

Analysis of Non-Orthogonal Training in Massive MIMO under Channel Aging with SIC Receivers

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

Abstract—We analyze the effect of channel aging on the achievable rate of time division duplexed (TDD) massive multiple input multiple output (MIMO) systems serving a number of users under aging channels, using non orthogonal multiple access (NOMA) and orthogonal multiple access (OMA). Using the recently proposed shared uplink pilot based channel estimation for NOMA, we derive bounds on the channel estimation error variance for the two schemes. We then derive the achievable spectral efficiencies of the two schemes. Using numerical results, we show that, in slowly varying channels, using NOMA with shared pilots is preferable over OMA, while the reverse is true under fast varying channels.

I. INTRODUCTION

Multiuser MIMO systems with a large number of base station (BS) antennas serving a smaller number of users, is a promising technology for next generation wireless communications [1]–[3]. It has recently been shown that in addition to conventional impairments such as estimation errors and pilot contamination [4], channel aging [5]–[8], caused due to the dynamic nature of the wireless channel, is a major source of inaccuracy in the Channel State Information (CSI) available in massive MIMO systems. Channel aging manifests as a mismatch between the acquired CSI and the channel state at the time of data transmission [9], [10], and a consequent reduction in the achievable data rates [8]. Past work has not considered the effect of aging on different Multiple Access (MA) techniques such as Non-Orthogonal MA (NOMA), which is the focus of this paper.

NOMA is a technique for boosting the capacity of a wireless system when serving a combination of near and far users. It uses superposition coding and successive interference cancellation (SIC) techniques to allow users to access the same time-frequency slot [11], [12]. It has been shown in [13] that the idea of NOMA can be extended to a MIMO system with the design of suitable precoding and combining matrices. A scheme for using shared pilot based training for time division duplexed (TDD) massive MIMO systems was proposed in [14]. It was shown that the use of NOMA can increase the throughput of a massive MIMO system in some cases. In their scheme, pilot sequences are reused by pairs of users, allowing one to halve the training overhead, leading to a longer effective data duration. However, the effect of aging was not considered in [14], and the gain in sum rate due to

NOMA may quickly disappear in high mobility scenarios due to aging. Also, in NOMA, users need to decode signals in the presence of inter-user interference, which gets exacerbated by channel mismatches due to aging. Analyzing effect of channel aging on the sum rate of NOMA is the goal of this paper.

In this paper, we derive and compare bounds on the channel estimation errors and achievable downlink rates using the shared pilot based scheme introduced in [14] against the conventional orthogonal multiple access (OMA) scheme in a TDD massive MIMO system under channel aging. Our contributions are as follows:

- 1) We reformulate the effective estimate of the channel to a user pair as a sum of two independent channel vectors, and use it to derive upper bounds on the mean squared channel estimation error accounting for the effect of channel aging during training. (See Section III.)
- 2) We derive lower bounds on the per user achievable rates in the downlink with matched filter (MF) precoding. (See Section IV.)
- 3) We empirically compare the performance of the two multiple access schemes under different channel aging conditions. (See Section V.)

Our results reveal that the impact of channel aging depends heavily on the underlying protocol, and is often difficult to predict using intuition. It can be determined only by careful analysis and interpreting the results. We next discuss the channel and system model considered in this paper.

II. SYSTEM MODEL

Under NOMA transmission, we consider a single cell massive MIMO system with an N antenna BS and K user pairs, ($2K < N$), with each user pair consisting of a strong user and a weak user. Letting $\beta_{k,h}$ and $\beta_{k,g}$ denote the macroscopic fading coefficients of the k th strong and weak user, respectively. We assume that the large scale fading coefficients are the same for all BS antennas, and remain unaffected by aging [7], [8]. Each user pair satisfies $\beta_{k,h} > \beta_{k,g}$, and typically we have, $\beta_{k,h} \gg \beta_{k,g}$.

We consider a downlink TDD system with reverse link training, and assume the channels between the BS and the users to be reciprocal. In the NOMA scheme, each frame consisting of T channel uses is divided into a training phase of duration K and a transmission phase of duration $T - K$. During the training phase, all users simultaneously transmit pilot signals to the BS. The pilot sequences are orthogonal across user pairs, but both users in each pair share the same pilot sequence. The BS uses these contaminated pilots to obtain an estimate of a weighted linear combination of the

R. Chopra is with the Dept. of EEE, IIT Guwahati, India. (ribhufec@iitg.ac.in). C. R. Murthy is with the Dept. of ECE, IISc Bangalore, India. (cmurthy@iisc.ac.in). H. A. Suraweera is with the Dept. of EEE, University of Peradeniya, Sri Lanka (himal@ee.pdn.ac.lk). E. G. Larsson is with the Dept. of EE (ISY), Linköping University, Sweden (erik.g.larsson@liu.se). The work of E.G. Larsson was partially supported by ELLIIT and the Swedish Research Council (VR).

channels corresponding to the two users in a group [14]. Since the channel estimate contains non-zero components in the direction of the channels to both the users in a pair, a beam can be formed by the BS in the general direction of the user pair. The BS then transmits a combination of the symbols intended for the two users over the beam formed for the given user pair. In each user pair, the strong user first decodes the symbols intended for the weak user, removes interference from the weak user using SIC, and then decodes its own symbol. The weak user treats the interference due to the symbol intended for the strong user as noise.

The channel coefficient between the i th BS antenna and the k th strong and weak users at the n th instant are $\sqrt{\beta_{k,h}}h_{ki}[n]$, and $\sqrt{\beta_{k,g}}g_{ki}[n]$, respectively, such that $h_{ki}[n], g_{ki}[n] \sim \mathcal{CN}(0, 1)$ are the independent (across k and i , but not across n) fading coefficients for the two users. We define $\mathbf{q}_k^T \triangleq [q_{k1}, \dots, q_{kN}]$; $q \in \{h, g\}$ as the channel coefficient vectors from the BS to the users constituting the k th pair.

The wireless channels evolve in time as [8]

$$\mathbf{q}_k[n + \tau] = \rho[\tau]\mathbf{q}_k[n] + \bar{\rho}[\tau]\mathbf{z}_{q,k}[\tau], \quad q \in \{h, g\}, \quad (1)$$

where $\rho[\tau] \triangleq E[h_{ik}[n]h_{ik}^*[n - \tau]] = E[g_{ik}[n]g_{ik}^*[n - \tau]]$ denotes the channel correlation coefficient assumed to be the same for the channels to all users, $E[\cdot]$ denotes the expectation operator, and $\mathbf{z}_{h,k}[\tau], \mathbf{z}_{g,k}[\tau] \sim \mathcal{CN}(\mathbf{0}_N, \mathbf{I}_N)$ are the channel innovation components, such that $E[\mathbf{h}_k[n]\mathbf{z}_{h,k}^H[\tau]]$ and $E[\mathbf{g}_k[n]\mathbf{z}_{g,k}^H[\tau]]$ both equal the $N \times N$ all zero matrix. Throughout the paper, a variable with a bar over it, e.g., $\bar{\rho}$, is defined in terms of the variable without the bar, e.g., ρ , as $\bar{\rho} = \sqrt{1 - \rho^2}$. The random processes $\mathbf{z}_{h,k}[n]$ and $\mathbf{z}_{g,k}[n]$ are assumed to be stationary and ergodic. The channel is assumed to evolve either according to the Jakes' model [5], [15], or as a first order autoregressive (AR1) process [10]. In case the channel evolution is assumed to follow the Jakes' model [15], then $\rho[n] = J_0(2\pi f_d T_s n)$, 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 [16, Eq. (9.1.18)].

For OMA, we consider an N antenna BS serving $2K$ users indexed as $k \in \{1, 2, \dots, 2K\}$ with the channel from the BS to the k th user at the n th instant being given as $\sqrt{\beta_k}\mathbf{h}_k[n]$. As with NOMA, we consider a downlink TDD system with reverse link training, and assume channel reciprocity. However, since the available pilots are limited to K , the frame is split into two subframes. Each has a duration $\frac{T}{2}$, with the strong users being served in the first subframe and the weak users being served in the second subframe to avoid near far effects [14]. At the start of each subframe, the respective subset of K users transmit K orthogonal pilot sequences. The BS uses the received pilot signals to estimate the channel to each of the users, and form a beam to individual users. The channel vectors $\mathbf{h}_k[n]$ are assumed to age as (1).

III. CHANNEL ESTIMATION

For NOMA transmission, the first K slots in each frame are used for channel estimation. During this time, each user pair transmits orthogonal pilots to the BS. Both the users in the k th pair transmit identical pilots, $\psi_k[n]$, to the BS synchronously, but with different powers. The stronger user transmits at a

pilot power $\mathcal{E}_{p,k,h}$ and the weaker user transmits at a pilot power $\mathcal{E}_{p,k,g}$. The signal received at the i th BS antenna during the n th channel use ($1 \leq n \leq K$) is given by

$$y_i[n] = \sum_{k=1}^K (\sqrt{\beta_{k,h}}\mathcal{E}_{p,k,h}h_{ki}[n] + \sqrt{\beta_{k,g}}\mathcal{E}_{p,k,g}g_{ki}[n])\psi_k[n] + \sqrt{N_0}w_i[n], \quad (2)$$

with $w_i[n] \sim \mathcal{CN}(0, 1)$ being the additive white Gaussian noise (AWGN) at the BS. Note that, without loss of generality, we normalize the noise variance at the BS to unity.

Since data transmission starts from the $(K + 1)$ th instant, we are interested the channel estimate at the $(K + 1)$ th instant. The aging model can be used to express the channels between the i th BS antenna and the k th user pair at the n th ($1 \leq n \leq K$) instant in terms of the channels at the $(K + 1)$ th instant as

$$q_{ki}[n] = \rho^*[K + 1 - n]q_{ki}[K + 1] + \bar{\rho}[K + 1 - n]\zeta_{q,ki}[n], \quad (3)$$

with $q \in \{h, g\}$, $\zeta_{h,ki}[n], \zeta_{g,ki}[n] \sim \mathcal{CN}(0, 1)$ such that $E[q_{ik}[n]\zeta_{q,ki}^*[n]] = 0$ being the time reversed innovation components for the channels of the strong and weak users, respectively. Substituting these into (2), and defining $\lambda_k \triangleq \sqrt{\frac{\beta_{k,h}\mathcal{E}_{p,k,h}}{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g}}}$, $f_{ki}[n] \triangleq \lambda_k h_{ki}[n] + \bar{\lambda}_k g_{ki}[n]$ and $\xi_{ki}[n] \triangleq \lambda_k \zeta_{h,ki}[n] + \bar{\lambda}_k \zeta_{g,ki}[n]$ we obtain

$$y_i[n] = \sqrt{N_0}w_i[n] + \rho^*[K + 1 - n] \sum_{k=1}^K \sqrt{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g}} f_{ki}[K + 1]\psi_k[n] + \bar{\rho}[K + 1 - n] \sum_{k=1}^K \sqrt{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g}} \xi_{ki}[n]\psi_k[n].$$

We assume $u_{ki} = \sum_{n=1}^K c_k[n]y_i[n]$ with $c_k[n]$ chosen such that the orthogonality condition $\sum_{n=1}^K \psi_m[n]c_k[n]\rho[K + 1 - n] = \delta[m - k]$ described in [8] is satisfied. Letting $c_k[n] = (\rho[K + 1 - n])^{-1}\delta[n - k]$ and $\psi_m[n] = \delta[n - m]$, we can write the minimum mean squared error (MMSE) estimate \hat{f}_{ki} of $f_{ki}[K + 1]$ as

$$\hat{f}_{ki} = \rho[K + 1 - k] \frac{\sqrt{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g}}}{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g} + N_0} u_{ki}. \quad (4)$$

In general, the user pair scheduled in the first training slot experiences the largest estimation error due to channel aging. Considering this, it is easy to show that, regardless of the slot in which the k th user's pilots are scheduled, its (worst-case) mean squared channel estimation error including the effect of channel aging can be upper bounded as

$$\sigma_{ki}^2 \leq \frac{|\bar{\rho}[K]|^2 (\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g})}{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g} + N_0}. \quad (5)$$

Letting $b_k^2 \triangleq |\rho[K]|^2 \frac{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g}}{\beta_{k,h}\mathcal{E}_{p,k,h} + \beta_{k,g}\mathcal{E}_{p,k,g} + N_0}$ be the effective pilot signal-to-noise ratio (SNR) for the k th user pair, after some algebra, we can write the cross correlations between the estimate and the true channels as $E[\hat{f}_{ki}h_{ki}^*[K + 1]] = b_k^2 \lambda_k$; $E[\hat{f}_{ki}g_{ki}^*[K + 1]] = b_k^2 \bar{\lambda}_k$. Since the realizations of the strong user and the weak user channels are independent of

each other, we can express \hat{f}_{ki} as $\hat{f}_{ki} = \lambda_k \hat{h}_{ki} + \bar{\lambda}_k \hat{g}_{ki}$, such that, $E[\hat{g}_{ki} \hat{h}_{ki}^*] = 0$. Here \hat{h}_{ki} is the component of the channel estimate in the direction of the strong channel, and \hat{g}_{ki} is the component of the channel estimate in the direction of the weak channel. Therefore, the channel coefficient between the i th BS antenna and the strong and weak users of the k th user pair at the $(K + 1)$ th instant can be written as $q_{ki}[K + 1] = b_k \hat{q}_{ki} + \bar{b}_k \tilde{q}_{ki}$, with $E[\hat{q}_{ki} \tilde{q}_{ki}^*] = 0$, and $q \in \{h, g\}$.

Consequently, the channel coefficient at the n th instant can be expressed in terms of its component in the channel estimate available at the BS as $q_{ki}[n] = \rho[n - (K + 1)] b_k \hat{q}_{ki} + \rho[n - (K + 1)] \bar{b}_k \tilde{q}_{ki} + \bar{\rho}[n - (K + 1)] z_{q,ki}[n]$, $q \in \{h, g\}$.

In case of OMA transmission, during the first (second) subframe, the K strong (weak) users transmit orthogonal pilots to the BS. This is similar to the channel estimation in aging channels with K users [8], and it can be argued that the pilot SNR for the k th user can be lower bounded as $b_{k,q}^2 \geq \frac{[\rho[K]]^2 \beta_{k,q} \mathcal{E}_{p,k,q}}{\beta_{k,q} \mathcal{E}_{p,k,q} + N_0}$, $q \in \{h, g\}$.

IV. ACHIEVABLE DOWNLINK RATES

We first consider NOMA transmission. Let the symbols to be transmitted to the k th strong and weak user at the n th instant be denoted by $s_{k,h}[n]$ and $s_{k,g}[n]$, respectively. Also, letting the power allocated to the k th user pair be $\frac{\mathcal{E}_{s,k}}{N}$, with a fraction α_k^2 being allocated to the strong user, we can write the symbol transmitted over the k th data stream (corresponding to the data sent over a beam formed in the direction of the k th user pair) as

$$x_k[n] = \alpha_k s_{k,h}[n] + \bar{\alpha}_k s_{k,g}[n]. \quad (6)$$

Note that, since the power allocated by the BS to the weak user in a user pair is greater than that allotted to the strong user, $\bar{\alpha}_k > \alpha_k$.

Let $\mathbf{x}[n] \triangleq [x_1[n], x_2[n], \dots, x_K[n]]^T$ denote the concatenated $K \times 1$ symbol vector. Then, the symbol transmitted by the BS for MF precoding is $\mathbf{v}[n] = \hat{\mathbf{F}}^* \text{diag} \left(\sqrt{\frac{\mathcal{E}_s}{N}} \right) \mathbf{x}[n]$, with $\mathcal{E}_s = [\mathcal{E}_{s,1}, \mathcal{E}_{s,2}, \dots, \mathcal{E}_{s,K}]^T$, and $\hat{\mathbf{F}}$ being the channel estimate matrix with \hat{f}_{ki} as its (k, i) th entry.

The signal received at the k th strong/weak user at the n th instant ($K + 1 \leq n \leq T$) is

$$y_{k,q}[n] = \sqrt{\beta_{k,h}} \mathbf{q}_k^T [n] \mathbf{v}[n] + \sqrt{N_0} w_{k,q}[n], \quad (7)$$

$q \in \{g, h\}$. Substituting (III) and (6) into (7), the symbol received at the k th strong user can be expanded as

$$\begin{aligned} y_{k,h}[n] &= \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \rho[n - (K + 1)] b_k \lambda_k \alpha_k \hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* s_{k,h}[n] \\ &+ \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \rho[n - (K + 1)] b_k \lambda_k \bar{\alpha}_k \hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* s_{k,g}[n] \\ &+ \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \rho[n - (K + 1)] \bar{b}_k \lambda_k \tilde{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* x_k[n] \\ &+ \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \bar{\lambda}_k \mathbf{h}_k^T [n] \hat{\mathbf{g}}_k^* x_k[n] \\ &+ \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \bar{\rho}[n - (K + 1)] \lambda_k \mathbf{z}_{k,h}^T [n] \hat{\mathbf{h}}_k^* x_k[n] \end{aligned}$$

$$+ \sum_{\substack{l=1 \\ l \neq k}}^K \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,l}}{N} \mathbf{h}_k^T [n] \hat{\mathbf{f}}_l^* x_l[n] + \sqrt{N_0} w_{k,h}[n]. \quad (8)$$

The data rate of the weak user is selected such that its data symbol $s_{k,g}[n]$ is decodable when the interference from the strong user is treated as noise. Then, since the channel gain of the strong user is strictly better than that of the weak user, $s_{k,g}[n]$ is decodable at the strong user, allowing it to use SIC to subtract and remove the interference due to the weak user. Also, since no CSI is available at the users, it is assumed that the strong user only knows the expected value of the corresponding channel, and hence the signal at the k th user after subtracting $\rho[n - (K + 1)] \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} b_k \lambda_k \bar{\alpha}_k E[\hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^*] s_{k,h}[n]$ at the n th instant becomes

$$\begin{aligned} r_{k,h}[n] &= \sqrt{\frac{\beta_{k,h} \mathcal{E}_{s,k}}{N}} \rho[n - (K + 1)] b_k \lambda_k \alpha_k \\ &\times E \left[\hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* \right] s_{k,h}[n] + \sqrt{\frac{\beta_{k,h} \mathcal{E}_{s,k}}{N}} \rho[n - (K + 1)] \\ &\times b_k \lambda_k \alpha_k \left(\hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* - E \left[\hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* \right] \right) s_{k,h}[n] \\ &+ \sqrt{N_0} w_{k,h}[n] + \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \rho[n - (K + 1)] b_k \lambda_k \bar{\alpha}_k \\ &\times \left(\hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* - E \left[\hat{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* \right] \right) s_{k,g}[n] \\ &+ \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \rho[n - (K + 1)] \bar{b}_k \lambda_k \tilde{\mathbf{h}}_k^T \hat{\mathbf{h}}_k^* x_k[n] \\ &+ \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \bar{\lambda}_k \mathbf{h}_k^T [n] \hat{\mathbf{g}}_k^* x_k[n] + \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,k}}{N} \bar{\rho}[n - (K + 1)] \\ &\times \lambda_k \mathbf{z}_{k,h}^T [n] \hat{\mathbf{h}}_k^* x_k[n] + \sum_{\substack{l=1 \\ l \neq k}}^K \sqrt{\beta_{k,h}} \frac{\mathcal{E}_{s,l}}{N} \mathbf{h}_k^T [n] \hat{\mathbf{f}}_l^* x_l[n]. \end{aligned}$$

In the above, the first term corresponds to the desired signal, and all the other terms correspond to noise and interference. Note that the BS uses $\hat{\mathbf{f}}_k$ for precoding, and the UEs use the expected value of the corresponding (scaler) downlink channels for data detection. Therefore, explicit knowledge of $\hat{\mathbf{g}}_k$ and $\hat{\mathbf{h}}_k$ is not required anywhere. Since the channel estimate $\hat{\mathbf{h}}_k$ is uncorrelated with $\hat{\mathbf{g}}_k$, $\tilde{\mathbf{h}}_k$, \mathbf{z}_k , and $\hat{\mathbf{f}}_l$ for $l \neq k$, we can use the worst case noise theorem [17] and treat the interference as Gaussian noise. With some algebra [2], it can be shown that SINR and achievable rate at the k th strong receiver at the n th instant are given by

$$\gamma_{k,h}[n] = \frac{N \beta_{k,h} \mathcal{E}_{s,k} b_k^2 \lambda_k^2 \alpha_k^2 \rho^2 [n - (K + 1)]}{\beta_{k,h} \sum_{l=1}^K \mathcal{E}_{s,l} + N_0}, \quad (9)$$

$$R_{k,h}[n] = \frac{1}{T} \sum_{n=K+1}^N \log_2 \left(1 + \gamma_{k,h}[n] \right). \quad (10)$$

By following a similar analysis, the SINR at the k th weak receiver at the n th instant can be expressed as

$$\gamma_{k,g}[n] =$$

$$\frac{N\beta_{k,g}\mathcal{E}_{s,k}b_k^2\bar{\lambda}_k^2\bar{\alpha}_k^2\rho^2[n-(K+1)]}{\rho^2[n-(K+1)]N\beta_{k,g}\mathcal{E}_{s,k}b_k^2\bar{\lambda}_k^2\bar{\alpha}_k^2 + \beta_{k,g}\sum_{l=1}^K\mathcal{E}_{s,l} + N_0}$$

In the above, the first denominator term arises because interference is treated as noise. The achievable rate of the weak user over a frame of length T is thus given by

$$R_{k,g} = \frac{1}{T} \sum_{n=K+1}^T \log_2 \left(1 + \gamma_{k,g}[n] \right). \quad (11)$$

The above sum rate expressions are valid lower bounds on the capacity for any number of antennas. However, since they are based on the use-and-forget technique from [2], they are only tight when the number of antennas is reasonably large.

A. Achievable Rates With OMA

For OMA, the channel estimate matrix $\hat{\mathbf{H}}$ ($\hat{\mathbf{G}}$) for transmissions to the strong (weak) users is known at the BS. Defining $\mathbf{s}_h[k] = [s_{1,h}[n]s_{2,h}[n]\dots s_{K,h}[n]]^T$ ($\mathbf{s}_g[k] = [s_{1,g}[n]s_{2,g}[n]\dots s_{K,g}[n]]^T$) and as the concatenated symbol vector to the strong (weak) users, we can write the transmitted symbol with MF precoding as $\mathbf{v}[n] = \hat{\mathbf{Q}}^H \text{diag} \left(\sqrt{\frac{\mathcal{E}_{s,q}}{N}} \right) \mathbf{s}_q[n]$, where $\mathcal{E}_s = [\mathcal{E}_{s,q,1}\mathcal{E}_{s,q,2}\dots\mathcal{E}_{s,q,K}]$, $q \in \{g,h\}$, and $\hat{\mathbf{Q}} = \hat{\mathbf{H}}(\hat{\mathbf{G}})$ is the precoding matrix for the strong (weak) users. The SINR of the k th strong/weak user at the n th instant can then be computed as $\gamma_{k,q}[n] = \frac{\rho^2[n-(K+1)]N\beta_{k,q}\mathcal{E}_{s,q,k}b_k^2}{\beta_{k,q}\sum_{l=1}^K\mathcal{E}_{s,q,l} + N_0}$, for $q \in \{h,g\}$ [8]. The achievable rate over a frame containing a total of T symbols therefore becomes [8], $R_{k,q}[n] = \frac{1}{T} \sum_{n=K+1}^{T/2} \log_2 \left(1 + \gamma_{k,h}[n] \right) \left(1 + \gamma_{k,g} \left[\frac{T}{2} + n \right] \right)$.

V. NUMERICAL RESULTS AND DISCUSSION

In this section, we numerically quantify and compare the performance of the NOMA and OMA schemes. We consider a single cell system containing a 256 antenna BS serving 64 user pairs transmitting at a carrier frequency (f_c) of 2 GHz with a signal bandwidth 1 MHz. We also assume that the BS samples at the Nyquist rate of the complex baseband signal, i.e., at 1 MHz. For the purpose of these simulations, we assume both the pilot and data SNRs to be 10 dB. The channel is assumed to age according to the Jakes model, i.e. $\rho[n] = J_0(2\pi f_d T_s n)$, and the frame duration (T) is fixed at 1000 symbols [8]. For each drop of users (uniformly over the cell), the K users with the lowest path loss are designated as strong users, and the remainder designated as weak users.

In Fig. 1, we plot the average achievable throughput as a function of the user mobility, for the strong and weak users for different schemes, and with the BS's transmit power being equally allotted to all the users. We see that NOMA results in significantly larger per user rates for the strong users as compared to OMA, at a cost of marginally reduced rates for the weak users. However, since NOMA frames have a longer duration than OMA frames, the effect of channel aging is more pronounced in NOMA, rendering the gain in the sum rates due to NOMA negligibly small.

In Fig. 2, we plot the average throughput as a function of the user mobility, for the strong and weak users, with statistical channel inversion based power control [2]. In this

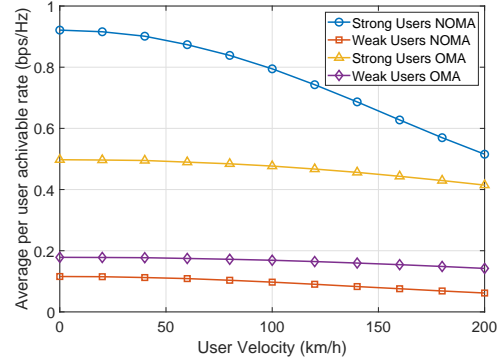


Fig. 1: Average achievable throughputs of NOMA and OMA at different user mobilities for 64 users with equal power transmission from the BS.

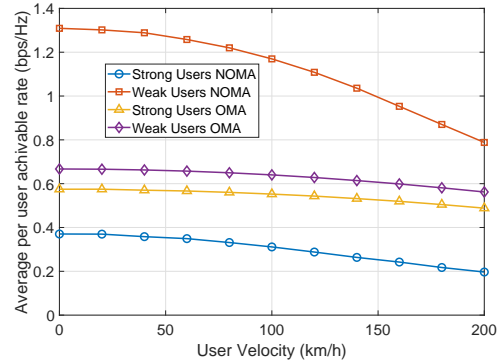


Fig. 2: Average achievable throughputs of NOMA and OMA at different user mobilities for 64 users statistical channel inversion based power control.

case, all the systems are interference-limited. Typically, the power levels of the weak users are comparable, while that of strong users see a larger variation. Hence, strong users tend to see larger interference compared to weak users. Under NOMA, the strong users bear the brunt of the interference, as they are interfered by the higher power signal transmitted to the weak users. The main contribution to the interference at the weak users is from the signals transmitted other weak users with comparable powers, as the strong users' signals are transmitted at lower powers. Hence, the relative behavior of strong and weak users is different from the equal power allocation case. Again, the effects of channel aging are observed to affect NOMA more than OMA, with the latter outperforming the former at higher user mobilities.

VI. CONCLUSIONS

In this paper, we considered the downlink performance of a massive MIMO system under channel aging and a constraint on the available number of orthogonal pilots for NOMA and OMA. We then used the derived bounds on channel estimation errors to obtain the achievable rates of both NOMA and OMA. Finally, using numerical results, we illustrated that at low user mobilities, the performance of the NOMA based scheme is superior to the OMA scheme. However, the relative performance improvement of NOMA over OMA reduces with increasing user mobility, with OMA becoming preferable to NOMA at user velocities beyond 150 km/h.

REFERENCES

- [1] 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.
- [2] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, Cambridge, UK, 2016.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] R. Chopra, C. R. Murthy, H. A. Suraweera, and E. G. Larsson, "Performance analysis of FDD massive MIMO systems under channel aging," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1094–1108, Feb. 2018.
- [9] S. Kashyap, C. Mollén, E. Björnson, and E. G. Larsson, "Performance analysis of TDD massive MIMO with Kalman channel prediction," in *Proc. Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2017)*, New Orleans, LA, March 2017, pp. 3554–3558.
- [10] 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.
- [11] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (noma) for cellular future radio access," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, June 2013, pp. 1–5.
- [12] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C. L. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.
- [13] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, Jan. 2016.
- [14] H. V. Cheng, E. Björnson, and E. G. Larsson, "Performance analysis of NOMA in training-based multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 372–385, Jan. 2018.
- [15] W. C. Jakes and D. C. Cox, Eds., *Microwave Mobile Communications*, IEEE Press, New York.
- [16] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. New York: Dover, 1964.
- [17] 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.