Active Attack on User Load Achieving Pilot Design in Massive MIMO Networks

Noman Akbar and Shihao Yan
Research School of Engineering, Australian National University, Acton, ACT, 2601, Australia
Emails: noman.akbar@anu.edu.au, shihao.yan@anu.edu.au

***Abstract*—In this paper, we propose an active attacking strategy on a massive multiple-input multiple-output (MIMO) network, where the pilot sequences are obtained using the user load-achieving pilot sequence design. The user load-achieving design ensures that the signal-to-interference-plus-noise ratio (SINR) requirements of all the users in the massive MIMO networks are guaranteed even in the presence of pilot contamination. However, this design has some vulnerabilities, such as one known pilot sequence and the correlation among the pilot sequences, that may be exploited by active attackers. In this work, we first identify the potential vulnerabilities in the user load-achieving pilot sequence design and then, accordingly, develop an active attacking strategy on the network. In the proposed attacking strategy, the active attackers transmit known pilot sequences in the uplink training and artificial noise in the downlink data transmission. Our examination demonstrates that the per-cell user load region is significantly reduced by the proposed attacking strategy. As a result of the reduced per-cell user load region, the SINR requirements of all the users are no longer guaranteed in the presence of the active attackers. Specifically, for the worst affected users the SINR requirements may not be ensured even with infinite antennas at the base station.**

I. Introduction

Massive multiple-input multiple-output (MIMO) technology is considered as one of the key enablers of the future fifth generation (5G) wireless networks. In a massive MIMO network, base stations (BSs) are equipped with hundreds of antennas. Massive MIMO provides a number of lucrative advantages over the conventional MIMO systems. One of these benefits is the increase in the spectral and energy efficiency [1]. In addition, the use of massive MIMO technologies achieves a higher throughput and reliability [2]. Another important advantage of massive MIMO is that the channels between BSs and users become increasingly orthogonal [3] as the number of antennas at BSs increases, which leads to the fact that the interference in the network will be significantly reduced. Recent research in the context of massive MIMO focused on resolving some specific key issues that limit the performance of massive MIMO. Among these issues, pilot contamination is considered as the most severe performance degrading factor in massive MIMO networks [3].

Pilot contamination occurs when the number of users in a cell is larger than the number of orthogonal pilot sequences, i.e., when it is not possible to allocate orthogonal pilot sequences to all the users and thus the pilot sequences are reused in the network. Pilot contamination is a performance bottleneck in massive MIMO networks [4]–[6], because it still exists even when the number of antennas at the BSs approaches infinity. As such, a lot of recent research works focused on mitigating or reducing the detrimental affects of pilot contamination in massive MIMO networks (e.g., [3], [7]–[11]). The recent research in pilot contamination can be generally categorized into five groups: protocol based methods [8], precoding based methods [9], angle-of-arrival based methods [12], blind methods [11], and pilot sequence design methods [3], [7], [13]–[15].

Recently, a user load-achieving pilot sequence design algorithm has been proposed for a multi-cell multi-user massive MIMO network [3] and the thorough performance analysis of this algorithm has been conducted [7]. The key idea of the user load-achieving pilot design is to first determine the user load region of the network under pilot contamination. Then, the algorithm allocates pilot sequences for all users in a distributed manner, which requires very little BS cooperation. The algorithm also allocates the downlink transmit power for all the users at BSs, such that the signal-to-interference-plus-noise ratio (SINR) requirements of all the users in the network can be guaranteed. We note that the user load-achieving pilot sequence design [7] is also referred to as the user capacity-achieving pilot sequence design in [3]. The main advantage of the user load-achieving pilot design is that it guarantees the SINR requirements for all the users in the network when some specific conditions are met. We would like to highlight that as long as the SINR requirements are within the per-cell user load region, the pilot sequence design and downlink transmit power allocation can ensure the SINR requirements of all the users. Otherwise, the pilot sequence design may not be feasible to guarantee the SINR requirements of all the users in the considered massive MIMO network.

In this paper, we propose a strategy for an active attacker who aims at exploiting the vulnerability in the user load-achieving pilot sequence design to degrade the performance of massive MIMO networks. In the proposed strategy, the attacker exploits the known properties of the user load-achieving pilot design to deliberately increase the pilot contamination in the uplink training phase. In addition, during the downlink transmission phase, the active attacker transmits artificial noise (AN) to increase the interference to each user in the network. Notably, the attack strategy carefully exploits the design of the user load-achieving pilot design and degrades its performance, such that the SINR requirements for all the users in the network are no longer guaranteed with a certain number of antennas at the BS. We recall that the goal of the user load-achieving pilot design is to ensure the SINR requirements for all the users in the network. This goal cannot be achieved in the the presence of the active attacker with the proposed attack strategy. The main contributions of this work are summarized as follows.

1) We identify potential vulnerabilities in the user load-achieving pilot sequence design. As shown in this work, these vulnerabilities can be exploited by an active attacker to significantly degrade the performance of a massive MIMO network, such that the SINR requirements of all the users in the network cannot be guaranteed by the user load-achieving pilot sequence design.
2) We propose an active attacking strategy on the user load-achieving pilot sequence design in massive MIMO networks. Our examination shows that the user load region achieved by the user load-achieving pilot sequence design is significantly reduced by the active attacker, such that the diverse range of SINR requirements is no longer supported. Specifically, with the active attack the SINR requirements for some users cannot be guaranteed even with an infinite number of antennas at the BSs.

II. User Load-Achieving Pilot Sequence Design and Its Vulnerabilities

For the sake of completeness, in this section we first present the process for the user load-achieving pilot sequence design [3], [7] and then identify its vulnerabilities that can be exploited by an active attacker.

A. User Load-Achieving Pilot Sequence Design

In the user load-achieving pilot sequence design, the user load is defined as the number of users that can be simultaneously served in a pilot-contaminated massive MIMO network, such that the SINR requirement of each individual user is guaranteed. To this end, the user load-achieving pilot design first determines the user load region of the network and then designs the pilot sequences accordingly. The key benefit achieved by this design is that it is capable of ensuring a diverse range of the SINR requirements for all the users simultaneously in the pilot-contaminated massive MIMO network. A thorough comparison of the design with existing pilot sequence designs demonstrates that the user load-achieving design can achieve a larger user load region and support a greater and diverse range of the SINR requirements. In addition, the user load-achieving pilot design guarantees the SINR requirement of all the users in the network with a finite N_t, where N_t is the number of antennas at each BS. Meanwhile, the existing pilot designs are unable to support such diverse SINR requirements even with an infinite N_t. We next present the vulnerabilities in the user load-achieving pilot sequence design, which can be exploited by an active attacker.

B. Vulnerabilities in the User Load-Achieving Pilot Design

We first find that the user load-achieving pilot sequence design always outputs the pilot sequence of the form $[1, 0, \cdots, 0]^T$ for one user in each cell. As such, when the length of the pilot sequence is known, the attacker can figure out the pilot sequence assigned to at least one user in each cell. Furthermore, we note that the user load-achieving pilot sequence design modifies the SINR requirements for all the users in the network such that the SINR requirements lie on the upper surface boundary of the user load region. This SINR modification ensures that the benefits offered by the large user load region of the user load-achieving pilot design are fully utilized. On the other side of the coin, this SINR modification introduces the potential vulnerability in the pilot design (i.e., the known sequence $[1, 0, \cdots, 0]^T$). Importantly, in the presence of active attackers, the SINR requirements may no longer remain inside the user load region. Consequently, the SINR requirements of all the users in the network will not be guaranteed. Another vulnerability in the user load-achieving design is that all the pilot sequences designed for the network are correlated with each other. As such, if the attacker even knows one pilot sequences in the network, it can potentially contaminate the channel estimates of all the users in the massive MIMO network. We note that an attacker only needs to know two network parameters for successfully exploiting the user load-achieving pilot design, i.e., the length of the pilot sequence, and the information that user load-achieving pilot design is being used in the network. We highlight that these parameters are easy to obtain in any network. Throughout this paper, we assume that the attacker has knowledge of these network parameters.

III. Multi-Cell Massive MIMO Networks with Active Attackers

In this section, we first detail the adopted system model and related assumptions. Then, we present the channel estimation and data transmission in the presence of the active attackers.

A. System Model and Adopted Assumptions

We consider a multi-cell multi-user massive MIMO network, where there are L cells and each of them has K single-antenna users, as depicted in Fig. 1. One BS is located in the center of each cell and is equipped with N_t antennas. We assume that there is one active attacker present in each cell and thus totally there are L active attackers in the network. We also assume that the communication channels in the network suffer from both large-scale and small-scale propagation effects. We denote the large-scale propagation factor from the jth user in the ith cell to the BS in the lth cell as $\beta_{i_j l}$. Additionally, we denote the small-scale propagation factor from the jth user in the ith cell to the nth BS antenna in the lth cell as $h_{i_j l_n}$. Consequently, the uplink propagation factor from the jth user in the ith cell to the nth BS antenna in the lth cell is represented as $\sqrt{\beta_{i_j l}} h_{i_j l_n}$. Furthermore, we assume the the small-scale propagation factor is Rayleigh distributed, i.e., $h_{i_j l_n} \sim \mathcal{CN}(0, 1)$. We assume that the network operates in the time-division-duplex (TDD) mode. The entire transmission, consisting of the uplink training and the downlink data transmission, occurs within one coherence block. As

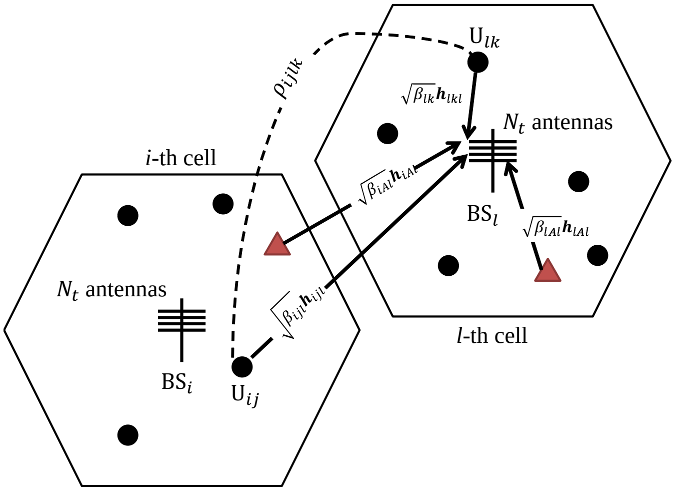
U$_{lk}$
ρ_{ijlk}
$\sqrt{\beta_{lk}}\boldsymbol{h}_{lkl}$
N_t antennas
i-th cell
$\sqrt{\beta_{iAl}}\boldsymbol{h}_{iAl}$
BS$_l$
$\sqrt{\beta_{lAl}}\boldsymbol{h}_{lAl}$
N_t antennas
$\sqrt{\beta_{ijl}}\boldsymbol{h}_{ijl}$
l-th cell
BS$_i$
U$_{ij}$

Fig. 1. An illustration of a multi-cell multi-user massive MIMO network in the presence of active attackers. Legitimate users are represented by black circles and active attackers are represented by red triangles.

such, we assume that the uplink and the downlink channels remain unchanged for the entire transmission. As a result, the channel estimates in the uplink can be utilized for the downlink precoding [16], [17]. We next detail the uplink channel estimation and downlink data transmission in the following two subsections.

B. Channel Estimation with Active Attackers

During the uplink training phase, the BS estimates the propagation factors from the users in a cell to the same-cell BS. Each user in a cell transmits a pre-assigned pilot sequence to the same-cell BS. We assume that all pilot sequences have unit energy. Additionally, the length of a pilot sequence is τ, which is assumed to be the same for all the users in the network. In the presence of active attackers, the observation received at the BS in the lth cell during the uplink training phase, denoted by $\mathbf{s}_l$, is given by

$$\mathbf{s}_l=\sum_{i=1}^{L}\sum_{j=1}^{K}\eta_{i_jl}\mathbf{Q}_{i_j}\mathbf{h}_{i_jl}+\sum_{m=1}^{L}\eta_{m_Al}\mathbf{Q}_{m_A}\mathbf{h}_{m_Al}+\mathbf{n}_l,\quad(1)$$

where $\eta_{i_jl}=\sqrt{p_{i_j}\beta_{i_jl}}$, $\mathbf{Q}_{i_j}=\mathbf{q}_{i_j}\otimes\mathbf{I}_{N_t}$ is the pilot matrix, $\mathbf{q}_{i_j}$ is the pilot sequence assigned to the jth user in the ith cell, $\otimes$ represents the Kronecker product, $\mathbf{I}_{N_t}$ denotes the $N_t\times N_t$ identity matrix, p_{i_j} is the pilot power for the jth user in the ith cell, $\mathbf{h}_{i_jl}=[h_{i_jl_1},h_{i_jl_2},\ldots,h_{i_jl_n}]^T$ is an $N_t\times 1$ uplink channel vector from the jth user in the ith cell to the BS in the lth cell, and $\mathbf{n}_l$ is the additive white Gaussian noise (AWGN) at the BS in the lth cell. We highlight that the second term in the observation given in (1) is due to the presence of the L active attackers in the network.

Exploiting the structure of the pilot sequences generated by the user load-achieving pilot sequence design, we assume that all the attackers in the network transmit the known pilot sequence, i.e., $\mathbf{Q}_{m_A}=\mathbf{Q}_{1_j}$, where the pilot sequence assigned to $\mathbf{Q}_{1_j}$ is in the form $[1,0,\cdots,0]^T$. Accordingly, the uplink channel from the kth user in the lth cell to the BS in the lth cell is obtained by utilizing the property of the pilot sequence matrix, given by $\mathbf{Q}_{l_k}^T\mathbf{Q}_{l_k}=\mathbf{I}_{N_t}$. Based on (1) and assuming that the uplink power control is enabled with $\eta_{l_kl}=1$, the BS in the lth cell obtains the least square (LS) channel estimate for $\mathbf{h}_{l_kl}$ as $\hat{\mathbf{g}}_{l_kl}=\mathbf{Q}_{l_k}^T\mathbf{s}_l$. We rewrite $\hat{\mathbf{g}}_{l_kl}$ as

$$\hat{\mathbf{g}}_{l_kl}=\mathbf{h}_{l_kl}+\sum_{i,j\neq l,k}\eta_{i_jl}\rho_{i_jl_k}\mathbf{h}_{i_jl}+\sum_{m=1}^{L}\eta_{m_Al}\rho_{m_Al_k}\mathbf{h}_{m_Al}+\bar{\mathbf{n}}_l,\tag{2}$$

where $\mathbf{Q}_{l_k}^T$ denotes the matrix transpose of $\mathbf{Q}_{l_k}$, $\bar{\mathbf{n}}_l=\mathbf{Q}_{l_k}^T\mathbf{n}_l$, $\sum_{i,j\neq l,k}=\sum_{i=1}^{L}\sum_{j=1}^{K}$ with the condition $(i,j)\neq(l,k)$, and $\rho_{i_jl_k}$ is the correlation coefficient between pilot sequences, defined as $\rho_{i_jl_k}=\mathbf{q}_{l_k}^T\mathbf{q}_{i_j}$, $k\in\{1,2,\ldots,K\}$. We highlight that the uplink power control is not applied for the active attackers because the BSs are not aware of their presence.

In the user load-achieving pilot design, we have the value range of the correlation $\rho_{i_jl_k}$ in (2) as $-1\leq\rho_{i_jl_k}\leq+1$. If all the users are assigned orthogonal pilot sequences, we have $\rho_{i_jl_k}=0$ and thus no pilot contamination. In massive MIMO networks, $\rho_{i_jl_k}$ is nonzero due to the limited number of orthogonal pilot sequences and thus pilot contamination always exists. With the knowledge of pilot sequence assigned to one user in each cell, the active attackers can deteriorate the quality of the channel estimate by increasing pilot contamination, which is confirmed by the second term in (2). We note that the pilot sequences obtained from the user load-achieving design are correlated with each other. As such, the attacks not only affect the users with the pilot sequence of the form $[1,0,\cdots,0]^T$, but all the users in the network.

C. Data Transmission via the Downlink

We now focus on the downlink data transmission in the massive MIMO network. We assume that the attackers are active during this phase and transmit AN, while each BS transmits the downlink data symbols to the same-cell users. We denote the data symbol intended for the kth user in the lth cell as x_{l_k}. We also assume that the downlink transmit power for the symbol x_{l_k} at the BS is given as $\mathbb{E}\left[x_{l_k}^Hx_{l_k}\right]=P_{l_k}$, where $\mathbb{E}[\cdot]$ denotes the expectation operation. Based on the channel estimates obtained during the uplink training and the reciprocity between the uplink and downlink channels for the TDD mode, the BS performs a linear precoding using a vector $\mathbf{a}$. Thus, the received signal at the kth user in the lth cell is given by

$$\hat{y}_{l_k}=\sum_{m=1}^{L}\sum_{n=1}^{K}\sqrt{\beta_{l_km}}\mathbf{h}_{l_km}^H\left(\mathbf{a}_{m_n}x_{m_n}\right)+w_{l_k},\tag{3}$$

where $w_{l_k}=\sum_{m=1}^{L}P_{m_A}w_{m_A}+\bar{w}_{l_k}$. Specifically, w_{m_A} is the AN generated by the active attacker in the mth cell with transmit power P_{m_A}, and $\bar{w}_{l_k}$ is the AWGN at the kth user in the lth cell. Assuming that users only have the statistical

$$\phi_{l_k,N_t} = \frac{\left(\mathbb{E}\left[\mathbf{h}_{l_k l}^H \mathbf{a}_{l_k}\right]\right)^2 \beta_{l_k l} P_{l_k}}{\text{var}\left[\mathbf{h}_{l_k l}^H \mathbf{a}_{l_k}\right] \beta_{l_k l} P_{l_k} + \sum_{m,n\neq l,k} \mathbb{E}\left[|\mathbf{h}_{l_k m}^H \mathbf{a}_{m_n}|^2\right] \beta_{l_k m} P_{m_n} + \sigma_w^2}. \tag{5}$$

$$\phi_{l_k,N_t} = \frac{\beta_{l_k l} P_{l_k}}{\left(\delta_{l_k} + \alpha_{l_k}\right)\left[\sum\limits_{m,n\neq l,k} \frac{\rho^2_{l_k m_n} \eta^2_{l_k m} \beta_{l_k m} P_{m_n}}{(\delta_{m_n} + \alpha_{m_n})} + \frac{1}{N_t}\left(\sum_{m=1}^{L}\sum_{n=1}^{K} \beta_{l_k m} P_{m_n} + \sigma_w^2\right)\right]}. \tag{11}$$

information of the channel [9], [18], we rewrite $\hat{y}_{l_k}$ in (3) as

$$\hat{y}_{l_k} = \sqrt{\beta_{l_k l}}\mathbb{E}\left[\mathbf{h}_{l_k l}^H \mathbf{a}_{l_k}\right] x_{l_k} + \sqrt{\beta_{l_k l}}\left(\mathbf{h}_{l_k l}^H \mathbf{a}_{l_k} - \mathbb{E}\left[\mathbf{h}_{l_k l}^H \mathbf{a}_{l_k}\right]\right) x_{l_k} + \sum_{m,n\neq l,k} \sqrt{\beta_{l_k m}} \mathbf{h}_{l_k m}^H \left(\mathbf{a}_{m_n} x_{m_n}\right) + w_{l_k}. \tag{4}$$

We now present the expressions for the achievable SINR for the kth user in the lth cell. We denote the achievable downlink SINR at the kth user in the lth cell by ϕ_{l_k,N_t}. We note that the first term in (4) represents the signal intended for the kth user in the lth cell. We assume that the remanding terms in (4) are uncorrelated with the intended signal and are treated as the effective noise. Accordingly, we express SINR ϕ_{l_k,N_t} as (5) given on the top of the page, where var$[\cdot]$ denotes the variance operation, and σ_w^2 denotes the variance of w_{l_k}. We note that the SINR expression (5) is a generalised expression valid for any type of linear precoding vector $\mathbf{a}_{l_k}$. Notably, we observe that the linear precoding vector is based on the channel estimates obtained by the uplink training phase. As such, pilot contamination in the uplink training affects the downlink data transmission.

In this work, we consider that the BS performs maximum-ratio transmission (MRT) precoding [3], [18], which is given by

$$\mathbf{a}_{l_k} = \frac{\hat{\mathbf{g}}_{l_k l}}{\|\hat{\mathbf{g}}_{l_k l}\|} = \frac{\hat{\mathbf{g}}_{l_k l}}{\sqrt{N_t\left(\hat{\mathbf{g}}_{l_k l}^H \hat{\mathbf{g}}_{l_k l}/N_t\right)}}, \tag{6}$$

where $\|\cdot\|$ denotes the l_2 norm. We now simplify the denominator in (6) by utilizing the fact that the channels in massive MIMO become increasingly orthogonal when the number of antennas at the BS (i.e., N_t) increases. This phenomenon is known as channel hardening and is represented as

$$\frac{1}{N_t}\mathbf{h}_{i_j i}^H \mathbf{h}_{l_k l} = \begin{cases} 1, & \forall\ (i,j) = (l,k) \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Using (7), we now simplify the denominator in (6) as

$$\begin{aligned} \frac{\hat{\mathbf{g}}_{l_k l}^H \hat{\mathbf{g}}_{l_k l}}{N_t} &= \sum_{i=1}^{L}\sum_{j=1}^{K} \eta^2_{i_j l}\rho^2_{i_j l_k} + \sum_{m=1}^{L} \eta^2_{m_A l}\rho^2_{m_A l_k} + \sigma^2_{n_l} \\ &= \left(\delta_{l_k} + \alpha_{l_k}\right), \end{aligned} \tag{8}$$

where $\alpha_{l_k} = \sum_{m=1}^{L} \eta^2_{m_A l}\rho^2_{m_A l_k}$ and

$$\delta_{l_k} = \sum_{i=1}^{L}\sum_{j=1}^{K} \eta^2_{i_j l}\rho^2_{i_j l_k} + \sigma^2_{n_l}. \tag{9}$$

Substituting (8) into (6), we obtain the precoding vector as

$$\mathbf{a}_{l_k} = \frac{\hat{\mathbf{g}}_{l_k l}}{\sqrt{N_t\left(\delta_{l_k} + \alpha_{l_k}\right)}}. \tag{10}$$

We highlight that α_{l_k} in (10) appears due to the pilot contamination caused by the active attackers. As such, the channel estimate $\hat{\mathbf{g}}_{l_k l}$ suffers from increased pilot contamination in the presence of the active attackers. We next present the closed form expression for ϕ_{l_k,N_t}, when the BSs adopt the precoding vector given in (10) and the channel estimates are obtained using the LS channel estimation given in (2).

Lemma 1: When the BSs adopt the precoding vector given in (10) and the channel estimates are obtained using the LS channel estimation given in (2), the SINR at kth user in the lth cell is given in (11) at the top of the page.

In massive MIMO, each BS is equipped with a large number of antennas. We next present the asymptotic SINR when $N_t \rightarrow \infty$. Following (11), as $N_t \rightarrow \infty$ the asymptotic SINR expression for ϕ_{l_k,N_t}, denoted by $\phi_{l_k,\infty}$, is given by

$$\phi_{l_k,\infty} = \frac{\beta_{l_k l} P_{l_k}}{\left(\delta_{l_k} + \alpha_{l_k}\right)\left(\sum\limits_{m=1}^{L}\sum\limits_{n=1}^{K} \frac{\rho^2_{l_k m_n}\eta^2_{l_k m}\beta_{l_k m} P_{m_n}}{(\delta_{m_n} + \alpha_{m_n})}\right) - \beta_{l_k l} P_{l_k}}. \tag{12}$$

We obtain some interesting observations from the the asymptotic SINR expression given by (12). The expression reveals that the pilot contamination still exists and limits the performance of massive MIMO, even when each BS is equipped with an infinite number of antennas. Furthermore, the increased pilot contamination due to the active attackers does not disappear in massive MIMO regime, i.e., α_{m_n} still exists when $N_t \rightarrow \infty$.

IV. User Load Region in Massive MIMO Networks with Active Attackers

In this section, we present the user load region of the massive MIMO network in the presence of the active attackers, while the user load region without the active attackers is provided as a benchmark.

The user load-achieving pilot sequence design [3], [7] guarantees the SINR requirements of all the users under the user load region. For comparison, we first represent the per-cell user load region of the massive MIMO network without active attackers, which is given by

$$\sum_{i=1}^{L}\sum_{j=1}^{K}\left(\frac{\gamma_{i_j}}{1+\gamma_{i_j}}\right) \leq \frac{\tau}{L}, \tag{13}$$

where γ_{i_j} is the SINR requirement of the ith user in cell. Furthermore, $\gamma_{i_j}/(1+\gamma_{i_j})$ denotes the effectiv width of the ith user in the jth cell. The bound on load signifies the region under which the user load is a which means that the SINR requirements of all the the network are guaranteed.

We now examine the impact of the active attacker user load region. As evident from (11), the active at the network lead to the reduction in the achievable SI derivations for user load region with the active attac omitted here due to space limitations. Following the approach as given in [3], [7], the per-cell user load r the presence of the active attackers is given by

$$\sum_{j=1}^{K}\left(\frac{\gamma_{i_j}}{1+\gamma_{i_j}}\right)+\left(\frac{\gamma_{m_A}}{1+\gamma_{m_A}}\right)\leq\frac{\tau}{L},$$

where $\gamma_{m_A}/(1+\gamma_{m_A})$ denotes the effective bandwid active attacker in the mth cell.

We note that having active attackers in the networ in the reduction of the user load region. The BSs d[illegible] user-load achieving pilot sequences based on the load region given in (13). We also note that the SINR requirements of all the users in the network can only be guaranteed within the user load region given in (13). With the active attackers, the user load region of the network is reduced to the one given in (14). As such, the SINR requirements of all the users in the network may not be supported by the BSs in the presence of the active attackers.

V. Numerical Results

In this section, we numerically evaluate the proposed active attack strategy and compare the performance of the massive MIMO network with and without the active attackers. Specifically, we present numerical results to demonstrate the performance degradation cased by the active attackers. Throughout this section, we consider a two-cell massive MIMO network, i.e., $L=2$. In addition, we set that there are eight users in the network and four users in each cell, i.e., $K=4$. Furthermore, the length of the pilot sequence used during the channel estimation is 3, i.e., $\tau=3$. Each BS designs the pilot sequences based on the user load-achieving pilot sequence design proposed in [3], [7]. Additionally, the downlink power for all the users in the network is set according to the pilot design [3], [7], where $P_{l_k}=\frac{\delta_{l_k}\gamma_{l_k}}{1+\gamma_{l_k}}$. Throughout this section, we assume that during the downlink transmission phase the active attackers transmit the AN with unit power, i.e., $P_{m_A}=1$.

We first compare the user load region of the network with and without the active attackers. In this comparison, we assume that the BS in each cell designs the pilot sequences separately. Accordingly, we compare the per-cell user load region. The SINR requirements for the users in the network are set as $\boldsymbol{\gamma}_1=\boldsymbol{\gamma}_2=[\gamma_{1_1},\gamma_{1_2},\gamma_{1_3},0.3]$. Additionally, we set that the effective bandwidth of the active attacker is 0.4. Fig.2 depicts the upper surface boundary of the per-cell user load region with and without the active attackers in the network. In this figure, the surface for the per-cell user load region without the active attackers, labeled as GWBE, is obtained using (13) and the surface for the per-cell user load region with the active attackers, labeled as GWBE$_A$, is obtained using (14). We note that the user load region is significantly reduced by the active attackers, even with only one attacker in a cell. Specifically, having one active attacker in each cell reduces the user load region by approximately 24.69%. The reduction in the user load region indicates that a group of users with high SINR requirements can no longer be successfully served in the pilot contaminated massive MIMO network in the presence of the active attackers.

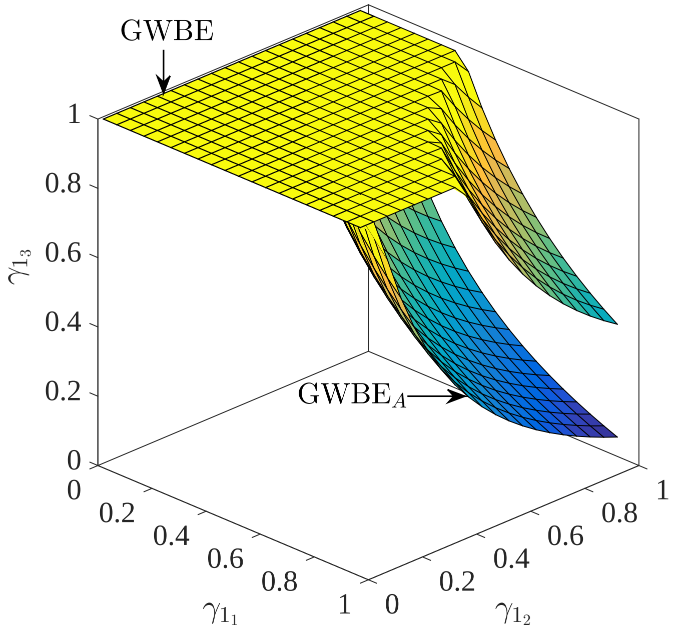

Fig. 2. The upper surface boundary of the per-cell user load regions versus the SINR requirements for the user load-achieving design with and without an active attacker.

In Fig. 3, we present the achievable downlink SINR performance with a finite number of antennas at the BSs with and without the active attackers. In this comparison, we generate results using (11) and consider that $L=2$, $\sigma_w^2=p_{l_k}=1$, and $\beta_{l_km}=1$, where $l=m$, $\beta_{l_km}=0.95$, and $l\neq m$. The SINR requirements for the users in the two cells are set as $\boldsymbol{\gamma}_1=[0.91,0.74,0.64,0.23]$, $\boldsymbol{\gamma}_2=[0.94,0.82,0.45,0.10]$. We note that the SINR requirements are carefully selected such that they remain inside the user load reagin of the network depicted as GWBE in Fig. 2. Additionally, we assume that the active attacker in each cell transmits the pilot sequence assigned to the first user in each cell. We clarify that the active attacker does not need to design the pilot sequences or know the full pilot sequence set designed for the entire network. Instead, the knowledge that the user load-achieving design is being used in the network is sufficient for the attacker to know the pilot sequences assigned to at least L users in the network. Fig. 3 depicts the achievable SINR for the two users

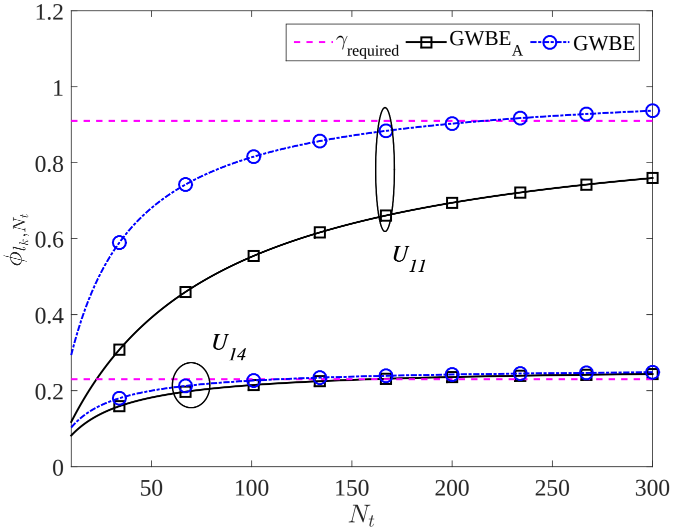

Fig. 3. The achievable SINR versus the number of antennas for the first user (U_{11}) and the fourth user (U_{14}) in the first cell for the user load-achieving design with and without an active attacker.

in the first cell with and without one active attacker in the cell. We highlight that without the active attacker in the cell, the SINR targets of all the users in the network can be guaranteed. However, in the presence of the active attacker in the cell, the SINR requirements of all the users in the cell cannot be satisfied. For example, we observe that in the presence of the active attacker, the achievable SINR for the first user in the cell reduces from 0.90 to 0.69 when $N_t = 200$. As such, there is an approximately 23.07% reduction in the achievable SINR for this user. Importantly, the achievable SINR never meets the SINR requirement even when the number of antennas at the BS is infinite. In other words, the first user never achieves the SINR target in the presence of the active attacker. We note that the pilot sequence used by the first user in the network is the same as the one used by the active attacker. One important observation found in Fig. 3 is that the impact of the active attacker on the fourth user (i.e., U_{14}) in the cell is negligible. This is due to the fact that the correlation between the pilot sequence used by the fourth user and the pilot sequence adopted by the attacker is very small.

VI. Conclusions

In this paper, we proposed an active attacking strategy on the user load-achieving pilot sequence design in a massive MIMO network. To this end, we first identified the potential vulnerabilities in the user load achieving pilot sequence design. Then, we proposed to increase the pilot contamination through transmitting known pilots by the active attackers in the uplink training. We demonstrated that the SINR requirements of all the users in the network are no longer guaranteed in the presence of the proposed active attacks. This is due to the fact that the per-cell user load region is significantly reduced by the active attackers. Specifically, the SINR requirements of the worst affected users by the attack may not be satisfied even with an infinite number of antennas at each BS in the massive MIMO network.

Acknowledgments

This work was supported by the Australian Research Council's Discovery Projects (DP150103905) and the Australian Government Research Training Program (RTP) Scholarship.

References

[1] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[2] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Di Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.

[3] N. Akbar, N. Yang, P. Sadeghi, and R. A. Kennedy, "Multi-cell multiuser massive MIMO networks: user capacity analysis and pilot design," *IEEE Trans. on Commun.*, vol. 64, no. 12, pp. 5064-5077, Dec. 2016.

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

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

[6] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July 2012, pp. 1137–1141.

[7] N. Akbar, N. Yang, P. Sadeghi, and R. A. Kennedy, "User load analysis and pilot sequence design for multi-cell massive MIMO networks," in *IEEE Global Communications Conference*, Dec. 2016, pp. 1-6.

[8] F. Fernandes, A. Ashikhmin, and T. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[9] J. Jose, A. Ashikhmin, T. 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.

[10] 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.

[11] R. Muller, L. Cottatellucci, and M. Vehkapera, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[12] N. Akbar, S. Yan, N. Yang, and J. Yuan, "Mitigating pilot contamination through location-aware pilot assignment in massive MIMO networks," in *IEEE Globecom Workshops (GC Wkshps)*, Dec 2016, pp. 1–6.

[13] S. Ulukus and R. Yates, "Iterative construction of optimum signature sequence sets in synchronous CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1989–1998, July 2001.

[14] P. Cotae, "Transmitter adaptation algorithm for multicellular synchronous DS-CDMA systems with multipath," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 1, pp. 94–103, Jan. 2006.

[15] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1984–1991, Sept. 1999.

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

[17] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. 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.

[18] J.-C. Shen, J. Zhang, and K. Letaief, "Downlink user capacity of massive MIMO under pilot contamination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3183–3193, June 2015.