Security performance analysis for cell-free massive multiple-input multiple-output system with multi-antenna access points deployment in presence of active eavesdropping

Xiaoyu Wang®, Yuanyuan Gao, Guangna Zhang, Mingxi Guo and Kui Xu

Abstract
This article investigates the influence of multiple-antenna deployment at access points on physical layer security under cell-free massive multiple-input multiple-output systems. To embody the concept of network multiple-input multiple-output, cell-free massive multiple-input multiple-output employs a great deal of geographically distributed access points, which jointly serve multiple users in the same frequency-time resource. Assuming a malicious active eavesdropper having knowledge of instantaneous channel state information, we derive the asymptotic closed-form expression of the achievable secrecy rate, which is used to evaluate network security performance. A distributed power allocation scheme is developed based on channel fading degrees of different users. All of access points distribute a larger proportion of local transmission power to the users with better channel conditions. We verified that the security performance of the multi-antenna access points outperforms that of the single one. Compared to equal power allocation, the proposed power control algorithm can further boost the network security performance.

Keywords
Cell-free massive multiple-input multiple-output, physical layer security, multi-antenna technology, active eavesdropping, achievable secrecy rate, distributed power allocation scheme

Date received: 13 March 2021; accepted: 16 May 2022
Handling Editor: Yanjiao Chen

Introduction
Telatar,1 Foschini and Gans,2 Bell Laboratory, in their pioneering paper proposed the concept of multiple-input multiple-output (MIMO) system, which can provide spatial multiplexing and diversity gained by using multiple antennas on transmitter and receiver. Thus, MIMO systems can significantly increase spectral efficiency throughout without additional frequency, power, and time resources. As a scalable multiuser MIMO implementation, massive MIMO technology has been derived by Marzetta.3,4 Each base station (BS) is equipped with a large array of antenna elements. Massive MIMO has become one of the key technologies for 5G networks with a great capability of

Arm y Engineering University of PLA, Nanjing, China

Corresponding author:
Yuanyuan Gao, Army Engineering University of PLA, Nanjing 210007, China.
Email: njyygao@sina.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
enhancing the spectrum and energy efficiency as well as reducing transmission power and inter-cell interference of the network through the deep mining of dimensional resources in a wireless communication environment.  

According to the deployment of antenna arrays on BS, massive MIMO systems can fall into collocated MIMO systems and distributed ones. The former is built with a compact large-scale antenna array in the BS, which has advantages of low data sharing cost and fronthaul requirements. The antenna of the latter is distributed over a wide area, which can provide uniformly good services for all authorized users as well as higher diversity against the shadow fading. Multicell distributed MIMO is a generalized concept, including network MIMO, coordinated multiple point (CoMP), and distributed antenna systems (DASs). In 2015, Nayebi et al. first proposed the network architecture of cell-free massive MIMO, which has received widespread attention. The different structure of massive MIMO systems is illustrated in Figure 1. Chien et al. investigated how to use power allocation to save load resources. To reduce the overhead for channel estimation, an aggregated channel estimation approach was introduced by Chien et al. These studies have explored more performance advantages for cell-free massive MIMO.

With the dramatic array gain, massive MIMO has great potential in physical layer security (PLS). Also, this topic has been developed in recent years. Due to the stronger directivity of the signal beam, the massive MIMO system has natural robustness to passive eavesdropping. To obtain more favorable information, smart eavesdroppers usually take action not only by overhearing but also by jamming. Time-division duplexing (TDD) mode has the advantage of making forward and reverse channels under massive MIMO systems reciprocal. It allows that only the uplink is used for channel estimation to acquire channel state information (CSI), dispensing with downlink estimation. At the same time, it provides a huge opportunity for active eavesdroppers to send spoofing pilot sequences, also called pilot spoofing attacks. The damage is to decrease the accuracy of channel estimation, and bring correlated precoding vectors between eavesdroppers and legitimate users, resulting in severe leaking of information. To mitigate the impact of active eavesdropping in massive MIMO systems, power control methods are commonly applied, such as trying to ensure the minimum transmit power required for reliable and secure communication, minimizing the signal-to-interference and -noise ratio (SINR) of eavesdroppers, and using game theory to maximize the SINR of legitimate users.

The security aspect of the cell-free massive MIMO network under pilot spoofing attacks was first explored by Hoang et al., which proposed a simple detection mechanism and designed power control schemes with the aim of security rate maximization and power consumption minimization. To continue handling the crisis of the active eavesdropper, Zhang et al. analyzed the security performance for the multigroup multicasting scenario. Besides, a comparison about the lower bound on security capacity between the collocated and cell-free massive MIMO has been presented by Timilsina et al. The achievable rate on legitimate users and the rate leaked to active eavesdroppers have been driven within the range of limited and unlimited numbers of APs. Research on security performance was expanded by introducing novel technologies. Alageli et al. have delved into the problem of secure communication assuming that there is an untrusted dual-antenna energy harvester in the network. One antenna is used to harvest energy, whereas the other one intercepts information signals as an active eavesdropper. Elhoushy et al. have enhanced the robustness of network security against pilot spoofing attacks by applying reconfigurable intelligent surface, which reconfigures the wireless propagation channels through software-controlled reflections. Based on more realistic
consideration of low-resolution analog-to-digital converters, secure communication with active eavesdroppers has been studied by Zhang et al.36 However, the abovementioned research on PLS is all limited to single-antenna APs.

To the best of the authors’ knowledge, there are no fruitful results with the cell-free massive MIMO system model established in multi-antenna APs. However, it is easier to realize channel hardening when each AP is deployed with 5–10 antennas.27 By channel hardening, fading multi-antenna channels can behave almost as deterministic scalar channels. This basic property simplifies the resource allocation28 and assists to deliver higher performance29 in the cellular massive MIMO system. Ngo et al.7 proved that the cell-free massive MIMO can also utilize channel hardening, but Chen and Bjoernson30 have proved that channel hardening is not always available to adopt in the cell-free massive MIMO network. So the studies on cell-free massive MIMO should deploy multi-antenna APs in the pursuit of relying on channel hardening. Vu et al.31 proposed a novel full-duplex mode, wherein the plurality of multi-antenna APs provided uplink and downlink services for multiple single-antenna users in the same frequency band. Considering imperfect CSI, non-orthogonal pilots, and power control, the spectral efficiency with respect to numbers of APs antennas and the users in cell-free massive MIMO system has been investigated by Mai et al.32

Motivated by the aforementioned discussion, we explore the PLS of cell-free massive MIMO systems with multi-antenna APs in the presence of an active eavesdropper. First of all, the closed-form expressions of downlink achievable rates at the legitimate users and the rate leaked to the wiretapper are derived. Second, we propose a distributed power control scheme that allocates power factors by estimated channel quality. This simple and feasible scheme puts power control operations locally. Moreover, the effect of the number of the antennas per AP is analyzed by the achievable secrecy rate. Simulation results show that the security improved by the multi-antenna APs consistently outperforms that of the single one in terms of achievable security rate. And dense nodes of APs cooperative with the aforesaid power allocation scheme can provide a respectable security performance.

The remainder of this article is organized as follows. In the section “Channel and System Model,” the system model including uplink channel estimation and downlink information transmission phase are presented. In section “Security Performance Analysis,” we carry out a set of general mathematical closed formulas for security performance analysis. In addition, we optimize the total downlink transmission power here. This is followed by the section “Numerical Results and Discussion,” where numerical results and discussion are presented. Finally, we provide our conclusion in the section “Conclusions.”

Notations: Boldface letters denote the vector or matrix, for example, \( \mathbf{x} \). \( \mathbf{I}_N \) denotes the \( N \times N \) identity matrix. Conjugate, transpose, Hermitian, and trace operator are indicated by \((\cdot)^*, (\cdot)^T, (\cdot)^H, \text{ and } tr(\cdot)\), respectively. \( E\{\} \) and \( \text{Var}\{\} \) denote expectation and variance operations of random variables in brackets, respectively. \( \mathbb{C}^{M \times N} \) stands for \( M \times N \) dimensional complex space. \( |\cdot| \) and \( \|\cdot\|_2 \) represent absolute value and the Euclidean norm, respectively. \( \mathbf{x} \sim \mathcal{CN}(\mathbf{x}, \mathbf{A}) \) is a complex Gaussian random vector with a mean vector of \( \mathbf{x} \) and covariance matrix of \( \mathbf{A} \).

### Channel and System Model

#### Channel Model

We consider a cell-free massive MIMO system with \( M \) APs and \( K \) users in the presence of an active eavesdropper (Eve). Each AP is equipped with \( N \) antennas, and each user and Eve have one single antenna. All node positions are randomly located. The central processing unit (CPU) connects every AP nodes to perform joint signal processing, beamforming, and computing tasks via backhaul, such as fiber, cable, and microwave.\(^6\) Let us denote the channel response matrix between the AP and the user as \( \mathbf{g}_{mk} \in \mathbb{C}^{N \times 1} \), which can be expressed as

\[
\mathbf{g}_{mk} = \beta_{mk}^{1/2} \mathbf{h}_{mk}
\]

where \( \beta_{mk} \) is large-scale fading including shadow fading and path loss. It is related to the distance between the \( m \)th AP and the \( k \)th user, which can be considered as a constant.\(^{33} \) \( \mathbf{h}_{mk} \in \mathbb{C}^{N \times 1} \) represents small-scale fading,
which is composed of the superposition of the signal components generated by the reflection, scattering, and diffraction of the channel through various obstacles.

In addition, we discuss the Rayleigh fading channel, that is, $[h_{mk}]_i \sim \mathcal{CN}(0, 1)$ is independent and identically distributed (i.i.d). The notation $[h_{mk}]_i$ is the $i$th element of $h_{mk}$, $i \in \{1, \ldots, N\}$.

**System model**

The system model of the cell-free massive MIMO network is shown in Figure 2. We assume that the system is operated in TDD mode. The length of a TDD coherent interval is $\tau^p$. Considering the downlink secure transmission, each coherence interval consists of two phases: uplink channel estimation and downlink information transmission.

**Uplink channel estimation:** During the pilot sequence training duration $\tau^p$, every legitimate user sends its own unique pilot sequence to all APs, and APs process these orthogonal mixed pilot sequences by multiplying the conjugation of the user’s own pilot sequence. In this article, we adopt the minimum mean square error (MMSE) estimation method, often referred to as the optimal estimate.

In detail, the system design $K$ vectors $p_1, p_2, \ldots, p_K$ as standard and public pilot sequences. It has to satisfy $p_i^H p_j = 1$ in the case of $i = j$, and $p_i^H p_j = 0$ in the case of $i \neq j$. The $k$th user sends $\sqrt{\tau^p p_k} \in \mathbb{C}^{\tau^p \times 1}$ to all APs, and Eve sends the same one with the target intercepted user. Generally assuming that Eve desires to wiretap the confidential messages from the $k$th user, that is, $p_E = p_k$, the process is completely synchronous. Hence, the received pilot vector at the $m$th AP is

$$Y_{p, m} = \sqrt{\tau^p} P_{ul} \sum_{k=1}^{K} g_{mk} p_k^H + \sqrt{\tau^p} P_{uc} h_{mk} p_E^H + W_{p, m} \tag{2}$$

where $P_{ul} \triangleq P_{ul} / \sigma^2$ and $P_{uc} \triangleq P_{uc} / \sigma^2$ are normalized signal-to-noise ratio (SNR) of each pilot symbol. $P_{ul}$ and $P_{uc}$ are the average uplink power of each user and Eve, respectively. $W_{p, m} \in \mathbb{C}^{N \times \tau_v}$ is an additive white Gaussian noise (AWGN) matrix. The elements of $W_{p, m}$ obey distribution with $\mathcal{CN}(0, 1)$, i.i.d. Then, the projection process can be put into $p_k^H$ as

$$Y_{lm} = Y_{p, m} p_k^H = \begin{cases} \sqrt{\tau^p} P_{ul} g_{mk} + W_{p, m} p_k^H, & k \neq k_0 \\ \sqrt{\tau^p} P_{ul} g_{mk} + \sqrt{\tau^p} P_{uc} h_{mk} + W_{p, m} p_k^H, & k = k_0 \end{cases} \tag{3}$$

The MMSE estimation of $g_{mk}$ and $g_{ml}$ can be given locally at the $m$th AP by

$$\hat{g}_{mk} = \sqrt{\tau^p} P_{ul} \beta_{mk} A_{km}^{-1} y_{km} \tag{4}$$

where

$$\begin{align*}
A_{km} &= E\left\{ y_{km} y_{km}^H \right\} \\
&= \begin{cases} (\tau^p P_{ul} \beta_{mk} + 1) I_N, & k \neq k_0 \\
(\tau^p P_{ul} \beta_{mk} + \tau^p P_{uc} \beta_{ml} + 1) I_N, & k = k_0 \end{cases} \tag{6}
\end{align*}$$

To facilitate later calculations, let us define

$$\begin{align*}
\gamma_{mk} &= E\left\{ \hat{g}_{mk} \hat{g}_{mk}^H \right\} = \tau^p P_{ul} \beta_{mk}^2 A_{km}^{-1} \tag{7}
\end{align*}$$

and

$$\begin{align*}
\gamma_{mk} &= E\left\{ \hat{g}_{mk} \hat{g}_{mk}^H \right\} = \tau^p P_{uc} \beta_{ml}^2 A_{km}^{-1} \tag{8}
\end{align*}$$

We can see that $\gamma_{mk}$ is a diagonal matrix with identical elements, and so is $\gamma_{mk}$. These two statistics describe the estimated large-scale fading of APs. Assuming the antennas by each AP are uncorrelated, they are expressed directly and briefly as $\gamma_{mk} = \lambda_{mk} I_N$, $\gamma_{mk} = \lambda_{mk} I_N$.

**Downlink data transmission:** During this phase, the $m$th AP expects to transmit the normalized signal $s_d$ for the $k$th user such that $E\{|s_d|^2\} = 1$ and utilizes the uplink estimated channel to perform beamforming. Then, the AP transmits the mixed signal for $K$ users as

$$s_m = \sum_{k=1}^{K} \sqrt{P_m} \omega_{mk} s_d \tag{9}$$

where $P_m$ is the downlink power allocated by the $m$th AP to the $k$th user. $\omega_{mk} \in \mathbb{C}^{N \times 1}$ represents the payload data precoding vector of the $m$th AP to the $k$th user. Maximum ratio transmission (MRT) is a simple linear precoding scheme that can provide nearly optimal performance. Therefore, using MRT scheme in this article, we define $\omega_{mk}$ as $\omega_{mk} = \hat{g}_{mk} / \sqrt{E\{|\hat{g}_{mk}|^2\}}$. For the sake of simplicity, the operation is aimed at $E\{|s_m|^2\} = \sum_{k=1}^{K} P_m$. After that, the received signal at the $k$th user from all APs is given by

$$x_k = \sum_{m=1}^{M} \hat{g}_{mk}^H s_m + w_{d, k} \tag{10}$$

where $a_{mk} = \sum_{m=1}^{M} \sqrt{P_m} \hat{g}_{mk}^H \omega_{mk}$. It means the strength of the message that the user has received is...
however intended for the k’th user. And 

\[ w_{d,k} \sim CN(0, \sigma^2) \]

is an AWGN.

As such, Eve will receive the information signal as

\[
z_E = \sum_{m=1}^{M} a_{ek}^H \mathbf{s}_m + w_{d,e}\]

\[
= a_{ek}^H s_{dk} + \sum_{k' \neq k} a_{ek'}^H s_{dk'} + w_{d,e},
\]

(11)

where \( a_{ek} \triangleq \sum_{m=1}^{M} \sqrt{P_{m}} \mathbf{u}_{mk}^H \omega_{mk} \) is the intercepted signal gain. \( w_{d,e} \) is AWGN satisfying \( CN(0, \sigma^2) \).

**Security performance analysis**

In this section, we analyze the security performance by achievable secrecy rate. The security rate can be defined as the difference between the information rate of legitimate users and Eve. Also, the achievable rate of legitimate users and the information leakage rate into Eve are deduced.

**Achievable rate of legitimate users**

On the basis of information theory, the achievable rate of the user can be given as

\[
R_k = \log_2 (1 + \gamma_k)
\]

(12)

where \( \gamma_k \) is the SINR of the \( k \)th user. Due to the fact that multi-antenna APs deployment makes it realize the property of channel hardening, legitimate users have access to statistic CSI without downlink training.\(^{27,30}\)

The achievable rate of legitimate users can be regarded as the lower bound for mutual information between \( x_k \) and \( y_k \). Based on Hoang et al.,\(^{21}\) and considering that the effective noises are pair-wisely uncorrelated, we can obtain the lower-bound SINR of the \( k \)th user as

\[
\gamma_k = \frac{\left| E \left\{ a_{kk} \right\} \right|^2}{\text{var}(a_{kk}) + \sum_{k' = 1, k' \neq k}^{K} E \left\{ \left| a_{kk} \right|^2 \right\} + \sigma^2}
\]

(13)

As the \( k \)th user only gains the statistic CSI, \( E \{a_{kk}\} \) is substituted for \( a_{kk} \) denoting the strength parts of \( s_{dk} \). \( a_{kk} - E \{a_{kk}\} \) is the strength difference of the desired signal caused by the inaccuracy of the statistic CSI, which expatiates on the source of the var\{a_{kk}\} in the above formula.

**Lemma 1.** Let \( \mathbf{x} \in \mathbb{C}^{L \times 1} \) be a complex-valued random vector with \( \mathbf{x} \sim CN(0, \Lambda) \), 

\[
E \left\{ \left| \mathbf{H}^\dagger \mathbf{x} \right|^2 \right\} = \left| \text{tr}(\Lambda) \right|^2 + \text{tr}(\Lambda^2)
\]

**Proof:** The reader could be referred to Bjoernson et al.,\(^{33}\) Lemma 2.

**Theorem 1.** When an active eavesdropper conducts a pilot spoofing attack, and all APs are configured with multi-antenna and conjugate beamforming technology, the closed-form expression of the \( k \)th user’s SINR could be condensed as equation (14) by Lemma 1, where Qiu et al.,\(^{34}\) has not coped completely.

\[
\gamma_k = \frac{\left( \sum_{m=1}^{M} \sqrt{P_{m}} \omega_{mk}^H \right)^2}{\sum_{m=1}^{M} \sum_{k' \neq k} P_{m} + \sigma^2}
\]

(14)

**Proof:** See Appendix 1.

**Information rate leaked into Eve**

In this article, we focus on the worst-case scenario in terms of network security. The Eve has knowledge of instantaneous channel gain to get a genie-aided rate. Based on Jensen’s inequality, \( E \{ f(x) \} \geq f(E \{ x \}) \), if \( f(x) \) is convex. The Eve gets the upper-bound eavesdropping information rate, which can be represented as

\[
R_E = \log_2 (1 + \gamma_E)
\]

\[
= \log_2 \left( 1 + \frac{E \{ \left| a_{kk} \right|^2 \}}{\sum_{k' \neq k} P_{m} + \sigma^2} \right)
\]

(15)

**Theorem 2.** The closed-form expression of the active eavesdropper’s SINR when every AP has multi-antenna and adopts conjugate beamforming is

\[
\gamma_E = \frac{\sum_{m=1}^{M} \beta_{mk} \beta_{mk'} + \left( \sum_{m=1}^{M} \sqrt{P_{m}} \omega_{mk} \right)^2}{\sum_{m=1}^{M} \sum_{k' \neq k} P_{m} + \sigma^2}
\]

(16)

**Proof:** See Appendix 2.

**Achievable secrecy rate**

The above analysis lays the foundation for the derivation of the lower-bound secrecy rate at the \( k \)th user. Following Wu et al.,\(^{14}\) and Jose et al.,\(^{16}\) the achievable secrecy rate can be expressed as

\[
R_{sec} = \left[ R_k - R_E \right]^+
\]

(17)

where the notation is \( [x]^+ = \max(x, 0) \). We regard the worst-case secrecy performance for the network, which is of utmost importance. Under any other conditions,
such as legitimate users having perfect CSI, or Eve not being capable of achieving instantaneous CSI, the secrecy performance would be elevated spontaneously.

**Power allocation scheme**

In this subsection, we present a distributed power allocation (DPA) scheme, which is simple to implement for cell-free massive MIMO. Most power control algorithms take a certain utility metric as the optimization goal, depending on the powerful computing capability of the CPU, huge backhaul capacity, and perfect calibration of the hardware chains. Backhaul overhead is defined as the execution cost of computation task among cooperating APs.35 Because distributed antenna networks increase in size and density, it is urgent to reduce backhaul overhead.36 When the backhaul load is heavily limited, it is an alternative approach to conduct power control to APs locally. Assuming that the network does not detect the presence of active eavesdropping, each AP divides their power by the channel strength of different users. In this scenario, the influence of the unauthorized access of the active eavesdropper on the network can be more intuitive, credible, and persuasive. In our power allocation schemes, channel strength is measured on the basis of ${\gamma _{mk}} = {E[{|k_{mk}|^2}]} = \lambda _{mk}I_N$, so the downlink transmission power allocated from the $m$th AP to the $k$th user can be acquired as

$$\begin{align*}
    p_{mk} &= P_{dl} \frac{\lambda _{mk}}{\sum _{i=1} \lambda _{mk}}
\end{align*}$$

where $P_{dl}$ is the total transmission power of each AP.

Dealing with power control, it is traditional to adopt equal power allocation (EPA) scheme in collocated massive MIMO systems. Because the distance is so close among antennas in the BS, all are regarded to be set up in the same place. The channel quality between each antenna and users is in equilibrium. While CPU supports distributed scenarios in cell-free massive MIMO systems, APs have widely divergent locations. If a certain AP is relatively far from some users, or has serious interferences, it will be sensible to make these power control factors smaller. In consideration of the above, DPA is more appropriate in the light of the distribution network traits. It is also easy to implement after CSI estimated by each AP.

**Numerical results and discussions**

In this section, we present the numerical results of the security performance under the multi-antenna cell-free massive MIMO network in the presence of an active eavesdropper. Unless noted additionally, system parameters are set as shown in Table 1.

Besides, the amount of APs, users, and Eve nodes are uniformly distributed over the area of $1 \times 1$ km$^2$. Each point in the figure is shown by average value through 100 executions of the trivial for the location randomization. And there are $10^4$ times independent Monte-Carlo experiments verified for the following scenario: the large-scale fading by employing Hata-COST231 propagation model37 can be described as

$$\begin{align*}
    \beta _{mk} &= 10^{\frac{\alpha _{mk} + \beta _{mk}}{10}} \\
    \beta _{mE} &= 10^{\frac{\alpha _{mE} + \beta _{mE}}{10}}
\end{align*}$$

where $z_{mk} \sim N(0, 1)$ and $z_{mE} \sim N(0, 1)$ denote the normalized distribution of shadowing fading. $PL_{mk}$ and $PL_{mE}$ represent the path loss as

$$PL(d) = \begin{cases} 
    -L - 35\log _{10}(d) & \text{if } d > d_1 \\
    -L - 15\log _{10}(d_1) - 20\log _{10}(d) & \text{if } d_1 < d \leqslant d_2 \\
    -L - 15\log _{10}(d_1) - 20\log _{10}(d_0) & \text{if } d = d_0 
\end{cases}$$

where

$$L \triangleq 46.3 + 33.9 \times \log _{10}(f) - 13.82 \times \log _{10}(h_{AP})$$

$$- (1.1 \times \log _{10}(f) - 0.7) \times h_{UE} + (1.56 \times \log _{10}(f) - 0.8)$$

and where $PL(d)$ is the generalized notation in dB for $PL_{mk}$ or $PL_{mE}$. $d = d_{mk}$ (or $d = d_{mE}$) represents the distance between the $m$th AP and the $k$th user (or Eve). By the way, $d_0 = 10$ (m) and $d_1 = 50$ (m).

Moreover, the noise power can be given by

$$\sigma ^2 = B \times k_b \times T_0 \times \text{noise figure}$$

where $k_b$ is the Boltzmann constant, equal to $1.38 \times 10^{-23}$ (Joule/Kelvin), and the noise figure is 9 dB.

Figure 3 shows the cumulative distribution function (CDF) of the achievable secrecy rate by the $k_0$th user. We focus on the performance gain brought by multi-antenna APs deployment. When $N = 1$, it is the special case that all APs are single antenna in prior studies. Considering EPA for simplicity, no matter how many APs are deployed by single antenna or multiple antennas, the achievable secrecy rate is almost zero. With DPA, achievable secrecy rates are monotonically increasing with the antenna number of APs. This can be explained by the fact that when $N$ rises, the legitimate channels between APs and users are more favorable. 95%-likely performance7 is observable to analyze the level, which has a 95% chance to be achieved. Zooming in on the lower segment of the plot in Figure 3, the security rate of the $k_0$th user when each AP is set with 5 or 8 antennas deployment is about 2 bit/s/Hz, which is twice as high as that of only single antenna. It
also proves the conclusion by Irmer et al.\textsuperscript{9} that multi-antenna APs can inherit benefits from channel hardening.

As shown in Figure 4, we explore the intuitive influence on multi-antenna APs. As the amount of antennas by APs increases from 1 to 10, the achievable secrecy rate increases proportionally. With EPA, the number of antennas by APs and the number of legitimate users within the supported range, that is, $K_{\text{lt}}$, have no perceptible effect on the achievable secrecy rate. However, by applying DPA, the performance improvement is substantial. Even the number of legitimate users has reached its limit, that is, $K_{\text{lt}}$, which is the maximum interference from other legitimate users. From the growth trend, we can see the great promotion of achievable secrecy rate gains from single APs to the multi-antenna configuration. This is also the most challenging step for all network construction. In theory, random matrices are used to deduce the expected value for future simplification. In practical application, multi-antenna APs increase the multiplexing gain with moderate physical sizes.

In Figure 5, the achievable secrecy rate is plotted as a function of spoofing pilot power by Eve. If Eve cannot transmit training pilots, that is, $P_{\text{ue}} = 0$ W, the action will degenerate into passive eavesdropping. Between the active one and passive one, it is a significant decline of the achievable secrecy rate by EPA. But with our proposed DPA scheme, the performance gap is not such a great divergence, which indicates that power control is an effective method to restrain the harm of a pilot spoofing attack. When the power of Eve sending the training pilots is larger, the degree of pilot spoofing attack becomes more serious. Yet the damage to the target wiretapped users is not obvious. The reason is that DPA makes transmission power of each AP distributed to the legitimate users based on the channel conditions, which $g_{mk}$ is used as the performance metric. Particularly for the $k_{\text{th}}$ user, $g_{mk}$ suffers from the uplink power of Eve $P_{\text{ue}}$. If spoofing pilot power of Eve increases, channel estimation of the wiretapped user would be more inaccurate. This can lead to weaker estimated channel $g_{mk}$ and channel statistics $g_{mk}$, which makes the downlink transmission power received to Eve smaller. Therefore, DPA can exploit power control against Eve, which has strong spoofing pilot power.

Specifically, the achievable secrecy rate with single-antenna AP ranges from 1.769 bit/s/Hz with $P_{\text{ue}} = 0.1$ W to 1.675 bit/s/Hz with $P_{\text{ue}} = 1$ W. When the number of antennas per AP grows up to 5, the achievable secrecy rate value raises about 3.311 and 3.234 bit/s/Hz when pilot power of Eve increases 10 times from 0.1 W. In the same case with $N = 8$, the achievable secrecy rate shows a slighter downward trend from 3.996 to 3.862 bit/s/Hz. On one hand, the introduction of multi-antenna APs can eliminate some negative effects in the worst-case scenario in which Eve has strong power to send pilots. This means the beam-forming vector has to be more inclined toward the

| Symbol | Parameter meaning | Value |
|--------|-------------------|-------|
| $\tau_{\text{up}}$ | The length of pilot | 15 (samples) |
| $B$ | Bandwidth | 20 (MHz) |
| $f$ | Carrier frequency | $1.9 \times 10^3$ (MHz) |
| $h_{\text{AP}}$ | AP antenna height | 20 (m) |
| $h_{\text{UE}}$ | User/Eve antenna height | 1.5 (m) |
| $T_0$ | Noise temperature | 290 (Kelvin) |
| $\sigma_{\text{sh}}$ | The standard deviation of shadowing fading | 8 (dB) |
| $P_{\text{ue}}$ | Uplink pilot power of legitimate users | 0.6 (W) |
| $P_{\text{total}}$ | Total downlink power | 20 (W) |

AP: access point.

Figure 3. Cumulative distribution of achievable secrecy rate with $M = 100$, $K = 8$, and $P_{\text{ue}} = 0.2$ W.

Table 1. Setting of simulation parameters.

Wang et al.\textsuperscript{7}
direction of Eve. On the other hand, the power allocation scheme is of great assistance to the security performance enhancement. All APs only consider the channel condition of legitimate users to distribute themselves local power. Eve has no chance to change the received signal strength. During the downlink transmission phase, the identity of Eve is a passive eavesdropper. This DPA scheme transfers the difficult assignment from CPU to AP nodes. It can reduce the cell-free massive MIMO system workload and can handicap Eve’s position.

The achievable secrecy rate is depicted as a function of $M$ in terms of parameters with various amounts of antennas by APs shown in Figure 6. Under the worst-case security assumption, the network is coupled with a natural bias that Eve has more knowledge about CSI than legitimate users. Regardless of power control, there is terrible security performance whatever the number of antennas by APs or the number of APs by applying EPA. In our proposed DPA scheme, we generally demonstrate that a higher level of security performance can be gained from services provided by more APs. This also takes advantage of the typical narrow beam orientation of massive MIMO systems. As a result of each five antenna point intervals, the curves are not very smooth. We could still find that multi-antenna APs deployment is a bigger contributor to security performance improvement. When $N = 1$, the achievable secrecy rate climbs up from 0.950 to 2.143 bit/s/Hz. There are a dramatic rise of achievable secrecy rate in the case of $N = 5$ and $N = 8$, from 2.323 to 4.113 bit/s/Hz and from 2.897 to 4.630 bit/s/Hz between $M = 20$ and $M = 100$, respectively. Although densely random AP nodes in cell-free massive MIMO systems must have a closer geographical distance to Eve, the user-centric ideological concept can lead to significantly better security performance for legitimate users. Besides, in Figures 5 and 6, the simulation result is in keeping with the theoretical value to verify the correctness of secrecy performance analysis.

**Conclusion**

In this article, we have studied a cell-free massive MIMO network in the presence of an active eavesdropper, which is a smart attacker imitating the legitimate users’ training pilots and sending the same one to all APs. Combined with previous research, it has proven the channel hardening that exists when multiple antennas are configured at APs. Therefore, it is imperative to
consider the scenario of multi-antenna APs under cell-free massive MIMO networks, which is also the first time to focus on the security with arbitrary finite number of antennas deployed by APs. Taking MMSE estimation method and conjugate beamforming technology into accounts, we derive general closed-form expressions of the achievable rate at the legitimate user nodes, the information rate leaked into the eavesdropper, and the achievable secrecy rate. Moreover, a DPA scheme is proposed and contrasts with EPA. This power control algorithm could be used as a prototype for the user-centric network. Taking achievable secrecy rate as the performance security evaluation in simulations, our results reveal that it is superb to employ multi-antenna APs to enhance the security of wireless communication networks. That also corresponds with the application condition of the channel hardening mentioned above.

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 authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by China Postdoctoral Science Foundation (Grant No. 2021MD703980).

ORCID iD
Xiaoyu Wang @ https://orcid.org/0000-0001-7767-2145

References
1. Telatar IE. Capacity of multi-antenna Gaussian channels. *Eur Trans Telecommun* 1995, http://web.mit.edu/18.325/www/telatar_capacity.pdf
2. Foschini GJ and Gans MJ. On limits of wireless communication in a fading environment when using multiple antennas. *Wirel Person Commun* 1998; 6(3): 311–335.
3. Marzetta TL. Multi-cellular wireless with base stations employing unlimited numbers of antennas. In: *Proceedings of annual information theory & applications workshop UCSD*, February 2010, https://iwct.sjtu.edu.cn/Personal/zycchen/Massive%20MIMO.pdf
4. Marzetta TL. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans Wirel Commun* 2010; 9(11): 3590–3600.
5. Jungnickel V, Manolakis K, Zirwas W, et al. The role of small cells, coordinated multipoint, and massive MIMO in 5G. *IEEE Commun Mag* 2014; 52(5): 44–51.
6. Larsson EG, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems. *IEEE Commun Mag* 2014; 52(2): 186–195.
7. Ngo HQ, Ashikhmin A, Yang H, et al. Cell-free Massive MIMO versus small cells. *IEEE Trans Vehicul Technol* 2017; 66(3): 1834–1850.
8. Venkatesan S, Lozano A and Valenzuela R. Network MIMO: overcoming intercell interference in indoor wireless systems. In: *Proceedings of the 2007 conference record of the forty-first Asilomar conference on signals, systems and computers*, Pacific Grove, CA, 4–7 November 2007, pp.83–87. New York: IEEE.
9. Irmr R, Droste H, Marsch P, et al. Coordinated multi-point: concepts, performance, and field trial results. *IEEE Commun Mag* 2011; 49(2): 102–111.
10. You X, Wang D, Sheng B, et al. Cooperative distributed antenna systems for mobile communications [Coordinated and Distributed MIMO]. *IEEE Wirel Commun* 2010; 17(3): 35–43.
11. Nayebi E, Ashikhmin A, Marzetta TL, et al. Cell-free massive MIMO systems. In: *Proceedings of the 2015 49th Asilomar conference on signals, systems and computers*, Pacific Grove, CA, 8–11 November 2015, pp.695–699. New York: IEEE.
12. Chien TV, Bjornson E and Larsson EG. Joint power allocation and load balancing optimization for energy-efficient cell-free massive MIMO networks. *IEEE Trans Wirel Commun* 2020; 19(10): 6798–6812.
13. Chien TV, Ngo HQ, Chatzinotas S, et al. Reconfigurable intelligent surface-assisted cell-free massive MIMO systems over spatially-correlated channels, https://arxiv.org/pdf/2104.08648.pdf
14. Wu Y, Khisti A, Xiao C, et al. A survey of physical layer security techniques for 5G wireless networks and challenges ahead. *IEEE J Select Areas Commun* 2018; 36(4): 679–695.
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. Jose J, Ashikhmin A, Marzetta T, et al. Pilot contamination and precoding in multi-cell TDD systems. *IEEE Trans Wirel Commun* 2011; 10(8): 2640–2651.
17. Gutierrez-Estevez DM. Interference-aware flexible TDD design for massive MIMO 5G systems. In: *Proceedings of the 2017 IEEE wireless communications and networking conference*, San Francisco, CA, 19–22 March 2017. New York: IEEE.
18. Wu Y, Schober R, Ng DWK, et al. Secure massive MIMO transmission with an active eavesdropper. *IEEE Trans Inform Theor* 2016; 62(7): 3880–3900.
19. Lu X, Yang W, Guan X, et al. DCE-based secure transmission for massive MIMO relay system against active eavesdropper. *IEEE Trans Vehicular Technol* 2020; 69(11): 13045–13059.
20. Rawat DB, Neupane K and Song M. A novel algorithm for secrecy rate analysis in massive MIMO system with target SINR requirements. In: *Proceedings of the 2016 IEEE conference on computer communications workshops (INFOCOM WKSHPS)*, San Francisco, CA, 10–14 April 2016, pp.53–58. New York: IEEE.
21. 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.

22. Zhang X, Guo D, An K, et al. Secrecy analysis and active pilot spoofing attack detection for multigroup multicasting cell-free massive MIMO systems. *IEEE Access* 2019; 7: 57332–57340.

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

24. Alageli M, Ikhlef A, Alsifiany F, et al. Optimal downlink transmission for cell-free SWIPT massive MIMO systems with active eavesdropping. *IEEE Trans Inform Foren Sec* 2020; 15: 1983–1998.

25. Elhoushy S, Ibrahim M and Hamouda W. Exploiting RIS for limiting information leakage to active eavesdropper in cell-free massive MIMO. *IEEE Wirel Commun Lett* 2021; 11: 443–447.

26. Zhang X, Liang T and An K. Secrecy performance analysis of cell-free massive MIMO in the presence of active eavesdropper with low resolution ADCs. *Wirel Netw* 2021; 27: 4839–4852.

27. Chen Z and Bjoernson E. Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry. *IEEE Trans Commun* 2018; 66(11): 5205–5219.

28. Bjoernson E, Hoydis J and Sanguinetti L. Massive MIMO networks: spectral, energy, and hardware efficiency. *Found Trends Sig Process* 2017; 11(3–4): 154–655.

29. Marzetta TL, Larsson EG, Yang H, et al. *Fundamentals of massive MIMO*. Cambridge: Cambridge University Press, 2016.

30. Chen Z and Bjoernson E. Can we rely on channel hardening in cell-free massive MIMO? In: *Proceedings of the 2017 IEEE Globecom workshops*, Singapore City, Singapore, 4–8 December 2017, pp.1–6. New York: IEEE.

31. Vu TT, Ngo DT, Ngo HQ, et al. Full-duplex cell-free massive MIMO. In: *Proceedings of the 2019 IEEE international conference on communications*, Shanghai, China, 20–24 May 2019, pp.1–6. New York: IEEE.

32. Mai TC, Quoc Ngo H and Duong TQ. Cell-free massive MIMO systems with multi-antenna users. In: *Proceedings of the 2018 IEEE global conference on signal and information processing*, Anaheim, CA, 26–29 November 2018, pp.828–832. New York: IEEE.

33. Bjoernson E, Matthaiou M and Debbah M. Massive MIMO with non-ideal arbitrary arrays: hardware scaling laws and circuit-aware design. *IEEE Trans Wirel Commun* 2015; 14(8): 4353–4368.

34. Qiu J, Xu K, Xia X, et al. Downlink power optimization for cell-free massive MIMO over spatially correlated Rayleigh fading channels. *IEEE Access* 2020; 8: 56214–56227.

35. Pham QV, Le LB, Chung SH, et al. Mobile edge computing with wireless backhaul: joint task offloading and resource allocation. *IEEE Access* 2019; 7: 16444–16459.

36. Gao Z, Dai L, Mi D, et al. MmWave massive-MIMO-based wireless backhaul for the 5G ultradense network. *IEEE Trans Wirel Commun* 2015; 22(5): 13–21.

37. Rappaport TS. *Wireless communications: principles and practice*. 2nd ed. Hoboken, NJ: Prentice-Hall Press, 2002.

**Appendix I**

**Proof of equation (14)**

By substituting the concrete expression of $a_{kk}$ into equation (14), we can obtain

$$\gamma_k = \frac{\left(\sum_{l=1}^{M} \sqrt{\alpha_{ll}} w_{ll} \right)^2 + a^2}{\sum_{l=1}^{M} \sqrt{\alpha_{ll}} w_{ll}} (24)$$

Initially, we handle $\omega_{mk}$ into the closed form of $\hat{g}_{mk}/\sqrt{tr(\gamma_{mk})}$ through operating

$$E\left\{\left(\hat{g}_{mk}^H \hat{g}_{mk}\right)^*\right\} = E\left\{tr(\hat{g}_{mk}^H \hat{g}_{mk})\right\} = \left(\gamma_{mk}\right) = N\lambda_{mk}$$

Second, we solve the numerator term in equation (24) by calculating

$$E\left\{\hat{g}_{mk}^H \hat{g}_{mk}\right\} = \left\{\alpha \right\} E\left\{\hat{g}_{mk}^H + \hat{g}_{mk}^H \hat{g}_{mk}\right\}$$

$$= \left\{\alpha \right\} E\left\{\hat{g}_{mk}^H \hat{g}_{mk} + \hat{g}_{mk}^H \hat{g}_{mk} \right\}$$

$$= \left\{\alpha \right\} E\left\{\hat{g}_{mk}^H \hat{g}_{mk}\right\}$$

$$= tr(\gamma_{mk}) = N\lambda_{mk}$$

where $\alpha$ is achieved by $\hat{g}_{mk} = \hat{g}_{mk} + \hat{g}_{mk}$ with $\hat{g}_{mk}$ being the channel estimation error between the $m$th AP and the $k$th user. $\left\{\alpha \right\}$ is the result of independence principle between $\hat{g}_{mk}$ and $\hat{g}_{mk}$, attributed to MMSE estimator. It has key implications of $\hat{g}_{mk} \sim CN(0, \beta_{mk} I_N - \gamma_{mk})$.

Furthermore, in order to obtain the explicit expression of the first denominator term in equation (24), we continue to calculate

$$E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk} - \left|\hat{g}_{mk}^H \hat{g}_{mk}\right|\right|^2\right\} = E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk} - tr(\gamma_{mk})\right|^2\right\}$$

$$= E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|^2 + tr(\gamma_{mk}) E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|^2\right\} - 2tr(\gamma_{mk}) E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|\right\}\right\}$$

$$= E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|^2\right\} - |tr(\gamma_{mk})|^2$$

$$= \left\{\alpha \right\} E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|^2\right\}$$

$$= \left\{\alpha \right\} E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|^2\right\} - |tr(\gamma_{mk})|^2$$

$$= \left\{\alpha \right\} E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|^2\right\} - |tr(\gamma_{mk})|^2$$

$$= \left\{\alpha \right\} tr(\gamma_{mk}) + E\left\{\left|\hat{g}_{mk}^H \hat{g}_{mk}\right|^2\right\} - |tr(\gamma_{mk})|^2$$

$$= \left\{\alpha \right\} tr(\gamma_{mk}) + tr((\beta_{mk} I_N - \gamma_{mk})\gamma_{mk})$$

$$= \beta_{mk} tr(\gamma_{mk}) = N\beta_{mk}\lambda_{mk}$$

where $\alpha$ is acquired based on Lemma 1. For the second denominator term in equation (24), we need to make a calculation in the case of $k' \neq k$,.
Substituting equations (26), (27), and (28) into equation (24), we can obtain equation (14).

**Proof of equation (16)**

We first rewrite the SINR of Eve as

\[
\gamma_E = \frac{E\left\{ \sum_{m=1}^{M} \frac{P_{mk}}{r(\gamma_{mk})} \hat{g}_{mk}^H \hat{g}_{mk} \right\}^2}{\sum_{m=1}^{M} \sum_{n=1}^{M} \sum_{m \neq n}^{M} \frac{P_{mk}}{r(\gamma_{mk})} E\left\{ \hat{g}_{mk}^H \hat{g}_{mk} \right\} + \sigma^2}
\]  

(29)

And then, the numerator term of equation (29) can be figured out by

\[
E\left\{ \sum_{m=1}^{M} \frac{P_{mk}}{r(\gamma_{mk})} \hat{g}_{mk}^H \hat{g}_{mk} \right\} = \sum_{m=1}^{M} \frac{P_{mk}}{r(\gamma_{mk})} E\left\{ \hat{g}_{mk}^H \hat{g}_{mk} \right\}
\]

\[
+ E\left\{ \sum_{m=1}^{M} \frac{P_{mk}}{r(\gamma_{mk})} \hat{g}_{mk}^H \sum_{n \neq m}^{M} \frac{P_{nk}}{r(\gamma_{nk})} \hat{g}_{nk} \hat{g}_{mk} \right\}
\]

\[
= \sum_{m=1}^{M} \frac{P_{mk}}{r(\gamma_{mk})} \left( E\left\{ \hat{g}_{mk}^H \hat{g}_{mk} \right\} + E\left\{ \hat{g}_{mk}^H \hat{g}_{mk} \right\} \right)
\]

\[
+ E\left\{ \sum_{m=1}^{M} \sum_{n \neq m}^{M} \frac{P_{mk}P_{nk}}{r(\gamma_{mk})r(\gamma_{nk})} \hat{g}_{mk}^H \hat{g}_{mk} \hat{g}_{nk} \right\}
\]

\[
= \sum_{m=1}^{M} \frac{P_{mk}}{r(\gamma_{mk})} \left( P_{mk} \hat{g}_{mk}^H \hat{g}_{mk} + \sum_{n \neq m}^{M} \frac{P_{nk}}{r(\gamma_{nk})} \hat{g}_{mk}^H \hat{g}_{nk} \hat{g}_{mk} \right)
\]

\[
+ \sum_{m=1}^{M} \sum_{n \neq m}^{M} \frac{P_{mk}P_{nk}}{r(\gamma_{mk})r(\gamma_{nk})} E\left\{ \hat{g}_{mk}^H \hat{g}_{mk} \hat{g}_{nk} \right\}
\]

(30)

Third, we address the residual non-closed part of equation (29) through

\[
E\left\{ \hat{g}_{mk}^H \hat{g}_{mk} \right\} = E\left\{ \hat{g}_{mk}^H \hat{g}_{mk} \right\}
\]

\[
= E\left\{ (\hat{g}_{mk}^H \hat{g}_{mk})^2 \right\}
\]

(31)

Last, using equations (30) and (31), the derivation of \( \gamma_k \) is easily obtained.