

Data-Aided Secure Massive MIMO Transmission with Active Eavesdropping

Yongpeng Wu, Chao-Kai Wen, Wen Chen, Shi Jin, Robert Schober, and Giuseppe Caire

Abstract—In this paper, we study the design of secure communication for time division duplexing multi-cell multi-user massive multiple-input multiple-output (MIMO) systems with active eavesdropping. We assume that the eavesdropper actively attacks the uplink pilot transmission and the uplink data transmission before eavesdropping the downlink data transmission phase of the desired users. We exploit both the received pilots and data signals for uplink channel estimation. We show analytically that when the number of transmit antennas and the length of the data vector both tend to infinity, the signals of the desired user and the eavesdropper lie in different eigenspaces of the received signal matrix at the base station if their signal powers are different. This finding reveals that decreasing (instead of increasing) the desired user's signal power might be an effective approach to combat a strong active attack from an eavesdropper. Inspired by this result, we propose a data-aided secure downlink transmission scheme and derive an asymptotic achievable secrecy sum-rate expression for the proposed design. Numerical results indicate that under strong active attacks, the proposed design achieves significant secrecy rate gains compared to the conventional design employing matched filter precoding and artificial noise generation.

I. INTRODUCTION

Wireless networks are widely used in civilian and military applications and have become an indispensable part of our daily lives. Therefore, security is a critical issue for future wireless networks. Conventional security approaches based on cryptographic techniques have many well-known weaknesses. Therefore, new approaches to security based on information theoretical concepts, such as the secrecy capacity of the propagation channel, have been developed and are collectively referred to as physical layer security [1]–[4].

Massive MIMO is a promising approach for efficient transmission of massive amounts of information and is regarded as one of the “big three” 5G technologies [5]. Most studies on physical layer security in massive MIMO systems assume that the eavesdropper is passive and does not attack the

communication process of the systems [6]–[9]. However, a smart eavesdropper can perform the pilot contamination attack to jeopardize the channel estimation process at the base station [10]. Due to the channel hardening effect caused by large antenna arrays, the pilot contamination attack results in a serious secrecy threat to time division duplexing (TDD)-based massive MIMO systems [10].

The authors of [11] propose a secret key agreement protocol for single-cell multi-user massive MIMO systems under the pilot contamination attack. An estimator for the base station (BS) is designed to evaluate the information leakage. Then, the BS and the desired users perform secure communication by adjusting the length of the secrecy key based on the estimated information leakage. Other works have studied how to combat the pilot contamination attack. The authors of [12] investigate the pilot contamination attack problem for single-cell multi-user massive MIMO systems over independent and identically distributed (i.i.d.) fading channels. The eavesdropper is assumed to only know the pilot signal set whose size scales polynomially with the number of transmit antenna. For each transmission, the desired users randomly select certain pilot signals from this set, which are unknown to the eavesdropper. In this case, it is proved that the impact of the pilot contamination attack can be eliminated as the number of transmit antenna goes to infinity. For the more pessimistic assumption that the eavesdropper knows the exact pilot signals of the desired users for each transmission, the secrecy threat caused by the pilot contamination attack in multi-cell multi-user massive MIMO systems over correlated fading channels is analyzed in [10]. Based on this, three transmission strategies for combating the pilot contamination attack are proposed. However, the designs in [10] are not able to guarantee a high (or even a non-zero) secrecy rate for weakly correlated or i.i.d. fading channels under a strong pilot contamination attack.

In this paper, we investigate secure transmission for i.i.d. fading¹ TDD multi-cell multi-user massive MIMO systems under a strong active attack. We assume the system performs first uplink training followed by an uplink data transmission phase and a downlink data transmission phase. The eavesdropper jams the uplink training phase and the uplink data transmission phase and then eavesdrops the downlink data transmission².

We utilize the uplink transmission data to aid the channel estimation at the BS. Then, based on the estimated channels, the BS designs precoders for the downlink transmission.

This paper makes the following key contributions:

¹For simplicity of presentation, we assume i.i.d. fading to present the basic idea of data-aided secure massive MIMO transmission. The results can be extended to the general case of correlated fading channels by combining the techniques in [10] with those in this paper. This will be considered in extended journal version of this paper.

The work of Y. Wu was supported in part by NSFC No. 61701301. The work of C.-K. Wen was supported by the Ministry of Science and Technology of Taiwan under Grant MOST 106-2221-E-110-019. The work of W. Chen is supported by Shanghai STCSM 16JC1402900 and 17510740700, by National Science and Technology Major Project 2017ZX03001002-005 and 2018ZX03001009-002, by NSF China 61671294, and by Guangxi NSF 2015GXNSFDA139037. The work of S. Jin was supported in part by the NSFC under Grant 61531011.

Y. Wu and W. Chen are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Minhang 200240, China (e-mail: yongpeng.wu@sjtu.edu.cn; wenchen@sjtu.edu.cn).

C. K. Wen is with the Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan (Email: chaokai.wen@mail.nsysu.edu.tw).

S. Jin is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, 210096, P. R. China. (Emails: jinshi@seu.edu.cn).

R. Schober is with Institute for Digital Communications, Universität Erlangen-Nürnberg, Cauerstrasse 7, D-91058 Erlangen, Germany (Email: schober@lnt.de).

G. Caire is with Institute for Telecommunication Systems, Technical University Berlin, Einsteinufer 25, 10587 Berlin, Germany (Email: caire@tu-berlin.de).

- 1) We prove that when the number of transmit antennas and the amount transmitted data both approach infinity, the desired users' and the eavesdropper's signals lie in different eigenspaces of the uplink received signal matrix due to their power differences. Our results reveal that increasing the power gap between the desired users' and the eavesdropper's signals is beneficial for separating the desired users and the eavesdropper. This implies that when facing a strong active attack, decreasing (instead of increasing) the desired users' signal power could be an effective approach to enable secure communication.
- 2) Inspired by this observation, we propose a joint uplink and downlink data-aided transmission scheme to combat strong active attacks from an eavesdropper. Then, we derive an asymptotic achievable secrecy sum-rate expression for this scheme. The derived expression indicates that the impact of an active attack on the uplink transmission can be completely eliminated by the proposed design.
- 3) Our numerical results reveal that the proposed design achieves a good secrecy performance under strong active attacks, while the conventional design employing matched filter precoding and artificial noise generation (MF-AN) [10] is not able to guarantee secure communication in this case.

Notation: Vectors are denoted by lower-case bold-face letters; matrices are denoted by upper-case bold-face letters. Superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ stand for the matrix transpose, conjugate, and conjugate-transpose operations, respectively. We use $\text{tr}(\mathbf{A})$ and \mathbf{A}^{-1} to denote the trace and the inverse of matrix \mathbf{A} , respectively. $\text{diag}\{\mathbf{b}\}$ denotes a diagonal matrix with the elements of vector \mathbf{b} on its main diagonal. $\text{Diag}\{\mathbf{B}\}$ denotes a diagonal matrix containing the diagonal elements of matrix \mathbf{B} on the main diagonal. The $M \times M$ identity matrix is denoted by \mathbf{I}_M , and the $M \times N$ all-zero matrix and the $N \times 1$ all-zero vector are denoted by $\mathbf{0}$. The fields of complex and real numbers are denoted by \mathbb{C} and \mathbb{R} , respectively. $E[\cdot]$ denotes statistical expectation. $[\mathbf{A}]_{mn}$ denotes the element in the m th row and n th column of matrix \mathbf{A} . $[\mathbf{a}]_m$ denotes the m th entry of vector \mathbf{a} . \otimes denotes the Kronecker product. $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_N)$ denotes a circularly symmetric complex vector $\mathbf{x} \in \mathbb{C}^{N \times 1}$ with zero mean and covariance matrix \mathbf{R}_N . $\text{var}(a)$ denotes the variance of random variable a . $[x]^+$ stands for $\max\{0, x\}$. $a \gg b$ means that a is much larger than b .

II. UPLINK TRANSMISSION

Throughout the paper, we adopt the following transmission protocol. We assume the uplink transmission phase, composing the uplink training and the uplink data transmission, which is followed by a downlink data transmission phase.

We assume the main objective of the eavesdropper is to eavesdrop the downlink data. The eavesdropper chooses to attack the uplink transmission phase to impair the channel estimation phase at the BS. The resulting mismatched channel estimation will increase the information leakage in the subsequent downlink transmission. In the downlink transmission phase, the eavesdropper does not attack but focuses on eavesdropping the data.

We study a multi-cell multi-user system with $L + 1$ cells. We assume an N_t -antenna BS and K single-antenna users are present in each cell. The cells are indexed by $l = (0, \dots, L)$,

where cell $l = 0$ is the cell of interest. We assume an N_e -antenna active eavesdropper³ is located in the cell of interest and attempts to eavesdrop the data intended for all users in the cell. The eavesdropper sends pilot signals and artificial noise to interfere channel estimation and uplink data transmission⁴, respectively. Let T and τ denote the coherence time of the channel and the length of the pilot signal, respectively. Then, for uplink transmission, the received pilot signal matrix $\mathbf{Y}_p^m \in \mathbb{C}^{N_t \times \tau}$ and the received data signal matrix $\mathbf{Y}_d^m \in \mathbb{C}^{N_t \times (T-\tau)}$ at the BS in cell m are given by⁵

$$\mathbf{Y}_p^m = \sqrt{P_0} \sum_{k=1}^K \mathbf{h}_{0k}^m \boldsymbol{\omega}_k^T + \sum_{l=1}^L \sum_{k=1}^K \sqrt{P_l} \mathbf{h}_{lk}^m \boldsymbol{\omega}_k^T + \sqrt{\frac{P_e}{KN_e}} \mathbf{H}_e^m \sum_{k=1}^K \mathbf{W}_k + \mathbf{N}_p^m \quad (1)$$

$$\mathbf{Y}_d^m = \sqrt{P_0} \sum_{k=1}^K \mathbf{h}_{0k}^m \mathbf{d}_{0k}^T + \sum_{l=1}^L \sum_{k=1}^K \sqrt{P_l} \mathbf{h}_{lk}^m \mathbf{d}_{lk}^T + \sqrt{\frac{P_e}{N_e}} \mathbf{H}_e^m \mathbf{A} + \mathbf{N}_d^m \quad (2)$$

where $P_0, \boldsymbol{\omega}_k \in \mathbb{C}^{\tau \times 1}$, and $\mathbf{d}_{0k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{T-\tau})$ denote the average transmit power, the pilot sequence, and the uplink transmission data of the k th user in cell of interest, respectively. It is assumed that the same K orthogonal pilot sequences are used in each cell where $\boldsymbol{\omega}_k^H \boldsymbol{\omega}_k = \tau$ and $\boldsymbol{\omega}_k^H \boldsymbol{\omega}_l = 0$. P_l and \mathbf{d}_{lk} denote the average transmit power and the uplink transmission data of the k th user in the l th cell, respectively. $\mathbf{h}_{lk}^p \sim \mathcal{CN}(\mathbf{0}, \beta_{lk}^p \mathbf{I}_{N_t})$ denotes the channel between the k th user in the l th cell and the BS in the p th cell, where β_{lk}^p is the corresponding large-scale path loss. \mathbf{H}_e^l and P_e denote the channel between the eavesdropper and the base station in the l th cell and the average transmit power of the eavesdropper, respectively. We assume the columns of \mathbf{H}_e^l are i.i.d. with Gaussian distribution $\mathcal{CN}(\mathbf{0}, \beta_e^l \mathbf{I}_{N_t})$, where β_e^l is the large-scale path loss for the eavesdropper. For the training phase, the eavesdropper attacks all the users in cell of interest. Therefore, it uses the attacking pilot sequences $\sum_{k=1}^K \mathbf{W}_k$ [12], where $\mathbf{W}_k = [\boldsymbol{\omega}_k \cdots \boldsymbol{\omega}_k]^T \in \mathbb{C}^{N_t \times \tau}$. For the uplink data transmission phase, the eavesdropper generates artificial noise $\mathbf{A} \in \mathbb{C}^{N_t \times (T-\tau)}$, whose elements conform i.i.d. standard Gaussian distribution. $\mathbf{N}_p^m \in \mathbb{C}^{N_t \times \tau}$ and $\mathbf{N}_d^m \in \mathbb{C}^{N_t \times (T-\tau)}$ are noise matrices whose columns are i.i.d. Gaussian distributed with $\mathcal{CN}(\mathbf{0}, N_0 \mathbf{I}_{N_t})$.

We define $\mathbf{Y}_0 = [\mathbf{Y}_p^0 \ \mathbf{Y}_d^0]$ and the eigenvalue decomposition $\frac{1}{TN_t} \mathbf{Y}_0 \mathbf{Y}_0^H = [\mathbf{v}_1, \dots, \mathbf{v}_{N_t}] \boldsymbol{\Sigma} [\mathbf{v}_1, \dots, \mathbf{v}_{N_t}]^H$, where the eigenvalues on the main diagonal of matrix $\boldsymbol{\Sigma}$ are originated in ascending order. For the following, we make the important assumption that due to the strong active attack and the large-scale path loss difference between the cell of interest and other cells, $P_e \beta_e^0$, $P_0 \beta_{0k}^0$, and $P_l \beta_{lk}^0$ have the

³An N_e -antenna eavesdropper is equivalent to N_e cooperative single-antenna eavesdroppers.

⁴We note that if the eavesdropper only attacks the channel estimation phase and remains silent during the uplink data transmission, then the impact of this attack can be easily eliminated with the joint channel estimation and data detection scheme in [13]. Therefore, a smart eavesdropper will attack the entire uplink transmission.

⁵For notation simplicity, we assume the users in each cell use the same transmit power [6]. Following the similar techniques in this paper, the results can be easily extended to the case of different transmit powers of the users in each cell.

relationship $P_e\beta_e^0 \gg P_0\beta_{0k}^0 \gg P_l\beta_{lk}^0$. Let $M = (L+1)K + N_e$ and vector $(\theta_1, \dots, \theta_M)$ has the same element as vector $(P_1\beta_{11}^0, \dots, P_L\beta_{LK}^0, P_0\beta_{01}^0, \dots, P_0\beta_{0K}^0, P_e\beta_e^0, \dots, P_e\beta_e^0)$ but with the elements originated in ascending order whose index $1 \leq i_1 \leq i_2 \leq \dots \leq i_K \leq M$ satisfies $\theta_{i_k} = P_0\beta_{0k}^0$, $k = 1, 2, \dots, K$. Define $\mathbf{V}_{eq}^0 = [\mathbf{v}_{N_t-M+i_1}, \mathbf{v}_{N_t-M+i_2}, \dots, \mathbf{v}_{N_t-M+i_K}]$. Define $\mathbf{H}_0 = [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]$ and $\mathbf{H}_I = [\mathbf{h}_{11}^0, \dots, \mathbf{h}_{1K}^0, \dots, \mathbf{h}_{L1}^0, \dots, \mathbf{h}_{LK}^0]$. Then, we have the following theorem.

Theorem 1. Let $\mathbf{Z}_{0p} = \frac{1}{\sqrt{TN_t}} (\mathbf{V}_{eq}^0)^H \mathbf{Y}_p^0 = [\mathbf{z}_{0p,1}, \dots, \mathbf{z}_{0p,K}]$ and $\mathbf{H}_{eq}^0 = \frac{1}{\sqrt{TN_t}} (\mathbf{V}_{eq}^0)^H \mathbf{H}_0 = [\mathbf{h}_{eq,01}, \dots, \mathbf{h}_{eq,0K}]$. Then, when $T \rightarrow \infty$ and $N_t \rightarrow \infty$, the minimum mean square error (MMSE) estimate $\hat{\mathbf{h}}_{eq,0k}$ of $\mathbf{h}_{eq,0k}$ based on \mathbf{Z}_{0p} is given by

$$\hat{\mathbf{h}}_{eq,0k} = \frac{\sqrt{P_0}}{P_0\tau + N_0} (\sqrt{P_0}\tau \mathbf{h}_{eq,0k} + \mathbf{n}_{eq}) \quad (3)$$

where $\mathbf{n}_{eq} = \mathbf{V}_{eq}^0 \tilde{\mathbf{n}}_{eq}$ and $\tilde{\mathbf{n}}_{eq} \sim \mathcal{CN}(0, \tau N_0 \mathbf{I}_{N_t})$.

Proof. Please refer to Appendix A. \square

Remark 1: The basic intuition behind Theorem 1 is that when $T \rightarrow \infty$ and $N_t \rightarrow \infty$, each channel tends to be an eigenvector of the received signal matrix. As a result, we project the received signal matrix along the eigenspace which corresponds to the desired users' channel. In this case, the impact of the strong active attack can be effectively eliminated.

Remark 2: In Theorem 1, we assume that the coherence time of the channel is significantly larger than the symbol duration [14]. This assumption can be justified based on the expression for the coherence time in [14, Eq. (1)]. For typical speeds of mobile users and typical symbol duration, the coherence time can be more than hundreds symbol durations or even more.

Remark 3: The simulation results in Section IV indicate that a sufficient power gap between P_0 and P_e can guarantee a good secrecy performance when the number of transmit antennas and the coherence time of the channel are large but not infinite. We note that allocating more power to the desired users to combat a strong active attack is not needed. In contrast, the larger gap between $P_0\beta_{0k}^0$ and $P_e\beta_e^0$ will be beneficial to approach the channel estimation result in Theorem 1. This implies that *decreasing* the power of the desired users can be an effective secure transmission strategy under a strong active attack.

Remark 4: We can use large dimension random matrix theory [15] to obtain a more accurate approximation for the eigenvalue distribution of $\frac{1}{TN_t} \mathbf{Y} \mathbf{Y}^H$ for the case when N_t and T are large but not infinite. Then, power design policies for P_0 , P_l , and P_e can be obtained. This will be discussed in the extended journal version of this work.

Based on Theorem 1, we can design the precoders for downlink transmission.

III. DOWNLINK TRANSMISSION

In this section, we consider the downlink transmission. We assume the BSs in all $L+1$ cells perform channel estimation according to Theorem 1 by replacing $\hat{\mathbf{h}}_{eq,0k}$, $\mathbf{h}_{eq,0k}$, P_0 , and \mathbf{V}_{eq}^0 by $\hat{\mathbf{h}}_{eq,lk}$, $\mathbf{h}_{eq,lk}$, P_l , and \mathbf{V}_{eq}^l , respectively. Then, the l th BS designs the transmit signal as follows

$$\mathbf{x}_l = \sqrt{P} \sum_{k=1}^K \mathbf{t}_{lk} s_{lk}, \quad l = 0, \dots, L, \quad (4)$$

where P is the downlink transmission power, $\mathbf{t}_{lk} = (\mathbf{V}_{eq}^l)^H \frac{\hat{\mathbf{h}}_{eq,lk}}{\|\hat{\mathbf{h}}_{eq,lk}\|}$, and s_{lk} is the downlink transmitted signal for the k th user in the l th cell.

For the proposed precoder design, the base station only needs to know the statistical channel state information of the eavesdropper $P_e\beta_e^0$ in order to construct \mathbf{V}_0 . This assumption is justified in [10].

Because each user in the cell of interest has the risk of being eavesdropped, an achievable ergodic secrecy sum-rate can be expressed as [16]

$$R_{\text{sec}} = \sum_{k=1}^K [R_k - C_k^{\text{eve}}]^+ \quad (5)$$

where R_k and C_k^{eve} denote an achievable ergodic rate between the BS and the k th user and the ergodic capacity between the BS and the eavesdropper seeking to decode the information of the k th user, respectively.

The received signal y_{0k} at the k th user in the cell of interest is given by

$$\begin{aligned} y_{0k} &= \sum_{l=0}^L (\mathbf{h}_{lk}^0)^H \mathbf{x}_l + n_d \\ &= \sqrt{P} (\mathbf{h}_{0k}^0)^H (\mathbf{V}_{eq}^0)^H \frac{\hat{\mathbf{h}}_{eq,0k}}{\|\hat{\mathbf{h}}_{eq,0k}\|} s_{0k} \\ &\quad + \sqrt{P} (\mathbf{h}_{0k}^0)^H (\mathbf{V}_{eq}^0)^H \sum_{t=1, t \neq k}^K \frac{\hat{\mathbf{h}}_{eq,0t}}{\|\hat{\mathbf{h}}_{eq,0t}\|} s_{0t} \\ &\quad + \sqrt{P} \sum_{l=1}^L (\mathbf{h}_{lk}^0)^H (\mathbf{V}_{eq}^l)^H \sum_{t=1}^K \frac{\hat{\mathbf{h}}_{eq,lt}}{\|\hat{\mathbf{h}}_{eq,lt}\|} s_{lt} + n_d. \end{aligned} \quad (6)$$

where $n_d \sim \mathcal{CN}(0, N_{0d})$ is the noise in the downlink transmission.

We use a lower bound for the achievable ergodic rate R_k as follows [17]

$$\bar{R}_k = \log(1 + \gamma_k) \quad (7)$$

where

$$\gamma_k = \frac{|E[g_{0k,k}^0]|^2}{N_{0d} + \text{var}(g_{0k,k}^0) + \sum_{t=1, t \neq k}^K E[|g_{0t,k}^0|^2] + \sum_{l=1}^L \sum_{t=1}^K E[|g_{lt,k}^0|^2]} \quad (8)$$

and $g_{lt,k}^0 = \sqrt{P} (\mathbf{h}_{lk}^0)^H (\mathbf{V}_{eq}^l)^H \frac{\hat{\mathbf{h}}_{eq,lt}}{\|\hat{\mathbf{h}}_{eq,lt}\|}$.

For C_k^{eve} , we adopt the same pessimistic assumption as in [10], i.e., we assume that the eavesdropper can eliminate all interference from intra and inter-cell users to obtain an upper bound of C_k^{eve} as follows

$$C_{k,\text{upper}}^{\text{eve}} = E \left[\log_2 \left(1 + \frac{P}{N_0} \frac{g_{\text{eve}}}{\|\hat{\mathbf{h}}_{eq,0k}\|^2} \right) \right] \quad (9)$$

where

$$g_{\text{eve}} = (\hat{\mathbf{h}}_{eq,0k})^H (\mathbf{V}_{eq}^0)^H (\mathbf{H}_e^0)^H (\mathbf{H}_e^0)^H (\mathbf{V}_{eq}^0)^H \hat{\mathbf{h}}_{eq,0k}. \quad (10)$$

Based on (5), (7), and (9), we have the following theorem.

Theorem 2. For the considered multi-cell multi-user massive MIMO system, an asymptotic achievable secrecy sum-rate for

the transmit signal design in (4) is given by

$$R_{\text{sec, ach}} \xrightarrow{N_t \rightarrow \infty} \sum_{k=1}^K \log(1 + \bar{\gamma}_k) \quad (11)$$

where

$$\bar{\gamma}_k = \frac{a_1}{N_{0d} + P(a_2 - a_1) + P(K-1)\beta_{0k}^0 + PK \sum_{l=1}^L \beta_{lk}^0} \quad (12)$$

$$a_1 = \frac{P_0 \tau (P_0 \tau \beta_{0k}^0 (N_t + K - 1) + K N_0)}{(P_0 \tau + N_0)^2} \quad (13)$$

$a_2 =$

$$\frac{P_0 \tau (\beta_{0k}^0 N_t + \beta_{0k}^0 (K - 1))^2 + N_0 (N_t \beta_{0k}^0 + 3(K - 1) \beta_{0k}^0)}{P_0 \tau \beta_{0k}^0 (N_t + K - 1) + N_0} \quad (14)$$

Proof. Please refer to Appendix B. \square

Theorem 2 is a general expression which is valid for arbitrary K and L . Also, Theorem 2 indicates that when N_t tends to infinity, the impact of the active attack from the eavesdropper disappears if the proposed joint uplink and downlink transmission design is adopted.

IV. NUMERICAL RESULTS

In this section, we present numerical results to examine the proposed design and the obtained analytical results. We set $L = 3$, $N_t = 128$, $\beta_{0k}^0 = 1$, $k = 1, \dots, K$, $\beta_{lk}^0 = 0.2$, $k = 1, \dots, K$, $l = 1, \dots, L$, and $P_0 = P_1 = \dots = P_L$. We define the signal-to-noise ratio (SNR) as $\text{SNR} = P/N_{0d}$. Also, we define $\rho = P_E/P_0 K$.

Figure 1 plots the asymptotic and exact secrecy rate performance vs. the SNR for $T = 1024$, $P_0/N_0 = 5$ dB, $\rho = 30$, and different numbers of users, respectively. The exact secrecy rate is obtained based on Monte Carlo simulation of (8) and (9). We note from Figure 1 that the asymptotic secrecy rate in Theorem 2 provides a good estimate for the exact secrecy rate.

Figure 2 compares the secrecy performance of the proposed design and the MF-AN design in [10] for large but finite N_t and T as a function of ρ for $K = 5$, $P_0/N_0 = 5$ dB, $\text{SNR} = 5$ dB, and different values of T . We keep P_0 constant and increase P_e to increase ρ . We observe from Figure 2 that when the power of the active attack is strong, the MF-AN design cannot provide a non-zero secrecy rate. However, our proposed design performs well in the entire considered range of ρ . As ρ increases, the gap between $P_e \beta_e$ and $P_0 \beta_{0k}^0$ increases as well. Therefore, the secrecy rate increases with ρ for the proposed design. Moreover, Figure 2 reveals that increasing T is beneficial for the secrecy performance of the proposed design.

V. CONCLUSIONS

In this paper, we have proposed a data-aided secure transmission scheme for multi-cell multi-user massive MIMO systems which are under a strong active attack. We exploit the received uplink data signal for joint uplink channel estimation and secure downlink transmission. We show analytically that when the number of transmit antennas and the length of the data vector both approach infinity, the proposed design can effectively eliminate the impact of an active attack by an eavesdropper. Numerical results validate our theoretical

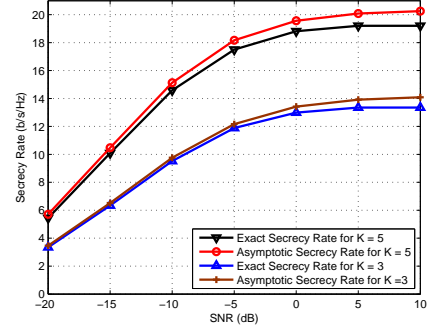


Fig. 1: Secrecy rate vs. SNR for $T = 1024$, $P_0/N_0 = 5$ dB, $\rho = 30$, and different numbers of users

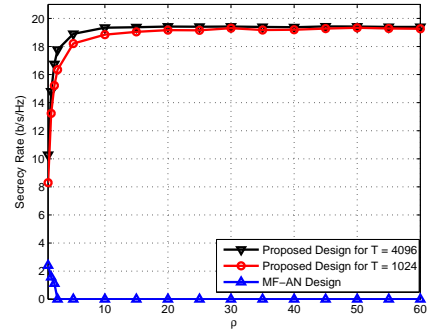


Fig. 2: Secrecy rate vs. ρ for $K = 5$, $P_0/N_0 = 5$ dB, $\text{SNR} = 5$ dB, and different values of T

analysis and demonstrate the effectiveness of the proposed design under strong active attacks.

APPENDIX A PROOF OF THEOREM 1

We define $\mathbf{\Omega}_0 = [\omega_1, \dots, \omega_K]^T$, $\mathbf{D}_0 = \sqrt{P_0} [\mathbf{d}_{01}, \dots, \mathbf{d}_{0K}]^T$, $\mathbf{\Omega}_L = [\sqrt{P_1} \mathbf{\Omega}_0^T, \dots, \sqrt{P_L} \mathbf{\Omega}_0^T]^T$, $\mathbf{D}_L = [\sqrt{P_1} \mathbf{d}_{11}, \dots, \sqrt{P_1} \mathbf{d}_{1K}, \dots, \sqrt{P_L} \mathbf{d}_{L1}, \dots, \sqrt{P_L} \mathbf{d}_{LK}]^T$, $\mathbf{X}_0 = [\sqrt{P_0} \mathbf{\Omega}_0 \quad \mathbf{D}_0]$, $\mathbf{X}_I = [\mathbf{\Omega}_L \quad \mathbf{D}_L]$, $\mathbf{X}_e = \left[\sqrt{\frac{P_E}{KN_e}} \sum_{k=1}^K \mathbf{W}_k \quad \sqrt{\frac{P_E}{N_e}} \mathbf{A} \right]$.

Based on (1) and (2), the received signal \mathbf{Y}_0 can be re-expressed as

$$\mathbf{Y}_0 = \mathbf{H}_0 \mathbf{X}_0 + \mathbf{H}_I \mathbf{X}_I + \mathbf{H}_e^0 \mathbf{X}_e + \mathbf{N} \quad (15)$$

where $\mathbf{N} = [\mathbf{N}_p^0 \quad \mathbf{N}_d^0]$.

When $T \rightarrow \infty$, based on [18, Corollary 1], we obtain (16) given at the top of the next page, where

$$\mathbf{U}_Y = \begin{bmatrix} \mathbf{U}_W & \mathbf{H}_I \mathbf{B}_I^{-1/2} & \mathbf{H}_e \beta_e^{-1/2} & \mathbf{H}_0 \mathbf{B}_0^{-1/2} \end{bmatrix} \quad (17)$$

$$\mathbf{B}_0 = \text{diag}(\beta_{01}^0, \dots, \beta_{0K}^0) \quad (18)$$

$$\mathbf{B}_I = \text{diag}(\beta_{11}^0, \dots, \beta_{1K}^0, \dots, \beta_{L1}^0, \dots, \beta_{LK}^0) \quad (19)$$

$$\mathbf{P}_I = \text{diag}(P_1, \dots, P_1, \dots, P_L, \dots, P_L) \quad (20)$$

and $\mathbf{U}_W \in \mathbb{C}^{N_t \times (N_t - M)}$ has orthogonal columns.

When $N_t \rightarrow \infty$, we have

$$\frac{1}{N_t} \mathbf{U}_Y^H \mathbf{U}_Y \xrightarrow{N_t \rightarrow \infty} \mathbf{I}_{N_t}. \quad (21)$$

$$\begin{aligned}
& \frac{1}{N_t T} \mathbf{Y}_0 \mathbf{Y}_0^H \xrightarrow{T \rightarrow \infty} \frac{1}{N_t T} \mathbf{H}_0 \mathbf{X}_0 \mathbf{X}_0^H \mathbf{H}_0^H + \frac{1}{N_t T} \mathbf{H}_I \mathbf{X}_I \mathbf{X}_I^H \mathbf{H}_I^H + \frac{1}{N_t T} \mathbf{H}_e^0 \mathbf{X}_e \mathbf{X}_e^H (\mathbf{H}_e^0)^H + \frac{N_0}{N_t} \mathbf{I}_{N_t} \\
& = \frac{1}{N_t} \begin{bmatrix} \mathbf{U}_W & \mathbf{H}_I \mathbf{B}_I^{-1/2} & \mathbf{H}_e \beta_e^{-1/2} & \mathbf{H}_0 \mathbf{B}_0^{-1/2} \end{bmatrix} \\
& \quad \begin{bmatrix} N_0 \mathbf{I}_{N_t-M} & \frac{\mathbf{B}_I^{1/2} \mathbf{X}_I \mathbf{X}_I^H \mathbf{B}_I^{1/2}}{T} + N_0 \mathbf{I}_{(L-1)K} & \frac{\beta_e \mathbf{X}_e \mathbf{X}_e^H}{T} + N_0 \mathbf{I}_{N_e} & \frac{\mathbf{B}_0^{1/2} \mathbf{X}_0 \mathbf{X}_0^H \mathbf{B}_0^{1/2}}{T} + N_0 \mathbf{I}_K \end{bmatrix} \begin{bmatrix} \mathbf{U}_W^H \\ \mathbf{B}_I^{-1/2} \mathbf{H}_I^H \\ \beta_e^{-1/2} \mathbf{H}_e^H \\ \mathbf{B}_0^{-1/2} \mathbf{H}_0^H \end{bmatrix} \\
& \xrightarrow{T \rightarrow \infty} \frac{1}{N_t} \mathbf{U}_Y \begin{bmatrix} N_0 \mathbf{I}_{N_t-M} & \mathbf{P}_I \mathbf{B}_I + N_0 \mathbf{I}_{(L-1)K} & (\beta_e^0 P_e + N_0) \mathbf{I}_{N_e} & P_0 \mathbf{B}_0 + N_0 \mathbf{I}_K \end{bmatrix} \mathbf{U}_Y^H \quad (16)
\end{aligned}$$

From (16)–(21), we know that for $T \rightarrow \infty$, $N_t \rightarrow \infty$, \mathbf{U}_Y is the right singular matrix of \mathbf{Y}_0 . Therefore, we obtain

$$\begin{aligned}
\mathbf{Z} &= \frac{1}{\sqrt{TN_t}} (\mathbf{V}_{eq}^0)^H \mathbf{Y}_p \xrightarrow{N_t \rightarrow \infty} \\
& \frac{1}{\sqrt{TN_t}} (\mathbf{V}_{eq}^0)^H \sqrt{P_0} \Omega_0 \mathbf{X}_0 + \frac{1}{\sqrt{TN_t}} (\mathbf{V}_{eq}^0)^H \mathbf{N}_p^0. \quad (22)
\end{aligned}$$

Define $\mathbf{z} = \text{vec}(\mathbf{Z}_{0p})$, where \mathbf{Z}_{0p} is defined in Theorem 1. From (22), we can re-express the equivalent received signal during the pilot transmission phase as follows

$$\mathbf{z} = \sqrt{P_0} \sum_{t=1}^K (\boldsymbol{\omega}_t \otimes \mathbf{I}_K) \mathbf{h}_{eq,0t} + \mathbf{n} \quad (23)$$

where

$$\mathbf{n} = \begin{bmatrix} (\mathbf{V}_{eq}^0)^H \mathbf{n}_{p1}^0 \\ \vdots \\ (\mathbf{V}_{eq}^0)^H \mathbf{n}_{p\tau}^0 \end{bmatrix} \quad (24)$$

and \mathbf{n}_{pt}^0 in (24) is the t th column of \mathbf{N}_p^0 .

Based on (23), the MMSE estimate of $\mathbf{h}_{eq,0k}$ is given by

$$\begin{aligned}
\hat{\mathbf{h}}_{eq,0k} &= \sqrt{P_0} (P_0 \tau \mathbf{I}_K + N_0 \mathbf{I}_K)^{-1} (\boldsymbol{\omega}_k \otimes \mathbf{I}_K)^H \mathbf{z} \\
&= \frac{\sqrt{P_0}}{P_0 \tau + N_0} \left(\sqrt{P_0} \tau \mathbf{h}_{eq,0k} + (\boldsymbol{\omega}_k \otimes \mathbf{I}_K)^H \mathbf{n} \right). \quad (25)
\end{aligned}$$

For the noise term in (25), we have

$$\begin{aligned}
(\boldsymbol{\omega}_k \otimes \mathbf{I}_K)^H \mathbf{n} &= (\mathbf{V}_{eq}^0)^H \sum_{t=1}^{\tau} \omega_{kt}^* \mathbf{w}_t \\
&= (\mathbf{V}_{eq}^0)^H \sum_{t=1}^{\tau} \omega_{kt}^* \mathbf{w}_t = (\mathbf{V}_{eq}^0)^H \tilde{\mathbf{n}}_{eq} \quad (26)
\end{aligned}$$

where ω_{kt} is the t th element of $\boldsymbol{\omega}_k$. Combining (25) and (26) completes the proof.

APPENDIX B PROOF OF THEOREM 2

First, based on the property of MMSE estimates, we know that $E[g_{0k,k}^0] = \sqrt{P} E[\|\hat{\mathbf{h}}_{eq,0k}\|]$.

Based on (3) and (16), we have

$$\begin{aligned}
\|\hat{\mathbf{h}}_{eq,0k}\|^2 &= \frac{P_0}{(P_0 \tau + N_0)^2} \\
&\times \left(\sqrt{P_0} \tau \mathbf{h}_{eq,0k} + \mathbf{V}_{eq}^0 \tilde{\mathbf{n}}_{eq} \right)^H \left(\sqrt{P_0} \tau \mathbf{h}_{eq,0k} + \mathbf{V}_{eq}^0 \tilde{\mathbf{n}}_{eq} \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{P_0}{(P_0 \tau + N_0)^2} \left[P_0 \tau^2 \frac{1}{N_t} (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} \right. \\
&\times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 + \frac{1}{N_t} \tilde{\mathbf{n}}_{eq}^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} \\
&\times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \tilde{\mathbf{n}}_{eq} \left. \right] \\
&= \frac{P_0}{(P_0 \tau + N_0)^2} \left[P_0 \tau^2 \frac{1}{N_t} \sum_{t=1}^K (\mathbf{h}_{0k}^0)^H (\beta_{0t}^0)^{-1} \mathbf{h}_{0t}^0 (\mathbf{h}_{0t}^0)^H \mathbf{h}_{0k}^0 \right. \\
&+ \frac{1}{N_t} \sum_{t=1}^K \tilde{\mathbf{n}}_{eq}^H \mathbf{h}_{0t}^0 (\beta_{0t}^0)^{-1} (\mathbf{h}_{0t}^0)^H \tilde{\mathbf{n}}_{eq} \left. \right]. \quad (27)
\end{aligned}$$

When $N_t \rightarrow \infty$, based on [18, Corollary 1], we have

$$\begin{aligned}
&\frac{1}{N_t} (\mathbf{h}_{0k}^0)^H \mathbf{h}_{0t}^0 (\beta_{0t}^0)^{-1} (\mathbf{h}_{0t}^0)^H \mathbf{h}_{0k}^0 \\
&\xrightarrow{N_t \rightarrow \infty} \frac{\beta_{0k}^0 (\beta_{0t}^0)^{-1}}{N_t} \text{tr} \left(\mathbf{h}_{0t}^0 (\mathbf{h}_{0t}^0)^H \right) \xrightarrow{N_t \rightarrow \infty} \beta_{0k}^0 \quad (28)
\end{aligned}$$

$$\frac{1}{N_t} (\mathbf{h}_{0k}^0)^H \mathbf{h}_{0k}^0 (\beta_{0k}^0)^{-1} (\mathbf{h}_{0k}^0)^H \mathbf{h}_{0k}^0 \xrightarrow{N_t \rightarrow \infty} \beta_{0k}^0 N_t \quad (29)$$

$$\frac{1}{N_t} \tilde{\mathbf{n}}_{eq}^H \mathbf{h}_{0t}^0 (\beta_{0t}^0)^{-1} (\mathbf{h}_{0t}^0)^H \tilde{\mathbf{n}}_{eq} \xrightarrow{N_t \rightarrow \infty} \tau N_0. \quad (30)$$

Substituting (28)–(30) into (27), we have

$$\begin{aligned}
&\|\hat{\mathbf{h}}_{eq,0k}\|^2 \xrightarrow{N_t \rightarrow \infty} \frac{P_0}{(P_0 \tau + N_0)^2} (P_0 \tau^2 \beta_{0k}^0 N_t + \\
&P_0 \tau^2 \beta_{0k}^0 (K-1) + K \tau N_0). \quad (31)
\end{aligned}$$

Next, we evaluate $\text{var}(g_{0k,k}^0)$. First, we obtain

$$\begin{aligned}
&(\mathbf{h}_{0k}^0)^H (\mathbf{V}_{eq}^0)^H \hat{\mathbf{h}}_{eq,0k} \hat{\mathbf{h}}_{eq,0k}^H \mathbf{V}_{eq}^0 \mathbf{h}_{0k}^0 = \frac{P_0}{(P_0 \tau + N_0)^2} (\mathbf{h}_{0k}^0)^H \\
&\times \frac{1}{\sqrt{N_t}} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1/2} \left(\sqrt{P_0} \tau \mathbf{h}_{eq,0k} + \mathbf{V}_{eq}^0 \tilde{\mathbf{n}}_{eq} \right) \\
&\times \left(\sqrt{P_0} \tau \mathbf{h}_{eq,0k} + \mathbf{V}_{eq}^0 \tilde{\mathbf{n}}_{eq} \right)^H \frac{1}{\sqrt{N_t}} \mathbf{B}_0^{-1/2} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 \\
&\xrightarrow{N_t \rightarrow \infty} \frac{P_0}{(P_0 \tau + N_0)^2} (\mathbf{h}_{0k}^0)^H \frac{1}{\sqrt{N_t}} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1/2} \left[\frac{1}{N_t} \right. \\
&\times P_0 \tau^2 \mathbf{B}_0^{-1/2} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H (\mathbf{h}_{0k}^0) (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \\
&\times \mathbf{B}_0^{-1/2} + \frac{1}{N_t} \mathbf{B}_0^{-1/2} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \tilde{\mathbf{n}}_{eq} \tilde{\mathbf{n}}_{eq}^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \left. \right] \quad (32)
\end{aligned}$$

$$\begin{aligned}
& \times \mathbf{B}_0^{-1/2} \left[\frac{1}{\sqrt{N_t}} \mathbf{B}_0^{-1/2} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 \right. \\
& = \frac{P_0}{(P_0\tau + N_0)^2} \left[\frac{P_0\tau^2}{N_t^2} (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} \right. \\
& \times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} \\
& \times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 + \frac{1}{N_t^2} (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} \\
& \times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \tilde{\mathbf{n}}_{eq} \tilde{\mathbf{n}}_{eq}^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} \\
& \times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 \left. \right]. \quad (33)
\end{aligned}$$

From (28) and (29), we have

$$\begin{aligned}
& \frac{1}{N_t} (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H (\mathbf{h}_{0k}^0) \\
& \xrightarrow{N_t \rightarrow \infty} \beta_{0k}^0 N_t + \beta_{0k}^0 (K-1) \\
& \frac{1}{N_t^2} (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \tilde{\mathbf{n}}_{eq} \\
& \times \tilde{\mathbf{n}}_{eq}^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 \quad (34)
\end{aligned}$$

Also, we have

$$\begin{aligned}
& = \frac{1}{N_t^2} \tilde{\mathbf{n}}_{eq}^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 (\mathbf{h}_{0k}^0)^H \\
& \times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \tilde{\mathbf{n}}_{eq} \\
& \xrightarrow{N_t \rightarrow \infty} \frac{1}{N_t^2} \tau N_0 \text{tr} \left([\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 \right. \\
& \times (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \left. \right) \\
& = \frac{\tau N_0}{N_t^2} (\mathbf{h}_{0k}^0)^H [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \\
& \times [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0] \mathbf{B}_0^{-1} [\mathbf{h}_{01}^0, \dots, \mathbf{h}_{0K}^0]^H \mathbf{h}_{0k}^0 \\
& = \frac{\tau N_0}{N_t^2} (\mathbf{h}_{0k}^0)^H \sum_{t=1}^K (\beta_{0t}^0)^{-1} \mathbf{h}_{0t}^0 (\mathbf{h}_{0t}^0)^H \\
& \times \sum_{p=1}^K (\beta_{0p}^0)^{-1} \mathbf{h}_{0p}^0 (\mathbf{h}_{0p}^0)^H \mathbf{h}_{0k}^0 \\
& \frac{1}{N_t^2} (\mathbf{h}_{0k}^0)^H (\beta_{0t}^0)^{-1} \mathbf{h}_{0t}^0 (\mathbf{h}_{0t}^0)^H (\beta_{0p}^0)^{-1} (\mathbf{h}_{0p}^0)^H \mathbf{h}_{0p}^0 \mathbf{h}_{0k}^0 \quad (35)
\end{aligned}$$

When $k \neq t \neq p$, we have

$$\begin{aligned}
& \xrightarrow{N_t \rightarrow \infty} \frac{\beta_{0k}^0}{N_t^2} \text{tr} \left((\beta_{0t}^0)^{-1} \mathbf{h}_{0t}^0 (\mathbf{h}_{0t}^0)^H (\beta_{0p}^0)^{-1} (\mathbf{h}_{0p}^0)^H \mathbf{h}_{0p}^0 \right) \\
& \xrightarrow{N_t \rightarrow \infty} \frac{\beta_{0k}^0}{N_t} \quad (36)
\end{aligned}$$

When $k = t = p$, we have

$$\begin{aligned}
& \frac{1}{N_t^2} (\mathbf{h}_{0k}^0)^H (\beta_{0k}^0)^{-1} \mathbf{h}_{0k}^0 (\mathbf{h}_{0k}^0)^H (\beta_{0k}^0)^{-1} \mathbf{h}_{0k}^0 (\mathbf{h}_{0k}^0)^H \mathbf{h}_{0k}^0 \\
& \xrightarrow{N_t \rightarrow \infty} N_t \beta_{0k}^0 \quad (37)
\end{aligned}$$

When $k = t \neq p$, $k = p \neq t$, $k \neq t = p$, $k \neq p = t$, we have

$$\begin{aligned}
& \frac{1}{N_t^2} (\mathbf{h}_{0k}^0)^H (\beta_{0t}^0)^{-1} \mathbf{h}_{0t}^0 (\mathbf{h}_{0t}^0)^H (\beta_{0t}^0)^{-1} (\mathbf{h}_{0t}^0)^H \mathbf{h}_{0t}^0 \mathbf{h}_{0k}^0 \\
& \xrightarrow{N_t \rightarrow \infty} \beta_{0k}^0 \quad (38)
\end{aligned}$$

Combining (31), (33)–(38), and the definition of $g_{0k,k}^0$, we have

$$\text{var}(g_{0k,k}^0) = a_2 - a_1. \quad (39)$$

where a_2 and a_1 are defined in (13) and (14).

For $E \left[|g_{0t,k}^0|^2 \right]$ and $E \left[|g_{lt,k}^0|^2 \right]$, we have

$$\begin{aligned}
E \left[|g_{0t,k}^0|^2 \right] & = PE \left[\frac{\hat{\mathbf{h}}_{eq,0t} \hat{\mathbf{h}}_{eq,0t}^H \text{tr} \left(\hat{\mathbf{h}}_{eq,0k} \hat{\mathbf{h}}_{eq,0k}^H \right)}{\left\| \hat{\mathbf{h}}_{eq,0t} \right\|^2} \right] \\
& = P \beta_{0k}^0 \quad (40)
\end{aligned}$$

$$\begin{aligned}
E \left[|g_{lt,k}^0|^2 \right] & = PE \left[\frac{\hat{\mathbf{h}}_{eq,lt} \hat{\mathbf{h}}_{eq,lt}^H \text{tr} \left(\hat{\mathbf{h}}_{eq,lk} \hat{\mathbf{h}}_{eq,lk}^H \right)}{\left\| \hat{\mathbf{h}}_{eq,lt} \right\|^2} \right] \\
& = P \beta_{lk}^0. \quad (41)
\end{aligned}$$

For $C_{k,\text{upper}}^{\text{reve}}$ in (9), we know from (16) that when $N_t \rightarrow \infty$, $(\mathbf{V}_{eq}^0)(\mathbf{H}_e^0)^H \rightarrow 0$. Therefore, we have

$$C_{k,\text{upper}}^{\text{reve}} \xrightarrow{N_t \rightarrow \infty} 0. \quad (42)$$

Substituting (31), (39), (40), (41), and (42) into (5) completes the proof.

REFERENCES

- [1] A. D. Wyner, "The wiretap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.
- [2] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, pp. 4961–4972, Aug. 2011.
- [3] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Trans. Veh. Technol.*, vol. 61, pp. 2599–2612, Jul. 2012.
- [4] Y. Wu, J.-B. Wang, J. Wang, R. Schober, and C. Xiao, "Secure transmission with large numbers of antennas and finite alphabet inputs," *IEEE Trans. Commun.*, vol. 65, pp. 3614–3628, Aug. 2017.
- [5] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, pp. 1065–1082, Jun. 2014.
- [6] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 4766–4781, Sep. 2014.
- [7] X. Chen, L. Lei, H. Zhang, and C. Yuen, "Large-scale MIMO relaying techniques for physical layer security: AF or DF?" *IEEE Trans. Wireless Commun.*, vol. 14, pp. 5135–5146, Sep. 2015.
- [8] J. Chen, X. Chen, W. H. Gerstacker, and D. W. K. Ng, "Resource allocation for a massive MIMO relay aided secure communication," *IEEE Trans. Inf. Forensics Security*, vol. 11, pp. 1700–1711, Aug. 2016.
- [9] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Analysis and design of secure massive MIMO systems in the presence of hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 2001–2016, Mar. 2017.
- [10] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Trans. Inf. Theory*, vol. 62, pp. 3880–3900, Jul. 2016.
- [11] S. Im, H. Jeon, J. Choi, and J. Ha, "Secret key agreement with large antenna arrays under the pilot contamination attack," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 6579–6594, Dec. 2015.
- [12] Y. O. Basciftci, C. E. Koksul, and A. Ashikhmin, "Securing massive MIMO at the physical layer," [Online]. Available: <http://arxiv.org/abs/1505.00396>.
- [13] C.-K. Wen, Y. Wu, K.-K. Wong, R. Schober, and P. Ting, "Performance limits of massive MIMO systems based on Bayes-optimal inference," in *Proc. Int. Conf. Commun. (ICC'2015)*, London, UK, Jun. 2015, pp. 1783–1788.
- [14] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topic Signal Process.*, vol. 8, pp. 773–786, Oct. 2014.
- [15] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge University Press, 2011.
- [16] G. Geraci, M. Egan, J. Yuan, A. Razi, and I. Collings, "Secrecy sum-rates for multi-user MIMO regularized channel inversion precoding," *IEEE Trans. Commun.*, vol. 60, pp. 3472–3482, Nov. 2012.
- [17] 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, pp. 2640–2651, Aug. 2011.
- [18] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2059–2078, Sep. 2000.