Secure performance analysis and pilot spoofing attack detection in cell-free massive MIMO systems with finite-resolution ADCs

Xianyu Zhang, Xiaoqiang Qiao, Tao Liang and Kang An

Abstract
In this article, the secure communication in cell-free massive multiple-input multiple-output system with low-resolution analog-to-digital converters is investigated in the presence of an active eavesdropper. Specifically, in this article, the deterioration caused by the analog-to-digital converter imperfections on the accuracy of the channel estimation and secure transmission performance is studied. Besides, the additive quantization noise model is utilized to analyze the impacts of the low-resolution analog-to-digital converters. The minimum mean square error channel estimation results show that there is a nonzero floor caused by the coarse analog-to-digital converters. Then, the closed-form expressions for both the legitimate users achievable ergodic rate and the information leakage to the eavesdropper are derived, respectively. Moreover, tight approximated ergodic secrecy rate expression is also presented with respect to analog-to-digital converters quantization bits, number of antennas, pilot power, and so on. To degrade the impacts of the pilot spoofing attack, an active attack detection approach based on random matrix theory is proposed which can only be operated at one access point. Simulation results are provided to corroborate the obtained results and analyze the impacts of various parameters on system secrecy performance. Also, the superiority of the proposed active attacks detection method is confirmed via simulation results.

Keywords
Cell-free massive multiple-input multiple-output, physical layer security, low-resolution analog-to-digital converters, channel estimation, spoofing attack, active attack detection

Date received: 22 April 2021; accepted: 15 November 2021

Handling Editor: Yanjiao Chen

Introduction
Massive multiple-input multiple-output (MIMO) is regarded as a promising wireless access technology for 5G and B5G wireless communication which can provide high spectral efficiency and high energy efficiency with simple signal processing.\cite{1,2} Cellular massive MIMO equips the large number of the antennas on a single base station (BS) which has been analyzed widely and deeply on various respects, such as performance analysis, beamforming, power control, secure communication, and so on.\cite{3,4} However, the performance of such cellular system is inevitably limited by the intercell interferences and channel correlation. One practical alternative is to install several antenna sets onto distributed physical locations which is also called distributed massive MIMO.\cite{5,6} As a scalable distributed massive MIMO concept, cell-free massive MIMO has been...
proved to be a promising wireless network architecture which can provide coherent service by geographically distributed access points (APs) for future wireless communication systems.\(^7\) Since its proposal, various works on performance evaluations and transmission schemes over cell-free massive MIMO have been investigated.\(^8\)–\(^10\) Due to the broadcast nature of wireless channels, facilitating secure and reliable wireless transmission is a critical topic for future wireless communication systems.\(^11\)–\(^13\) Despite the robustness of massive MIMO against passive eavesdropping, the active eavesdropping attacks are more relevant.\(^14\),\(^15\)

Currently, most related literatures merely focused on the secure transmission in co-located massive MIMO systems,\(^16\)–\(^18\) while secrecy provisions for cell-free massive MIMO have not been well explored. Hoang et al.\(^19\) discussed the secrecy performance and power allocation for cell-free massive MIMO with active eavesdropper. The effect of hardware impairments over secure cell-free massive MIMO system with active spoofing attack has been investigated in Zhang et al.\(^20\) However, all aforementioned works are based on adopting high-resolution analog-to-digital converters (ADCs) in massive MIMO systems. To reduce energy consumption and physical hardware costs, implementing finite-resolution ADCs in massive MIMO systems has been regarded as an effective approach. Several researchers have devoted to the implementation of finite-resolution ADCs in cell-free massive MIMO systems recently.\(^21\)–\(^23\)

Besides, the secure co-located massive MIMO with finite-resolution ADCs has been studied in Teeti.\(^24\) Nevertheless, the impact of ADCs quality on secrecy performance in cell-free massive MIMO has not been considered in existing research.

In a time-division duplex (TDD) massive MIMO system, the BS always acquires the channel state information (CSI) of the legitimate users via uplink training based on the reciprocity of the downlink and uplink channels. All legal users are required to send the redesigned and assigned pilot (training) signals to the BS in uplink training phase. Since all pilot signals are repeatedly used and publicly known by the BS and all users, a smart malicious eavesdropper can easily learn the target user’s pilot signals and transmit an identical pilot signal as that of a legitimate target user which is also called spoofing attack.\(^25\) Due to the spoofing attack operated by the active eavesdropper, the adversary could manipulate the CSI estimation which is the weighted sum of the target legitimate user and the eavesdropper’s channels. Consequently, the beamforming based on the estimated CSI during the data transmission results in information leakage to the eavesdropper. Some existing literatures show that the massive MIMO is robust against passive eavesdropping while the secrecy performance of the massive MIMO system is dramatically degraded by active attacks.\(^14\),\(^15\),\(^26\) Thus, detection of the active eavesdropper can be considered as an effective countermeasure to weaken the influence of active spoofing attacks. Kapetanović et al.\(^27\) proposed a detection scheme via random training pilots for massive MIMO system. However, some special pilot sequences are required to do the detection in this proposed method. Then, different schemes for active attack detection have been presented in Kapetanović et al.\(^28\) which are operated at different locations and based on different system parameters. It is noted that the precoder normalization is crucially important for these active eavesdropper detection schemes. Moreover, a simple detection protocol based on the statistical properties of the channel estimation has been provided in Al-Nahari,\(^15\) which is also determined by the correctness of the channel estimation. Besides, the minimum description length (MDL) source enumeration algorithm has been utilized for active eavesdropper detection in Tugnait,\(^29\),\(^30\) which can greatly improve the detection performance using short pilot sequences. According to Kritchman and Nadler\(^31\) and Tugnait,\(^32\) the random matrix theory (RMT) detection approach can achieve more precise results compared with MDL algorithm and has been proved to be highly effective in traditional MIMO system.

Motivated by the aforementioned observations, this article presents the first study of the security problem in cell-free massive MIMO system with finite-resolution ADCs against active eavesdropper. Furthermore, the tight asymptotic lower bound of achievable secrecy rate has been provided to enable the effects of ADC resolution, number of APs, each AP’s antenna number, and so on. Also, an active eavesdropping detection algorithm based on RMT is presented. And the proposed detection algorithm can be operated at one AP which only needs a simple procedure. The specific contributions of this article can be summarized as follows:

- To capture the aggregate impacts of low-resolution ADCs, the additive quantization noise model (AQNM) is utilized to analyze the performance of the secure cell-free massive MIMO system. Specifically, this article operates the channel estimation via uplink training under minimum mean square error (MMSE) criterion. Moreover, the results show that the low-quality ADCs lead to a nonzero floor on the channel estimation. Also, more pilot power of the eavesdropper would bring larger damnification for channel estimation.

- The downlink secrecy communication has been considered with the low-resolution ADCs. Based on the AQNM and the imperfect CSI, the closed-form expression of the secrecy rate has been derived with respect to key system
parameters, including the quantization bits of the ADCs, the number of the total antennas, the pilot and data transmission power, and so on. These novel results can provide an effective tool to quantitatively analyze the impacts of system parameters on system secrecy performance.

- To weaken the impacts taken by the active eavesdropping, an active attacks detection scheme based on the pilot self-contamination has been proposed which can only be operated at one AP. The detection is employed via determining the signal subspace rank which is estimated by RMT source enumeration approach. Also, the proposed detection approach has been illustrated accurate and effective via simulations.

The outline of the article is organized as follows. Section “System model” introduces the considered secure cell-free massive MIMO system model. Section “Uplink training” details the procedure of the uplink MMSE channel estimation. Section “Secrecy performance analysis” derives the achievable secrecy rate of the considered system. The RMT-based active eavesdropping detection is discussed in section “RMT-based pilot spoofing attack detection.” Simulation results are presented in section “Numerical results.” Finally, the article is concluded in section “Conclusion.”

**Notation**

Throughout this article, boldface lower case and upper case letters are utilized to represent vectors and matrices, for example, $\mathbf{y}$, $\mathbf{Y}$. The superscript $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ stand for the conjugate, transpose and Hermitian transpose operation, respectively. In addition, the Euclidean norm of a vector $\mathbf{g}$ and the absolute value of a scalar $g$ can be denoted as $||\mathbf{g}||$ and $|g|$, respectively. $E\{\cdot\}$, $\text{Var}\{\cdot\}$, $\text{rank}\{\cdot\}$, and $\text{Eig}\{\cdot\}$ represent operations of expectation, variance, rank, and eigenvalue of a matrix, respectively. $\mathbf{x} \sim \mathcal{CN}(\mathbf{m}, \mathbf{A})$ stands for a complex Gaussian random vector $\mathbf{x}$ with mean $\mathbf{m}$ and covariance matrix $\mathbf{A}$. $\mathbf{I}_M$ and $\mathbb{C}^{M \times N}$ are $M$-dimensional identity matrix and $M \times N$ dimensional complex space. Besides, independent and identically distributed is abbreviated as i.i.d. Finally, $[x]^+ = \max\{0,x\}$.

**System model**

As illustrated in Figure 1, the cell-free massive MIMO system comprising $M$ APs, $K$ one-antenna legitimate users, and an active eavesdropper (Eve) is considered in this article. Each AP is equipped with $N$ antennas. All APs are connected to a CPU through backhaul for sharing information. The Eve is equipped with single antenna and intends to contaminate the uplink training by pilot spoofing attack. The system is assumed operated in TDD mode, where the downlink and uplink channels are assumed to be reciprocal. Rayleigh block-fading channels are assumed between the APs and different terminals, where the channels remain constant in one time slot and vary independently afterward. Then, the channels between the $m$th AP and $k$th user (the Eve) can be denoted by $7$–$9$

$$
\mathbf{g}_{mk} = \sqrt{\beta_{mk}}\mathbf{h}_{mk} \in \mathbb{C}^{N \times 1} 
$$

$$
\mathbf{g}_{mk} = \sqrt{\beta_{mk}}\mathbf{h}_{mk} \in \mathbb{C}^{N \times 1} 
$$

where $\beta_{mk}(h_{mk})$ represents the large-scale fading, $h_{mk}(h_{mk})$ models the small-scale fading vector with elements being independent and identically distributed (i.i.d) $\mathcal{CN}(0, 1)$ random variables (RVs).

Since secure downlink transmission is investigated in this article, each time slot consists of uplink training and downlink transmission phases.

**Uplink training**

Due to the channel reciprocity, the CSI can be only acquired by uplink training. During this phase, all legitimate users send mutually orthogonal pilot sequences of length $\tau_p$ symbols for channel estimation at the APs, where $\tau_p = K$. Let $\sqrt{\tau_p} \mathbf{\phi} \in \mathbb{C}^{\tau \times 1}$ be the pilot vector sending by the $k$th user, where $\|\mathbf{\phi}\|^2 = 1$, $p_k$ represents the normalized pilot signal power. Since the pilot signals are repeatedly utilized and publicly predesigned, an intelligent eavesdropper can pretend to be a legitimate user by transmitting the identical pilot signals as the target user. No loss of generality, assuming that the Eve aims to eavesdrop the confidential information sending to the $k$th user, Eve’s pilot sequence can be denoted by $\sqrt{\tau_p} \mathbf{\varphi}_E$, where $\mathbf{\varphi}_E = \mathbf{\varphi}_k$, $p_E$ is the
normalized transmit power of Eve. The $N \times \tau_p$ received pilot matrix at the $m$th AP is given by

$$Y_{p,m} = \sqrt{\tau_p P_p} \sum_{k=1}^{K} g_{mk} \varphi_k^H + \sqrt{\tau_p P_E} g_{me} \varphi_e^H + \mathbf{W}_{p,m}$$  \hfill (3)$$

where $\mathbf{W}_{p,m} \in \mathbb{C}^{N \times \tau_p}$ denotes the additive white Gaussian noise matrix, whose elements are i.i.d $\mathcal{C}\mathcal{N}(0,1)$ RVs. By utilizing the additive quantization noise model (AQNM), the resulting signal after finite-resolution ADCs at $m$th AP can be given by 21,22

$$\hat{Y}_{p,m} = \gamma_m Y_{p,m} + \mathbf{W}_{p,m} \hfill (4)$$

where $\gamma_m \leq 1$ is scaling factor with respect to the ADC quantization bits $b_m$, whose accurate values are shown in Table 1.

Also, $\mathbf{W}_{p,m}$ is the quantization noise matrix with covariance matrix as

$$\mathbf{R}(\mathbf{W}_{p,m}) = \gamma_m (1 - \gamma_m) \text{diag} \left( \mathbb{E} \left\{ Y_{p,m} Y_{p,m}^H \right\} \right)$$ \hfill (5)

Then, projecting $\hat{Y}_{p,m}$ onto $\varphi_i$, the post-processing signal can be rewritten as

$$\hat{Y}_{p,m,i} = \gamma_m \left( \sqrt{\tau_p P_p} g_{mi} + \sqrt{\tau_p P_E} g_{me} \varphi_e^H \varphi_i + \mathbf{W}_{p,m} \right) + \hat{\mathbf{w}}_{p,m} \hfill (6)$$

where $\mathbf{W}_{p,m} = \mathbf{W}_{p,m} \varphi_i$, $\hat{\mathbf{w}}_{p,m} = \hat{\mathbf{W}}_{p,m} \varphi_i$.

By utilizing the MMSE estimation, the estimated channels can be written by 16,21

$$\mathbf{g}_m = \hat{\mathbf{g}}_m + \mathbf{g}_{me} \hfill (7)$$

$$\mathbf{g}_{me} = \hat{\mathbf{g}}_{me} + \mathbf{g}_{me} \hfill (8)$$

where the estimate $\hat{\mathbf{g}}_m$ (error $\mathbf{g}_{me}$) and estimation error $\mathbf{g}_{me}$ (true $\mathbf{g}_{me}$) are mutually independent, whose elements are zero mean i.i.d. RVs with variances $\lambda_m$ ($\lambda_{me}$) and $\beta_{ml} - \lambda_ml$ ($\beta_{me} - \lambda_ml$).

Moreover, the channel estimation $\hat{\mathbf{g}}_m$ can be expressed as

$$\hat{\mathbf{g}}_m^H = \frac{\mathbb{E}\left\{ \varphi_i^H \hat{Y}_{p,m,i} \right\}}{\mathbb{E}\left\{ \mathbf{Y}_{p,m,i} \hat{Y}_{p,m,i}^H \right\}} \varphi_i \hfill (9)$$

Then, one important theorem can be given as follows.

**Theorem 1.** The variances $\lambda_{ml}$ and $\lambda_{me}$ can be denoted by

$$\lambda_{ml} = \begin{cases} \frac{\gamma_m \tau_p P_p \beta_{ml}^2}{1 + \tau_p P_p \beta_{ml} + \tau_p P_E \beta_{me}}, & \text{if } l = k \\ \frac{\gamma_m \tau_p P_p \beta_{ml}^2}{1 + \tau_p P_p \beta_{ml}}, & \text{if } l \neq k \end{cases} \hfill (10)$$

$$\lambda_{me} = \frac{\gamma_m \tau_p P_E \beta_{me}^2}{1 + \tau_p P_p \beta_{ml} + \tau_p P_E \beta_{me}} \hfill (11)$$

**Proof.** Please see Appendix 1.

Since the downlink beamformer is based on the estimated CSI via uplink training, the accuracy of the channel estimation is crucially important. Then, the normalized minimum square error (NMSE) of the target user can be used to evaluate the estimation performance which is defined as

$$\text{NMSE} = \frac{\sum_{m=1}^{M} \mathbb{E}\left\{ \| \mathbf{g}_{mk} \|^2 \right\}}{\sum_{m=1}^{M} \sum_{l=1}^{L} (\beta_{mk} - \lambda_{mk})} \hfill (12)$$

To achieve some valuable insights, one limit behavior of the NMSE can be considered as follows. As $P_d \rightarrow +\infty$, we can get the limit value of $\lambda_{mk}$ as

$$\lim_{P_d \rightarrow +\infty} \lambda_{mk} = \lim_{P_d \rightarrow +\infty} \frac{\gamma_m \tau_p P_p \beta_{ml}^2}{1 + \tau_p P_p \beta_{ml} + \tau_p P_E \beta_{me}} \hfill (13)$$

Thus, there is a nonzero floor on the channel estimation error NMSE caused by the ADCs imperfection. Besides, this zero floor cannot be compensated by increasing the pilot power of the legitimate users.

**Secrecy performance analysis**

In this section, we will focus on the performance analysis as a measure to evaluate the performance of the considered secure communication system. 33–37 All APs perform conjugate beamforming to simultaneously transmit data to all users with the same time-frequency resource. Then, the signal transmitted by the $m$th AP can be represented by

$$\mathbf{x}_m = \sqrt{P_d} \sum_{l=1}^{K} \sqrt{\eta_{ml}} \mathbf{g}_{ml}^* q_l \hfill (14)$$
where \( q_l \) denotes the normalized signal intended to the \( l \)th user with \( |q_l| = 1 \), \( \eta_{ml} \) is the power control coefficient satisfying the power constraint as
\[
\sum_{l=1}^{K} \eta_{ml} \lambda_{ml} \leq P_d
\]  
(15)

Then, the received signal at the \( k \)th user and the Eve can be denoted by
\[
y_k = \sum_{m=1}^{M} g_{mk}^T x_m + w_k \tag{16}
\]
\[
y_E = \sum_{m=1}^{M} g_{mk}^T x_m + w_E \tag{17}
\]
where \( w_k \sim \mathcal{CN}(0, 1) \) and \( w_E \sim \mathcal{CN}(0, 1) \) represent received noise elements at the target user and the Eve, respectively. Operated by the low-resolution ADCs at the users, the output signal can be given by
\[
y_k^\prime = \gamma_u v_k + \tilde{w}_k \tag{18}
\]
where \( \gamma_u \) denotes the scaling factor of the ADCs at the users. Also, \( \tilde{w}_k \) is the quantization noise with the variance as
\[
R_{\tilde{w}_k} = \gamma_u(1-\gamma_u)\mathbb{E}\{|y_k|^2\} \tag{19}
\]

We consider the realistic case that the target user cannot obtain its instantaneous CSI. Hence, the \( k \)th user will detect its desired signal \( q_k \) by using the statistical CSI. Considering a pessimistic case, we assume that the Eve can obtain the perfect instantaneous CSI and has been equipped with high-quality hardware. Based on these settings, the lower bound on the target user’s achievable rate and the upper bound for the Eve’s ergodic capacity can be achieved. Furthermore, the signals at the \( k \)th user and the Eve can be rewritten as
\[
\begin{align*}
\hat{y}_k &= \gamma_u \sqrt{p_d} \mathbb{E}\left\{ \sum_{m=1}^{M} \sqrt{\eta_{mk}} \mathbf{g}_{mk}^T \mathbf{g}_{mk}^* q_k \right\} + w_k^{\text{eff}} \\
y_E &= \sqrt{p_d} \sum_{m=1}^{M} \sqrt{\eta_{mk}} \mathbf{g}_{mk}^T \mathbf{g}_{mk}^* q_k + w_E^{\text{eff}}
\end{align*}
\]  
(20)

where \( w_k^{\text{eff}} \) and \( w_E^{\text{eff}} \) denote the effective noise at the \( k \)th user and the Eve which are given by
\[
\begin{align*}
w_k^{\text{eff}} &= \gamma_u \sqrt{p_d} \sum_{m=1}^{M} \sqrt{\eta_{mk}} \mathbf{g}_{mk}^T \mathbf{g}_{mk}^* q_k \\
- \gamma_u \sqrt{p_d} \mathbb{E}\left\{ \sum_{m=1}^{M} \sqrt{\eta_{mk}} \mathbf{g}_{mk}^T \mathbf{g}_{mk}^* \right\} q_k \\
+ \gamma_u \sqrt{p_d} \sum_{l \neq k}^{K} \sum_{m=1}^{M} \sqrt{\eta_{ml}} \mathbf{g}_{ml}^T \mathbf{g}_{ml}^* q_l + \gamma_u w_k + \tilde{w}_k
\end{align*}
\]  
(21)

\[
w_E^{\text{eff}} = \sqrt{p_d} \sum_{l \neq k}^{K} \sum_{m=1}^{M} \sqrt{\eta_{ml}} \mathbf{g}_{ml}^T \mathbf{g}_{ml}^* q_l + w_E \tag{22}
\]

In this article, the ergodic secrecy rate is chosen to evaluate the secrecy performance of the considered system. By considering the difference of the channel capacities between the target user and the Eve, the ergodic secrecy rate is given by
\[
R_k^{\text{sec}} = [R_k - R_k^{\text{Eve}}]^+ = [\log_2(1 + \xi_k) - \log_2(1 + \xi_k^{\text{Eve}})]^+
\]  
(23)

where \( R_k \) (\( R_k^{\text{Eve}} \)) and \( \xi_k \) (\( \xi_k^{\text{Eve}} \)) denote the achievable ergodic rate and signal-to-interference-noise ratio (SINR) at the \( k \)th user (the Eve). Then, we can get that
\[
\begin{align*}
\xi_k &= \frac{\gamma_u^2 p_d \mathbb{E}\left\{ \sum_{m=1}^{M} \sqrt{\eta_{mk}} \mathbf{g}_{mk}^T \mathbf{g}_{mk}^* \right\}^2}{\mathbb{E}\left\{ \left| w_k^{\text{eff}} \right|^2 \right\}} \\
\xi_k^{\text{Eve}} &= \frac{p_d \mathbb{E}\left\{ \sum_{m=1}^{M} \sqrt{\eta_{mk}} \mathbf{g}_{mk}^T \mathbf{g}_{mk}^* \right\}^2}{\mathbb{E}\left\{ \left| w_E^{\text{eff}} \right|^2 \right\}}
\end{align*}
\]  
(24)

(25)

Subsequently, the closed-form expressions of \( R_k \) and \( R_k^{\text{Eve}} \) can be given in the following theorem.

**Theorem 2.** In the considered cell-free massive MIMO system, the SINRs at the \( k \)th user and the Eve can be given by equations (26) and (27) shown at the top of the next page

\[
\begin{align*}
\xi_k &= \frac{\gamma_u \left( \sum_{m=1}^{M} \sqrt{\eta_{mk}} \lambda_{mk} \right)^2}{\frac{1}{N_p} + \frac{1}{N_p} \sum_{l=1}^{K} \sum_{m=1}^{M} \sqrt{\eta_{mk}} \lambda_{mk} + (1-\gamma_u) \left( \sum_{m=1}^{M} \sqrt{\eta_{mk}} \lambda_{mk} \right)^2} \\
\xi_k^{\text{Eve}} &= \frac{p_d \sum_{m=1}^{M} \eta_{mk} \lambda_{mk}^2 + N \beta_{me} \lambda_{me}}{1 + N p_d \sum_{l \neq k}^{K} \sum_{m=1}^{M} \eta_{ml} \lambda_{ml} \beta_{me}}
\end{align*}
\]  
(26)

(27)

**Proof.** Please refer to Appendix 2.

**RMT-based pilot spoofing attack detection**

According to some existing papers, it is reasonable to operate the pilot spoofing attack detection based on the uplink training. Our proposed detection scheme is to perform source enumeration using RMT algorithm
which is based on both the distribution of the signal eigenvalues and noise eigenvalues with the presence of noise. Also, to improve the detection accuracy, the uplink training can be divided into two phases: uplink training and uplink active attack detection. In the detection, it is set that only the legitimate target user send the pilot signal which can be designed that a fraction $0 \leq \alpha \leq 1$ of the pilot power is allocated to a scalar random sequence $\phi_B$. The superimposing random sequence $\phi_B$ is normalized $(\|\phi_B\| = 1)$ and zero-mean whose elements can be chosen from finite alphabet: BPSK or QPSK. Therefore, instead of the pilot signal $\sqrt{P_p}\phi_1$, the pilot signal used for detection at the target user (the $k$th user) is given by

$$s_d = \sqrt{(1 - \alpha)P_p}\phi_1 + \sqrt{\alpha P_p}\phi_B$$  \hspace{1cm} (28)

To keep the detection processing confidential, the random sequence $\phi_B$ is unknown to the APs and to the Eve in advance as it is designed randomly at the user terminal. The detection problem is in fact a binary statistical hypothesis testing problem where, the two hypotheses can be made with the null hypothesis $H_0$ denoting no attacks and the alternative hypothesis $H_1$ representing existing active eavesdropping. The two hypotheses can be presented as follows

$$H_0: Y_{d,m} = \gamma_m g_{mk} s_d^H + V_m$$

$$H_1: Y_{d,m} = \gamma_m g_{mk} s_d^H + \gamma_m g_{mk} \phi_B^H + V_m$$  \hspace{1cm} (29)

where $V_m \in \mathbb{C}^{N \times r}$ is the effective noise component including the quantization noise matrix and additive noise matrix.

Obviously, the received signals under these two hypotheses can be modeled as Gaussian variable with different expectation. According to the existing literature, the deterministic signals can be used for source enumeration to detect the existing active attacks.

Then, we can define that the correlation matrix of the received signal as

$$R_{v,i} = \tau_p^{-1} \mathbb{E}\{Y_{d,m} Y_{d,m}^H | H_i\}$$  \hspace{1cm} (30)

Furthermore, the correlation matrix of the source signals is given by

$$R_{s,i} = \tau_p^{-1} \mathbb{E}\{(Y_{d,m} - V_m)(Y_{d,m} - V_m)^H | H_i\}$$  \hspace{1cm} (31)

As such, the matrix $R_{s,i}$ under the two hypotheses can be represented as

$$R_{s,0} = \tau_p^{-1}\gamma_m g_{mk} \mathbb{E}\{s_d^H s_d\} g_{mk}^H$$

$$R_{s,1} = \tau_p^{-1}\gamma_m g \cdot \mathbb{E}\{s_d^H s_d\} \cdot g^H$$  \hspace{1cm} (32)

where the matrices $G$ and $S_d$ are defined as

$$G = \sqrt{1 - \alpha g_{mk}} + \sqrt{\alpha g_{mk}} \in \mathbb{C}^{N \times 2}$$

$$S_d = \sqrt{\tau_p P_p} \phi_1 + \sqrt{\tau_p P_p} \phi_B \in \mathbb{C}^{r \times 2}$$  \hspace{1cm} (35)

It is noted that $\text{rank}(R_{v,0}) = 1$. Due to the fact that $g_{mk} \not\equiv g_{ml}$, we can derive that $\text{rank}(R_{v,1}) = 2$ with $\alpha > 0$. Moreover, we can derive that

$$R_{v,i} = R_{v,0} + \delta_i^2 I_{r_p}$$  \hspace{1cm} (36)

where $\delta_i^2$ is the variance of the noise matrix $V_m$. Let $s_{1,0} > 0$ and $s_{1,1} > s_{2,1} > 0$ denote the nonzero eigenvalues of correlation matrix $R_{v,0}$ and $R_{v,1}$. Hence, the eigenvalues of $R_{v,i}$ can be given by

$$\text{Eig}\{R_{v,0}\} = (s_{1,0} + \delta_1^2, \delta_2^2, \ldots, \delta_2^2)$$  \hspace{1cm} (37)

$$\text{Eig}\{R_{v,1}\} = (s_{1,1} + \delta_1^2, s_{2,1} + \delta_2^2, \delta_2^2, \ldots, \delta_2^2)$$  \hspace{1cm} (38)

Thus, the two hypotheses can be distinguished by estimating the dimension $d = 1$ or $d = 2$ of the signal subspace which can be operated by the RMT estimator. Thus, the two hypotheses in equation (29) can be reformulated as

$$H_0: \text{rank}(R_{v,0}) = d = 1$$

$$H_1: \text{rank}(R_{v,1}) = d > 1$$  \hspace{1cm} (39)

The RMT approach considers distribution of the noise eigenvalues of $R_{v,i}$ with the absence of signals, that is, $Y_{d,m} = V_m$. In particular, the distribution of the largest eigenvalue of a pure noise matrix based on a certain asymptotic limit of the distribution is used for matrix rank judgment. Then, we define the correction matrix as

$$R_v = \tau_p^{-1} \mathbb{E}\{V_m V_m^H\}$$  \hspace{1cm} (40)

The eigenvalues of $R_v$ can be ordered as $l_1 \geq l_2 \geq \ldots \geq l_N$. As the joint limit $N_\tau \rightarrow + \infty$ with $N_\tau \rightarrow c \geq 0$, the largest eigenvalue $l_1$ converges to a Tracy-Widom distribution of order 2, that is

$$\lim_{N_\tau \rightarrow + \infty, N_\tau \rightarrow c \geq 0} \mathbb{P}\{l_1 < \delta_i^2, \mu_{\tau_p N} + 2\delta_{\tau_p N}\} = F_{TW,2}(z)$$  \hspace{1cm} (41)

where $F_{TW,2}(z)$ represents the Tracy-Widom probability distribution function of order 2. Besides, $\mu_{\tau_p N}$ and $\delta_{\tau_p N}$ denote the centering and scaling parameters which are only determined by $N$ and $\tau_p$. Furthermore, the parameters $\mu_{\tau_p N}$ and $\delta_{\tau_p N}$ can be denoted by

$$\mu_{\tau_p N} = \frac{1}{\tau_p} \left(\sqrt{\tau_p + 0.5} + \sqrt{N + 0.5}\right)^2$$  \hspace{1cm} (42)
$$\delta_{\tau_p,N} = \sqrt{\frac{\mu_{\tau_p,N}}{\tau_p} \left( \frac{1}{\sqrt{\tau_p + 0.5}} + \frac{1}{\sqrt{N + 0.5}} \right)} \quad (43)$$

Then, we should consider the behavior of correlation matrix adding the signals. In this case, the eigenvalue of matrix $R_{y_i}$ also follows a Tracy-Widom distribution. Hence, the signal dimension can be estimated by judging whether the largest eigenvalue is over some deterministic threshold. Motivated by the above analysis, the RMT algorithm for signal subspace rank estimation can be defined by

$$d_{RMT} = \arg \min_k \left( \delta^2_{\tau_p,N} + z(P_{fa}) \delta_{\tau_p,N} \right)$$

(44)

where $k \in \{1, 2, \ldots, \min(N, \tau_p) - 1\}$, $z(P_{fa})$ satisfies that $F_{TW}(z(P_{fa})) = P_{fa}$. $P_{fa}$ denotes the false alarm probability. In addition, with the assumption of $K$ signals, the noise variance can be estimated by maximum likelihood estimator as

$$\delta^2_{\nu,k} = \frac{1}{N-K} \sum_{i=K+1}^{N} I_i$$

(45)

The above RMT approach can be used for signal subspace rank estimation based on the fact whether $d_{RMT} = 1$ or $d_{RMT} > 1$. Hence, the algorithm can only be performed in small set that is $k \in \{1, 2\}$. Furthermore, this detection algorithm can be only operated at one AP. To ensure the detection performance, we can choose the AP which is nearest to the user and the Eve for active attack detection. Therefore, the AP can achieve the largest SNR and can get a larger detection probability.

### Numerical results

In the simulations, all APs, users, and the Eve are randomly located over a circle cell of radius 300 m. A simplified path-loss model has been used as $\beta = (d_0/d)^\nu$, where $d$, $d_0$, and $\nu$ represent real geographical distance, reference distance, and loss coefficient, respectively. To quantitatively illustrate the results, key system parameters are set that $d_0 = 50m$, $\nu = 2.4$, $p_\nu = p_d = 10dBW$, $\tau_p = K = 20$, and $MN = 200$. Besides, the average power allocation scheme has been used in the simulations. All shown results were averaged over 5000 independent Monte Carlo trials.

Figure 2 depicts the NMSE of the channel estimation versus the SNR of the legitimate user's pilot signal as a function of the quantization bits and Eve’s pilot power. It is noted that NMSE is monotonically increasing for decreasing the quantization bits $b$. As expected, increasing the power of the active eavesdropper has a negative effect on the MMSE channel estimation. The dashed horizontally lines in Figure 2 denote the asymptotic limits of the NMSE in high power regime. The results illustrate that the coarse low-resolution ADCs result in a nonzero floor on NMSE which cannot be compensated by increasing the pilot power. On the contrary, there are no nonzero floors with implementing perfect ADCs at the transceivers which illustrates that equipping perfect ADCs could achieve an error-free channel estimation with larger pilot power.

Since the derived secrecy rate expression in Theorem 1 is an appropriate tool for secrecy performance measure, the tightness of expression (23) should be verified via numerical results. Figure 3 shows the analytical and numerical results versus the number of antennas.
where the results are achieved by averaging over 5000 channel realizations and randomly locations. It is observed that the analytical results can match exactly with the simulated values which can validate the correctness of the obtained analytical expressions. As can be readily noted, deploying more APs yields better secrecy performance which can be explained that both the channel responses and the diversity gains can be improved with deploying more APs. From Figure 1, another interesting observation is that the system suffers a trivial performance loss by low-resolution ADCs of the APs, but cannot tolerate low-quality ADCs at the users.

Figure 4 shows the achievable ergodic secrecy rate versus different ADC quantization bits with $p_E = 0$ dBW. Obviously, we can observe that the achievable ergodic secrecy rate increases with the ADC quantization bits. It can be explained that low-quality ADCs at APs and users bring larger channel estimation error and more signal quantization noise, which can cause a lower SINR (low achievable rate) at the target users. Besides, with the increase in the bits, the secrecy rate increases and converges to a finite rate limit (the dashed horizontally lines in Figure 4) which can be explained that the quantization distortion factor is very close to 1 when $b \geq 5$. As expected, the considered cell-free massive MIMO system can obtain more secrecy rate with high-resolution ADCs at the users. Hence, low-resolution ADCs can be widely used at APs instead of the users.

Figure 5 depicts the achievable secrecy rate as a function of the pilot power ratio between the Eve and the legitimate user, where $b_m$ and $b_u$ denote the quantization bits of low-resolution ADCs at APs and users, $+ \infty$ represents perfect high-resolution ADC. Interestingly, the results in Figure 2 further confirm the conclusions in Figure 1. Besides, Figure 5 illustrates that the antenna configuration at APs brings a significant impact on system performance. With the fixed total number of antennas, more APs can provide better quality of service, which is consistent with the results in Hu et al. Also, Figure 5 shows that the Eve can overhear information to one target user with a less power budget when the system exploits low-resolution ADCs. Hence, a trade-off between costs and performance is required in practical system design.

In Figure 6, the active attacks detection probability versus the pilot power of the Eve with various parameters are depicted when $p_p = 10$ dBW and $\alpha = 0.5$, where the results are obtained by 10,000 independent runs. The curves in Figure 6 illustrate that the detection performance improves with the increasing ADCs quantization bits and the antenna number of each AP. Moreover, it can be observed that the proposed RMT detector can detect active attacks at relative low power level. Also, we can note that the detection probability is over 95% when $p_E - p_p = -5$ dB for all parameter choices shown in Figure 6. Specifically, with more antennas at each AP ($N \geq 10$), the detection probability can reach to 1 when the Eve’s pilot power is about 6 dB lower than that of the legitimate user.

Figure 7 shows the detection probability of the proposed RMT-based approach versus the power allocation factor $\alpha$ with fixed total power $p_p = 10$ dBW and $p_E = 5$ dBW. At first, the results in Figure 7 indicate that the detection probability increases with the increasing $\alpha$. This can be explained that the correlation between the two column vectors of matrix $G$ defined in equation (47) decreases with increasing $\alpha$ which can lead to one relatively larger eigenvalue $s_{2,1}$. It is
propitious for promoting the RMT-based detection probability. The used pilot signal is publicly known and predesigned, while the detection signal is designed with adding a random sequence. Besides, a smart eavesdropper will wiretap and transmit the spoof pilot sequence. Thus, to make the detection procedure confidential and imperceptible, the power allocation factor should be as small as possible. For comparison, Figure 7 shows the detection performance under different $N$ and ADCs quantization bits $b_m$, which illustrates that the required allocation power for random sequence decreases with increasing $N$ or $b_m$ with the fixed detection probability. These results are valuable and significant for practical detection algorithm design and application.

Conclusion

This article adopted the well-established distortion model (AQNM) to reveal the secure performance of cell-free massive MIMO system with low-resolution ADCs. Uplink MMSE channel estimation has been operated in the presence of active eavesdropping attack which shows that there was a nonzero floor on the channel estimation error brought by the ADC impairments. Based on the imperfect CSI, conjugate beamforming has been performed for downlink impairments. Then, we derived the closed-form expression of the achievable ergodic secrecy rate. To weaken the influence of the active attacks, any effective resisting measure was to detect the presence of the active eavesdropper. This article exploited the RMT-based detection approach to detect the dimension of the signals embedded in noise space. Finally, numerical results have been presented to verify derived results and the effectiveness of the active attack detection scheme.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by China Postdoctoral Science Foundation under Grant 2021MD703980, in part by the National Natural Science Foundation of China under Grant 61901502.

ORCID iD

Xianyu Zhang https://orcid.org/0000-0002-2037-0231

References

1. Marzetta TL. Massive MIMO: an introduction. Bell Labs Tech J 2015; 20: 11–22.
2. Larsson EG, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems. IEEE Commun Mag 2014; 52(2): 186–195.
3. Björnson E, Larsson EG and Marzetta TL. Massive MIMO: ten myths and one critical question. IEEE Commun Mag 2016; 54(2): 114–123.
4. Zhu J, Schober R and Bhargava VK. Secure transmission in multicell massive MIMO systems. IEEE Trans Wirel Commun 2014; 13(9): 4766–4781.
5. Zhang J, Wen C, Jin S, et al. On capacity of large-scale MIMO multiple access channels with distributed sets of correlated antennas. IEEE J Select Area Commun 2013; 31(2): 133–148.
6. Guo K, Guo Y and Ascheid G. Security-constrained power allocation in MU-massive-MIMO with distributed antennas. IEEE Trans Wirel Commun 2016; 15(12): 8139–8153.
7. Ngo HQ, Ashikhmin A, Yang H, et al. Cell-free massive MIMO versus small cells. *IEEE Trans Wirel Commun* 2017; 16(3): 1834–1850.

8. Ngo HQ, Tran L, Duong TQ, et al. On the total energy efficiency of cell-free massive MIMO. *IEEE Trans Green Commun Netw* 2018; 2(1): 25–39.

9. Mai TC, Ngo HQ and Duong TQ. Downlink spectral efficiency of cell-free massive MIMO systems with multi-antenna users. *IEEE Trans Commun* 2020; 68(8): 4803–4815.

10. Björnson E and Sanguinetti L. Making cell-free massive MIMO competitive with MMSE processing and centralized implementation. *IEEE Trans Wirel Commun* 2020; 19(1): 77–90.

11. Yang N, Wang L, Geraci G, et al. Safeguarding 5G wireless communication networks using physical layer security. *IEEE Commun Mag* 2015; 53(4): 20–27.

12. Guo K, An K, Zhang B, et al. Physical layer security for multiuser satellite communication systems with threshold-based scheduling scheme. *IEEE Trans Veh Technol* 2020; 69(5): 5129–5141.

13. Guo K, An K, Zhou F, et al. On the secrecy performance of NOMA-based integrated satellite multiple-terrestrial relay networks with hardware impairments. *IEEE Trans Veh Technol* 2021; 70(4): 3661–3676.

14. Kapetanović D, Zheng G and Rusek F. Physical layer security for massive MIMO: an overview on passive eavesdropping and active attacks. *IEEE Commun Mag* 2015; 53(6): 21–27.

15. Al-Nahari A. Physical layer security using massive multiple-input and multiple-output: passive and active eavesdroppers. *IET Commun* 2016; 10(1): 50–56.

16. Zhu J, Schober R and Bhargava VK. Linear precoding of data and artificial noise in secure massive MIMO systems. *IEEE Trans Wirel Commun* 2016; 15(3): 2245–2261.

17. Zhang X, Guo D and Guo K. Secure performance analysis for multi-pair AF relaying massive MIMO systems in Ricean channels. *IEEE Access* 2018; 6: 57708–57720.

18. Wu Y, Wen C, Chen W, et al. Data-aided secure massive MIMO transmission under the pilot contamination attack. *IEEE Trans Commun* 2019; 67(7): 4765–4781.

19. Hoang TM, Ngo HQ, Duong TQ, et al. Cell-free massive MIMO networks: optimal power control against active eavesdropping. *IEEE Trans Commun* 2018; 66(10): 4724–4737.

20. Zhang X, Guo D, An K, et al. Secure communications over cell-free massive MIMO networks with hardware impairments. *IEEE Syst J* 2020; 14(2): 1909–1920.

21. Hu X, Zhong C, Chen X, et al. Cell-free massive MIMO systems with low resolution ADCs. *IEEE Trans Commun* 2019; 67(10): 6844–6857.

22. Zhang Y, Zhou M, Qiao X, et al. On the performance of cell-free massive MIMO with low-resolution ADCs. *IEEE Access* 2019; 7: 117968–117977.

23. Zhang Y, Zhou M, Cao H, et al. On the performance of cell-free massive MIMO with mixed-ADC under Rician fading channels. *IEEE Commun Lett* 2020; 24(1): 43–47.

24. Teeti MA. Downlink secrecy rate of one-bit massive MIMO system with active eavesdropping. *IEEE Access* 2020; 8: 37821–37842.

25. Xiong Q, Liang Y, Li KH, et al. Secure transmission against pilot spoofing attack: a two-way training-based scheme. *IEEE Trans Inform Foren Sec* 2016; 11(5): 1017–1026.

26. Zhang X, Guo D, Guo K, et al. Secure performance analysis and detection of pilot attack in massive multiple-input multiple-output system. *Int J Distrib Sens Netw* 2018; 14(5): 1–12.

27. Kapetanović D, Zheng G, Wong KK, et al. Detection of pilot contamination attack using random training and massive MIMO. In: Proceedings of the 2013 IEEE 24th annual international symposium on personal, indoor, and mobile radio communications (PIMRC), London, 8–11 September 2013, pp.13–18. New York: IEEE.

28. Kapetanović D, Al-Nahari A, Stojanovic A, et al. Detection of active eavesdroppers in massive MIMO. In: Proceedings of the 2014 IEEE 25th annual international symposium on personal, indoor, and mobile radio communication (PIMRC), Washington, DC, 2–5 September 2014, pp.585–589. New York: IEEE.

29. Tugnait JK. Self-contamination for detection of pilot contamination attack in multiple antenna systems. *IEEE Wirel Commun Lett* 2015; 4(5): 525–528.

30. Tugnait JK. Detection and identification of spoofed pilots in TDD/SDMA systems. *IEEE Wirel Commun Lett* 2017; 6(4): 550–553.

31. Kritchman S and Nadler B. Non-parametric detection of the number of signals: hypothesis testing and random matrix theory. *IEEE Trans Signal Process* 2009; 57(10): 3930–3941.

32. Tugnait JK. Pilot spoofing attack detection and countermeasure. *IEEE Trans Commun* 2018; 66(5): 2093–2106.

33. Guo K, Lin M, Zhang B, et al. On the performance of LMS communication with hardware impairments and interference. *IEEE Trans Commun* 2019; 67(2): 1490–1505.

34. Guo K, Lin M, Zhang B, et al. Performance analysis of hybrid satellite-terrestrial cooperative networks with relay selection. *IEEE Trans Veh Technol* 2020; 69(8): 9053–9067.

35. Guo K, An K, Zhang B, et al. On the performance of the uplink satellite multiterrestrial relay networks with hardware impairments and interference. *IEEE Syst J* 2019; 13(5): 2297–2308.

36. Zhou F, Chu Z, Sun H, et al. Artificial noise aided secure cognitive beamforming for cooperative MISO-NOMA using SWIPT. *IEEE J Select Area Commun* 2018; 36(4): 918–931.

37. Zhou F, Li Z, Cheng J, et al. Robust AN-aided beamforming and power splitting design for secure MISO cognitive radio with SWIPT. *IEEE Trans Wirel Commun* 2017; 16(4): 2450–2464.

38. Timilsina S, Kudathanthirige D and Amarasuriya G. Physical layer security in cell-free massive MIMO. In: *Proceedings of the 2018 IEEE global communications conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 9–13 December 2018, pp.1–7. New York: IEEE.
Appendix 1

Proof of Theorem 1

According to the MMSE channel estimation expression (9), we first focus on the term $E\left\{ g_{ml}^H \hat{y}_{p,m,l} \right\}$ as

$$
E\left\{ g_{ml}^H \hat{y}_{p,m,l} \right\} = \gamma_m E\left\{ g_{ml}^H \left( \sqrt{\tau_p \rho_p} \mathbf{e}_{ml} + \sqrt{\tau_p \rho_p} \mathbf{g}_{ml} \mathbf{f}_l \right) \right\} + \gamma_m E\left\{ g_{ml}^H \mathbf{w}_{p,m} \right\} + \gamma_m E\left\{ g_{ml}^H \mathbf{w}_{p,m} \right\}
$$

(46)

Since $g_{ml}$, $\mathbf{e}_{ml}$, $\mathbf{w}_{p,m}$, and $\mathbf{w}_{p,m}$ are all zero-mean and mutual independent, we can derive the result as

$$
E\left\{ g_{ml}^H \hat{y}_{p,m,l} \right\} = \gamma_m \sqrt{\tau_p \rho_p} E\left\{ g_{ml}^H \mathbf{g}_{ml} \right\}
$$

(47)

To calculate the term $E\left\{ \hat{y}_{p,m,l}^H \hat{y}_{p,m,l} \right\}$, the term $E\left\{ g_{ml}^H \mathbf{w}_{p,m} \right\}$ should be computed based on equation (4), which can be given by

$$
E\left\{ g_{ml}^H \mathbf{w}_{p,m} \right\} = \gamma_m (1 - \gamma_m) E\left\{ \mathbf{y}_{p,m,l} \mathbf{y}_{p,m,l} \right\} = N \gamma_m (1 - \gamma_m) \left( \tau_p \rho_p \beta_{ml} + \tau_p \rho_p \beta_{ml} \mathbf{f}_l \mathbf{f}_l + 1 \right)
$$

(48)

Furthermore, we can derive that

$$
E\left\{ \hat{y}_{p,m,l}^H \hat{y}_{p,m,l} \right\} = \gamma_m E\left\{ \mathbf{y}_{p,m,l} \mathbf{y}_{p,m,l} \right\} = N \gamma_m \left( \tau_p \rho_p \beta_{ml} + \tau_p \rho_p \beta_{ml} \mathbf{f}_l \mathbf{f}_l + 1 \right)
$$

(49)

Plugging equations (46) and (48) into equation (9), the norm-square of the channel estimation $\mathbf{g}_{ml}$ can be derived that

$$
E\left\{ \mathbf{g}_{ml}^H \mathbf{g}_{ml} \right\} = \left( \frac{E\left\{ g_{ml}^H \hat{y}_{p,m,l} \right\}}{E\left\{ \hat{y}_{p,m,l}^H \hat{y}_{p,m,l} \right\}} \right)^2 = N \lambda_{ml}
$$

(50)

Besides, according to the MMSE principle, the channel estimation $\mathbf{g}_{ml}$ and estimation error $\mathbf{e}_{ml}$ are uncorrelated. Hence, with the aid of $E\left\{ g_{ml}^H \right\} = 0$, it yields that

$$
E\left\{ \mathbf{g}_{ml}^H \mathbf{e}_{ml} \right\} = E\left\{ \mathbf{g}_{ml}^H \mathbf{g}_{ml} \right\} - E\left\{ \mathbf{g}_{ml}^H \mathbf{g}_{ml} \right\} = N (\beta_{ml} - \lambda_{ml})
$$

(51)

This finishes the proof.

Appendix 2

Proof of Theorem 2

To compute the SINR of the target legitimate user, we need to consider the terms $E\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\}$, $Var\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\}$, and $E\left\{ \mathbf{w}_{k}^2 \right\}$.

We start with the term $E\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\}$, which can be given by

$$
E\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\} = E\left\{ \sum_{m=1}^{M} \eta_{mk} (\mathbf{g}_{mk}^T + \mathbf{g}_{mk}^T)^2 \right\}
$$

(52)

Then, the term $Var\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\}$ can be calculated as

$$
Var\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\} = E\left\{ \left| \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right|^2 \right\} - \left| E\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\} \right|^2
$$

(53)

Furthermore, we can derive that

$$
E\left\{ \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\} = E\left\{ \sum_{m=1}^{M} \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\}
$$

(54)

Then, we compute the term $E\left\{ \left| \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right|^2 \right\}$ as

$$
E\left\{ \left| \sum_{m=1}^{M} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right|^2 \right\} = E\left\{ \sum_{m=1}^{M} \sum_{m=1}^{M} \eta_{mk} \eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk} \right\}
$$

(55)

$$
= N \sum_{m=1}^{M} \eta_{mk} \lambda_{mk} \beta_{mk}
$$
Next, we obtain the evaluation of $\mathbb{E}\{|\hat{w}_k|^2\}$ which can be given by

$$\mathbb{E}\{|\hat{w}_k|^2\} = \gamma_w(1 - \gamma_w)(1 + D_k) \quad (56)$$

where the term $D_k$ is defined as

$$D_k = \sum_{k' = 1}^{K} \sum_{k'' = 1}^{K} \sum_{m' = 1}^{M} \sum_{m'' = 1}^{M} \sqrt{\eta_{mk'}} \eta_{m'} \cdot \mathbb{E}\{\hat{\mathbf{g}}_{mk'}^T \hat{\mathbf{g}}_{mk''} \hat{\mathbf{g}}_{m'}^T \hat{\mathbf{g}}_{m''}\}$$

$$= N \sum_{k'' = 1}^{K} \sum_{m'' = 1}^{M} \eta_{mk} \lambda_{mk} \beta_{mk}$$

$$+ N^2 \sum_{m' = 1}^{M} \sum_{m = 1}^{M} \sqrt{\eta_{mk'} \eta_{m'} \lambda_{mk} \lambda_{m'k}} \quad (57)$$

Substituting equations (52), (53), (55), and (56) into equation (24), we can derive the SINR expression of equation (26).

Then, we discuss the SINR $\xi_k^{Eve}$ of the Eve. Due to the mutual independence of the effective signal and additive noise, the different parts in equation (17) can be separately derived. We first calculate that

$$\mathbb{E}\left\{\sum_{m = 1}^{M} \sqrt{\eta_{mk} \mathbf{g}_{mk}^T \mathbf{g}_{mk}}^2\right\}$$

$$= \sum_{m = 1}^{M} \eta_{mk} \mathbb{E}\{||\mathbf{g}_{mk}||^2\}$$

$$= \sum_{m = 1}^{M} \eta_{mk} \mathbb{E}\{||\mathbf{g}_{mk}\|^2\} + \mathbb{E}\{||\hat{\mathbf{g}}_{mk}\|^2\}$$

$$= \sum_{m = 1}^{M} \eta_{mk} \left(\mathbb{E}\{||\mathbf{g}_{mk}\|^2\} + \mathbb{E}\{||\hat{\mathbf{g}}_{mk}\|^2\}\right)$$

$$(a) = \sum_{m = 1}^{M} \eta_{mk} \left(\frac{N^2 \rho_{mk}^2}{\rho_{mk}^2 + \rho_{mk}} \lambda_{mk}^2 + N\beta_{mk} \lambda_{mk}\right)$$

Here, (a) follows two properties: (1) the independence of $\mathbf{g}_{mk}$ and $\hat{\mathbf{g}}_{mk}$ and (2) the equivalent $\mathbb{E}\{||\mathbf{g}_{mk}||^2\} = N(N + 1) \lambda_{mk}^2$. Similarly, using the mutual independence of $\mathbf{g}_{mk}$ and $\mathbf{g}_{ml}(l \neq k)$, we have

$$\mathbb{E}\{||\mathbf{g}_{mk}||^2\} = N\beta_{mk} \lambda_{mk} \quad (59)$$

Plugging these results into equation (25) yields the expression as equation (27), which completes the proof.