathbf{h}_k[n] \triangleq [h_{k1}[n], \ldots, h_{kN}[n]]^T$, follows [9]

$$\mathbf{h}_k[n + \tau] = \rho_k[\tau] \mathbf{h}_k[n] + \tilde{\rho}_k[\tau] \mathbf{z}_k[\tau],$$

where $\rho_k[\tau] \triangleq \mathbb{E}[h_{ki}[n]h_{ki}^*[n-\tau]]$ denotes the channel correlation coefficient assumed to be the same for all the channels to the $k$th user, $\mathbb{E}[\cdot]$ denotes the expectation operator, and $\mathbf{z}_k[\tau] \sim \mathcal{CN}(\mathbf{0}_N, \mathbf{I}_N)$ is the innovation component for the $k$th user’s channel for the lag $\tau$, such that $\mathbb{E}[\mathbf{h}_k[n]\mathbf{z}_k^H[\tau]]$ equals the $N \times N$ all zero matrix. The innovation process $\mathbf{z}_k[n]$ is assumed to be stationary and ergodic, but not white. Here, we note that this aging model is slightly different from the one discussed in [9], wherein $\rho_k[n]$ is assumed to be independent of $k$. The channel is assumed to evolve according to the Jakes’ model [5, 17], such that $\rho_k[n] = J_0(2\pi f_{dk} T_s n)$, where $f_{dk}$ is the Doppler frequency for the $k$th user, $T_s$ is the sampling period at the BS, and $J_0(\cdot)$ is the Bessel function of the first kind and zeroth order [18, Eq. (9.1.18)]. The Doppler frequency for the $k$th user, $f_{dk}$, is further defined in terms of $v_k$ as

$$f_{dk} \triangleq \frac{f_c v_k}{c},$$

with $f_c$ being the carrier frequency, and $c$ the velocity of light. Also, for any variable, $x$, $\tilde{x} \triangleq \sqrt{1 - |x|^2}$.

We assume the uplink data frame to consist of a total of $T(> K)$ channel uses, divided into two slots, training and transmission, consisting of $L(1 \leq L \leq K)$ and $T - L$ channel uses, respectively. During the training slot, $L$ scheduled users transmit orthogonal pilots to the BS during the $L$ channel uses. It is important to note that $L$ can be constant or, vary across frames, depending on the choice of the user scheduling algorithm. This is discussed in greater detail in Section 4. The BS then uses the received training symbols to update the available channel estimates for the $L$ scheduled users. Following this, during the next $T - L$ channel uses, all the $K$ users transmit their uplink data to the BS, decoded by the BS using the acquired CSI.

2.1. CSI acquisition

We assume that a Kalman filtering based technique, as discussed in [13–15], is used to update the CSI at the BS during each training period. Since the channel is assumed to vary according
to a linear regression model, it is easy to show that the Kalman filter minimizes the mean squared error at each stage [19, Chapter 10]. Letting $\tau_{m,k}$ denote the training instant for the $k$th user for the $m$th frame, we can write the signal received at the BS antennas at the $\tau_{m,k}$th instant as,

$$y[\tau_{m,k}] = h_k[\tau_{m,k}] \sqrt{\beta_k E_{p,k}} + \sqrt{N_0}w_k[\tau_{m,k}]$$

(3)

where $E_{p,k}$ is the pilot power, $N_0$ is the noise variance, and $w_k[I_{m,k}]$ is the additive ZMCSCG noise with unit variance, denoted as $w[\tau_{m,k}] \sim CN(0, I_K)$. It is important to note that since we consider a single cell massive MIMO system, the effects of pilot contamination are not accounted for. However, the extension of this model to a multi cell system, and the incorporation of the effects of pilot contamination is trivial.

Since we assume that the channel evolves according to (1), the BS can update the channel estimate acquired at the previous training instant, denoted as $\hat{h}_k[\tau_{l,k}]$ ($l < m$), using the training symbol received at $\tau_{k,n}$. Note that if all users are not scheduled for training during all the frames, $l$ is not necessarily equal to $m - 1$. The Kalman filter update equation to obtain $\hat{h}_k[\tau_{m,k}]$ from $\hat{h}_k[\tau_{l,k}]$ and $\hat{y}[\tau_{m,k}]$ is given as [13],

$$\hat{h}_k[\tau_{m,k}] = \rho_k[\tau_{m,k} - \tau_{l,k}]\hat{h}_k[\tau_{l,k}] + g_k[n]\alpha_k[\tau_{m,k}],$$

(4)

where,

$$\alpha_k[\tau_{m,k}] = y[\tau_{m,k}] - \rho_k[\tau_{m,k} - \tau_{l,k}] \sqrt{\beta_k E_{p,k}\hat{h}_k[\tau_{l,k}]},$$

(5)

is the innovation component,

$$g_k[\tau_{m,k}] = \frac{\psi_k[\tau_{m,k}, \tau_{l,k}] \sqrt{\beta_k E_{p,k}}}{\beta E_{p,k}\psi_k[\tau_{m,k}, \tau_{l,k}] + N_0},$$

(6)

is the Kalman gain,

$$\psi_k[\tau_{m,k}, \tau_{m,k} + \tau_0] = |\rho_k[\tau_0]|^2\psi_k[\tau_{m,k}] + |\tilde{\rho}_k[\tau_0]|^2,$$

(7)

is the mean squared estimation error for the channel at the $(\tau_{m,k} + \tau_0)$th instant, obtained using the information available till the $\tau_{m,k}$th instant, and

$$\psi_k[\tau_{m,k}] = \psi_k[\tau_{m,k}, \tau_{l,k}] - \frac{\sqrt{\beta_k E_{p,k}}}{\rho_k[(m-l)T]}g_k[\tau_{m,k}]$$

(8)
is the mean squared channel estimation error at the $\tau_{m,k}$th instant. The above recursive equations are initialized as $\hat{h}_k[\tau_{0,k}] = 0_N$, and $\psi_k[\tau_{1,k}, \tau_{0,k}] = 1$. It can be shown that $\psi_k[\tau_{m,k}]$ is upper bounded as [19, Chapter 10]

$$\psi_k[\tau_{m,k}] \leq \psi_k[\tau_{1,k}] = \frac{N_0}{\beta_k E_p,k + N_0},$$

(9)

with the upper bound being achieved for $m = 1$, when, in the absence of a prior estimate at the BS, the Kalman filter based estimator reduces to the one step MMSE estimator.

Now, letting $t_{k,n}$ denote the time of the most recent CSI acquisition for the $k$th user’s channel with respect to time instant $n$, and letting $\hat{h}_k[n] \triangleq \hat{h}_k[t_{k,n}]$, be the most recent estimate for the $k$th user’s channel. Also, letting $\psi_k[t_{k,n}]$, be the mean squared channel estimation error for $\hat{h}_k[n]$ at the time of CSI acquisition, it can be shown that the true channel at the $n$th instant, $h_k[n]$, can be expressed as [13]

$$h_k[n] = (1 - \psi_k[n, t_{k,n}])^{1/2} \hat{h}_k[n] + \psi_k^{(1/2)}[n, t_{k,n}] \tilde{h}_k[n],$$

(10)

with

$$\psi_k[n, t_{k,n}] \triangleq |\rho_k[n - t_{k,n}]| \psi_k[t_{k,n}] + |\bar{\rho}[n - t_{k,n}]|^2,$$

(11)

and $\tilde{h}_k[n]$ being the zero mean channel estimation error uncorrelated with $\hat{h}_k[n]$. We next use these results to obtain the achievable uplink rates for a maximal ratio combining (MRC) receiver at the BS.

### 3. Achievable Rate Analysis

During the data transmission phase, all the $K$ users simultaneously transmit their respective data symbols to the BS, such that the symbol vector received at the BS at the $n$th instant is,

$$y[n] = \sum_{k=1}^{K} \sqrt{\beta_k E_{s,k}} h_k[n] s_k[n] + \sqrt{N_0} w[n] = \sqrt{\beta_k E_{s,k}} h_k[n] s_k[n] + \sum_{l=k}^{K} \sqrt{\beta_l E_{s,k}} h_l[n] s_l[n] + \sqrt{N_0} w[n],$$

(12)

where $E_{s,k}$ is the uplink energy transmitted by the $k$th user, and $s_k[n]$ is the symbol being transmitted by the $k$th user at the $n$th instant. Using maximal ratio combining (MRC) at the BS, the
combining vector for the kth user’s channel is \( \hat{h}_k[n] \), and the combined signal for the kth user is given as

\[
r_k[n] = \hat{h}_k^H[n]y[n] = (1-\psi_k[n, t_{k,n}])^{1/2} \sqrt{\beta_k E_{s,k}} \hat{h}_k^H[n] \hat{h}_k[n] s_k[n] + \sqrt{\psi_k[n, t_{k,n}]} \beta_k E_{s,k} \hat{h}_k^H[n] \hat{h}_k[n] s_k[n] + \sum_{l=1}^{K} \sqrt{\beta_l E_{s,l}} \hat{h}_k^H[n] \hat{h}_l[n] s_l[n] + \sqrt{N_0} \hat{h}_k^H[n] w[n].
\]

(13)

In the above equation, the first term corresponds to the desired signal, the second to the self interference due to CSI impairments, the third to the inter user interference and the last to the additive noise. Treating interference as noise [20], the SINR for the kth user at nth instant for the given channel realization can be expressed as

\[
\gamma_{u,k}[n] = \frac{(1-\psi_k[n, t_{k,n}]) \beta_k E_{s,k} \hat{h}_k^H[n] \hat{h}_k[n]^2}{\psi_k[n, t_{k,n}] \beta_k E_{s,k} \hat{h}_k^H[n] \hat{h}_k[n]^2 + \sum_{l=1}^{K} \beta_l E_{s,l} \hat{h}_k^H[n] \hat{h}_l[n] s_l[n] + N_0}. 
\]

(14)

Since \( \hat{h}_k[n], \hat{h}_k[n], h_l, \) and \( w[n] \) are independent with i.i.d. \( \mathcal{CN}(0, 1) \) entries, it can be argued that each term in (14) converges to its deterministic equivalent (DE) [21]. Also, from the results in [5, Lemma 1], we can write,

\[
\gamma_{u,k}[n] \xrightarrow{a.s.} 0.
\]

(15)

Let us define,

\[
\gamma_{u,k}^{max} = \frac{N(1-\psi_k[t_{k,n}]) \beta_k E_{s,k}}{\psi_k[t_{k,n}] \beta_k E_{s,k} + \sum_{l=1}^{K} \beta_l E_{s,l} + N_0}
\]

(16)

as the maximum achievable SINR for the kth user, corresponding to the achievable SINR at the time of training. Then we consider the channel as being usable only when \( \gamma_{u,k}[n] \) stays above a fraction \( \mu \) (0 ≤ \( \mu \) ≤ 1) of \( \gamma_{u,k}^{max} \). Considering that the channel goes in outage whenever \( \gamma_{u,k}[n] < \mu \gamma_{u,k}^{max} \), we can implicitly define the outage time for the kth channel, \( \Delta_k(\mu) \), by the equation

\[
\gamma_{u,k}[t_{k,n} + \Delta_k(\mu)] = \mu \gamma_{u,k}^{max}.
\]

(17)

Now, since \( (1-\psi_k[n, t_{k,n}]) \) can be lower bounded as \(|\rho_k[n - t_{k,n}]|^2 \beta_k E_{p,k} \beta_k E_{p,k} + N_0|\), we can write,

\[
\gamma_{u,k}[n] \xrightarrow{a.s.} \frac{N \beta_k E_{s,k} |\rho_k[n - t_{k,n}]|^2 \beta_k E_{p,k} + N_0}{\beta_k E_{s,k} (\beta_k E_{p,k} + N_0) + \sum_{l=1}^{K} \beta_l E_{s,l} + N_0} \rightarrow 0.
\]

(18)
Consequently, $\Delta_k(\mu)$ can be obtained by solving the equation

$$
|\rho_k[\Delta_k(\mu)]|^2 \left( \frac{N_0}{\beta_k E_{s,k} \beta_k E_{p,k} + N_0} + \sum_{l \neq k}^K \beta_l E_{s,l} + N_0 \right)
$$

$$
= \mu \left( \frac{\left|\bar{p}_k[\Delta_k(\mu)]\right|^2}{\beta_k E_{p,k} + N_0} + \frac{N_0}{\beta_k E_{p,k} + N_0} \right) \beta_k E_{s,k} + \sum_{l \neq k}^K \beta_l E_{s,l} + N_0 \right),
$$

(19)

that can be further simplified as

$$
\Delta_k(\mu) = \rho_k^{-1} \left( \sqrt{\frac{\sum_{l=1}^K \beta_l E_{s,l} + N_0}{\frac{N_0}{\beta_k E_{p,k} + N_0} + \mu \beta_k E_{p,k} + \sum_{l \neq k}^K \beta_l E_{s,l} + N_0}} \right),
$$

(20)

with $\rho_k^{-1}()$ indicating the inverse function of $\rho_k()$.

Now, assuming the users to follow statistical channel inversion based power control for uplink pilot and data powers [9, 22], we get, $\beta_k E_{s,k} = E_s$, $\beta_k E_{p,k} = E_p$, and consequently,

$$
\gamma_{u,k}[n] - \frac{N E_s |p_k[n - t_{k,n}]|^2 E_p}{E_s \left( |\bar{p}_k[n - t_{k,n}]|^2 E_p + N_0 \right) + (K - 1) E_s + N_0} \xrightarrow{a.s.} 0.
$$

(21)

and therefore,

$$
\Delta_k(\mu) = \rho_k^{-1} \left( \sqrt{\frac{KE_s + N_0}{\frac{N_0}{E_p + N_0} + \mu \frac{E_p}{E_p + N_0} \beta_k E_{p,k} + \sum_{l \neq k}^K \beta_l E_{s,l} + N_0}} \right).
$$

(22)

It can be observed that the outage time is a highly nonlinear function of the pilot and data SNRs, as well as the user velocity, and is different for different users.

Also, in order to avoid the outage of the communication link to the $k$th user, it should be retrained before it goes into outage. Now, considering a standard massive MIMO frame where all the users are trained at the start of each frame, the frame duration $T$ becomes constrained by the outage time of the fastest moving user, such that,

$$
T \leq \min_k \Delta_k(\mu)
$$

(23)

for a given value of $\mu$. However, in this case, the users with a longer outage time will be trained unnecessarily, and since each pilot is transmitted during a channel use, this unwarranted training
results in a reduced throughput of the system. Therefore, it becomes important to consider a frame structure that retrains only a subset of the users in each frame.

If only \( L \) out of the total \( K \) users are trained per frame, then, the achievable spectral efficiency for the \( k \)th user can be expressed as,

\[
R_{u,k} = \frac{1}{T} \sum_{n=(m-1)T+1}^{mT} \log_2 \left( 1 + \frac{N E_s [\rho_k[n-t_{k,n}]]^2 E_s}{E_s (|\bar{\rho}_k[n-t_{k,n}]|^2 E_s + N_0) + (K-1)E_s + N_0} \right),
\]

and sum total uplink spectral efficiency for the \( m \)th frame as

\[
R_u = \sum_{k=1}^{K} R_{u,k}.
\]

From the expression for \( R_{u,k} \), it can be observed that a smaller value of \( L \) leads to a larger number of terms within the summation, potentially increasing the achievable rates. However, a smaller value of \( L \) also implies that only a small fraction of users is trained during the \( m \)th frame. This leads to an enhancement in the gap between the acquisition time of the previously acquired channel estimate, and the time instant of transmission, further increasing the argument of \( \rho_k[n-t_{k,n}] \). A decreased \( \rho_k[n-t_{k,n}] \) reduces the spectral efficiency per channel use for the said user, finally limiting the overall sum rate. On the other hand, while a large \( L \) ensures smaller lags between training and data transmission, it reduces the number of terms contained within the summation over \( n \), again limiting the data rates. The optimal \( L \) for the given frame can therefore be determined by maximizing \( R_u \) in terms of \( t_{k,n} \). However, since \( R_{u,k} \) and therefore, \( R_u \) involve the summation of a large number of logarithms of Bessel functions, a closed form solution is not feasible. However, it is easy to numerically optimize \( R_u \) via a one dimensional search over \( L \). In the next section, we present two algorithms for scheduling the training of various users.

4. User Scheduling Algorithms

4.1. Scheduling a fixed number of users in each frame

We first consider the case where the frame structure is designed to schedule a fixed number of users \( (L) \) in each frame. In this case, since different users age non-identically, it is intuitive to schedule the \( L \) (out of the total \( K \)) weakest users for training. Therefore, at the end of each
Algorithm 1 Scheduling algorithm for a fixed number of users per frame

GIVEN: $\gamma_{u,k}[t,k,mT]$ and $\gamma_{u,k}[mT]$ at the end of $m$th frame for $1 \leq k \leq K$.

CALCULATE: $b_k = \frac{\gamma_{u,k}[mT]}{\gamma_{u,k}[t,k,mT]}$.

SORT: $b_k$ in increasing order.

SCHEDULE: The $L$ users with the $L$ smallest $b_k$s.

Algorithm 2 Scheduling Algorithm with a variable number of users per frame

GIVEN: $\gamma_{u,k}[t,k,mT]$ and $\gamma_{u,k}[mT]$ at the end of $m$th frame for $1 \leq k \leq K$.

CALCULATE: $b_k = \frac{\gamma_{u,k}[mT]}{\gamma_{u,k}[t,k,mT]}$.

for $k = 1 \rightarrow K$ do
  if $b_k < \lambda$ then
    Schedule user $k$ for training in the $(m+1)$th frame.
  end if
end for

frame, the BS sorts the users in the increasing order of their received SINRs, and schedules the $L$ weakest users. This is summarized in Algorithm 1. As discussed in the previous section, the performance of this algorithm is determined by the choice of $L$, which, in turn, depends on the distribution of velocities among the users. We discuss the performance in greater detail in the section on numerical results.

4.2. Scheduling a variable number of users in each frame

If the frame structure provides a flexibility in terms of the number of users that can be scheduled in each frame, then all the users whose relative SINRs fall below a given threshold $\lambda$, can be scheduled for training at the end of each frame. This flexibility, however, comes at the cost of a more complex frame structure. Also, the choice of the threshold $\lambda$ is independent of the user mobility profile for the given cell. In addition to this, a threshold $\lambda$ can be chosen, and a suitable $L$ can be calculated numerically for a given user mobility profile. We discuss this in greater detail in the section on numerical results. A summary of this scheduling technique is given as Algorithm 2.
5. Numerical Results

For the purpose of these experiments, we consider a $N = 500$ antenna BS serving $K = 100$ users, and transmitting at a frequency of 2 GHz. We assume that the BS samples the complex baseband signal of bandwidth of 1 MHz, at the Nyquist rate, 1 MHz. Throughout this section, we assume that the users in the cell can be classified into slow (moving at velocities less than 50 km/h) and fast (moving at velocities between 50 km/h and 200 km/h), we assume that a fraction, $\alpha$, of the total number of users is slow moving. The frame duration is designed so that the SINR of the fastest moving user (at 200 km/h), does not fall below 50% of its maximum achievable value. In other words, the frame length is equivalent to the outage time of the fastest moving user for $\mu = 50\%$, which approximately corresponds to a frame length $T = 500$. We also assume statistical channel inversion based power control for both pilot and data transmission. It has been established in [6, 7, 9, 11], that results based on DE analysis can be used to closely approximate the performance of massive MIMO systems. Therefore, we omit the plots showing the equivalence of simulated results to DE analysis for the sake of brevity.

In Fig. 1, we plot the achievable per user rates for different user mobility profiles against $L$. It can be observed that as the fraction of users with smaller velocities increases, the number of users required to be trained per frame reduces. In addition to this a decrease in the number of fast moving users also results in better rates for the same training duration, which is in accordance with intuition. It is important to note that $L = 100$ corresponds to the case considered in [9].

In Fig. 2 we plot the fraction of users to be trained for maximizing the average per user achievable rate, referred to as the optimal fraction of users to be trained per frame, as a function of the number of users, for different user velocity profiles. It is again observed that the optimal fraction of users to be trained per frame decreases with both an increasing number of users, and a decreasing average user velocity. It can be concluded from the effect of the number of users on the optimal fraction of users to be trained per frame, that reduction in rates due to longer training durations affects the system more than the reduction in SINR due to aging. That is, the effect of reducing the number of time slots available for transmission outweighs the effects of a reduction in the individual values of the terms being summed.
In Fig. 3 we plot the achievable per user uplink rates for different user mobilities against the threshold $\lambda$ as discussed in Algorithm 2. The achievable rate is again observed to be a non monotonic function of $\lambda$, showing similar trends for all user velocity profiles, and peaking approximately at $\lambda = 0.9$. An appropriate training interval can therefore be obtained by substituting $\mu = 0.95$ in (22).

In Fig. 4 we plot the average number of training symbols required per frame, as a function of the threshold $\lambda$ for different user mobility profiles. It is observed that an increased user mobility, and hence smaller value of $\alpha$, leads to the requirement of training a larger number of users per frame to maintain the user SINRs above a given threshold. This plot can also be used to choose an optimal $L$ for Algorithm 1 for different user mobility profiles.

6. Conclusions

In this paper, we considered a generalized aging massive MIMO system where the channels to different users age non identically. We derived the expressions for the achievable per user rates and the sum rate, and argued that, it is suboptimal to train all the users in each frame. Then using
numerical techniques, we obtained the achievable sum rates for a 100 user system under different user mobility profiles. We showed that the number of users to be trained per frame should be chosen carefully based on the total number of users being served and the user mobility profile. In this paper, we limited the discussion to an uplink massive MIMO system, however, it is easy to extend these results to a downlink system as well.

Acknowledgment

We thank Prof. Chandra R. Murthy, Dept. of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India for his generous feedback on the language of this manuscript. Part of this work was funded by an IIT Guwahati Start-up grant, under grant number xEEESUGI-ITG01212RBHU001.


Figure 3: Achievable per user rates for different user mobility profiles as functions of the threshold $\lambda$.


Figure 4: Achievable uplink rates for different numbers of users under different configurations for a user velocity of 150 km/h.

Feb 2019.


