Vulnerabilities of Massive MIMO Systems Against Pilot Contamination Attacks

Berk Akgun, Marwan Krunz, and O. Ozan Koyluoglu

Abstract

We consider a single-cell massive MIMO system in which a base station (BS) with a large number of antennas transmits simultaneously to several single-antenna users in the presence of an attacker. The BS acquires the channel state information (CSI) based on uplink pilot transmissions. In this work, we demonstrate the vulnerability of CSI estimation phase to malicious attacks. For that purpose, we study two attack models. In the first model, the attacker aims at minimizing the sum-rate of downlink transmissions by contaminating the uplink pilots. In the second model, the attacker exploits its in-band full-duplex capabilities to generate jamming signals in both the CSI estimation and data transmission phases. We study these attacks under two downlink power allocation strategies when the attacker knows and does not know the locations of the BS and users. The formulated problems are solved using stochastic optimization, Lagrangian minimization, and game-theoretic methods. A closed-form solution for a special case of the problem is obtained. Furthermore, we analyze the achievable individual secrecy rates under a pilot contamination attack, and provide an upper bound on these rates. Our results indicate that the proposed attacks degrade the throughput of a massive MIMO system by more than 50%.

Index terms

Massive MIMO, pilot contamination, physical layer security, active attack, stochastic optimization.

An abridged version of this paper will appear in the IEEE CNS 2017 Conference, Las Vegas, October 9-11, 2017.

This research was supported in part by the National Science Foundation (grants # CNS-1409172, CNS-1513649, IIP-1265960, and CNS-1617335) and by Qatar Foundation (grant # NPRP 8-052-2-029). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the NSF and QF.

The authors are with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721 USA (email: berkakgun@email.arizona.edu, ozan@email.arizona.edu, krunz@email.arizona.edu).

I. Introduction

Massive multiple-input multiple-output (MIMO) is one of the key technologies in the upcoming 5G systems. It is envisioned that a cellular base station (BS) in 5G systems will be equipped with a very large antenna array, e.g., hundreds of antennas or more, boosting the transmission rate by orders of magnitude compared to conventional MIMO systems. Even though MIMO is a well-studied concept in wireless communications, massive MIMO requires novel techniques to overcome new design challenges, and as such it has received significant attention from researchers over the last few years (see, for example, [1], [2], [3], and the references therein).

One of the important issues in massive MIMO systems is pilot contamination (PC) [4]. Because of the large number of antennas at the BS and the relatively short channel coherence time, the channel state information (CSI) between the BS and various users must be estimated frequently using uplink pilot transmissions. Assuming channel reciprocity, the BS utilizes these CSI estimates for downlink data transmissions. However, due to the limited number of orthogonal pilot sequences (e.g., in the order of tens [4]), users in neighboring cells may share the same pilots. Interference among these pilots causes erroneous CSI estimates at the BS, leading to poor system performance.

In [5] the authors studied an attack that exploits vulnerabilities of the channel training phase in time division duplexing (TDD) systems. The key idea behind this attack is to contaminate uplink pilot transmissions and cause an erroneous uplink channel estimation. Typically, if the CSI is available, the BS would use MIMO beamforming techniques such as maximum-ratio transmission (MRT) to maximize the signal-to-noise-ratio (SNR) at users. However, the benefits of these techniques vanish rapidly if the CSI estimates are erroneous. A self-contamination technique was proposed in [6] to detect this type of attack. The authors in [7] proposed another approach in which the legitimate user transmits four random phase-shift keying symbols, and the BS checks the correlation matrix of the received signals. Based on the ratio of two largest eigenvalues of this matrix, the BS detects the attack. Secure transmissions for TDD-based massive MIMO systems was studied in [8] in the presence of an active eavesdropper. The authors derived the optimal power allocation for the information and artificial noise (AN) signals at the BS such that secrecy is asymptotically guaranteed, i.e., as the number of BS antennas (M) tends to infinity. In [9], the authors proposed providing secrecy against PC attacks by keeping pilot assignments hidden and using a pilot set that scales with M. However, there are two main problems with this

scheme. First, it requires a longer pilot transmission phase, which increases the overhead and decreases the throughput. Second, computationally intensive cryptographic methods are required to keep pilot assignments hidden. All of the papers discussed above consider an attacker that targets one user at a time. Even when a multiuser system is in place, the attacker randomly selects a given user and contaminates its pilot sequence. Given that one of the key aspects of massive MIMO systems is to serve tens of users simultaneously, the vulnerabilities of these systems to a multiuser pilot contamination attack should be investigated.

In this paper, we consider a single-cell multiuser massive MIMO network in the presence of an attacker. We study two attack models. In the first model, the attacker aims at minimizing the sum-rate of downlink transmissions by contaminating uplink pilot transmissions. We derive the downlink transmission rates with and without the PC attack by exploiting the *channel hardening effect* (effect of small-scale fading on channel gains vanishes as M tends to infinity) in massive MIMO. Optimal attack strategies are then investigated for two different cases: when the attacker knows the locations of the BS and users and when she does not have this information. Considering a fixed power allocation strategy for downlink data transmissions, convex problems are formulated for the optimal PC attack. These problems are solved via the interior-point and Lagrangian minimization methods. We obtain a closed-form solution for the case of perfect information, i.e., known topology at the attacker. This solution represents a lower bound on the downlink sum-rate of massive MIMO systems under an optimal PC attack and a fixed BS transmission power. Then, we study the scenario where the BS optimizes its own power allocation scheme in the presence of PC attacks. For this case, a game-theoretic problem formulation is considered in which the BS and attacker are the players of the game. In particular, we obtain a *convex-concave* game, and propose an iterative algorithm that converges to a Nash equilibrium (NE) of the game. This analysis provides an upper bound on the downlink sum-rate of massive MIMO systems under an optimal PC attack.

For the second attack model, the attacker generates jamming signals in both the pilot and downlink data transmission phases. For this hybrid attack, the attacker is required to have a full-duplex radio. Specifically, the attacker estimates the channels between users and itself while jamming the uplink pilot transmissions. These estimates are then used to strengthen the attack during the downlink data transmission phase. Stochastic optimization techniques are used to find the optimal power allocation at the attacker so as to minimize the downlink sum-rate of the system.

Massive MIMO systems are robust against passive eavesdropping, as the CSI at a legitimate receiver and an eavesdropper are near-orthogonal [7]. However, these systems are vulnerable to an active attacker that contaminates the uplink pilot transmissions. Therefore, we extend our work in [10] and analyze the secrecy performance of a massive MIMO system under a PC attack. Specifically, an attacker receives the information signals intended to users with a much higher signal power by using the PC attack. We study a problem where the attacker minimizes the maximum of the achievable individual secrecy rates at users. Our analysis provides an upper bound on the achievable individual secrecy rates in a given massive MIMO system under the PC attack. Moreover, by introducing chance constraints, we study the case where the attacker does not know the locations of users. The formulated problems are numerically solved by an iterative method. Numerical results show that the downlink sum-rate decreases significantly under a PC attack. Particularly, when the attacker is close to the BS, the downlink sum-rate of all users is reduced by more than 50%. Another important result of our work is that an attacker without perfect information about user locations is almost as devastating as one with perfect information. This fact emphasizes the vulnerability of massive MIMO systems to PC attacks. Further, we observe that even if the attacker moves farther from the BS, the maximum per-user secrecy rate is reduced by almost 30%.

The rest of the paper is organized as follows. Section II describes the system model. In Section III, we compute the downlink transmission rates with/without a PC attack. Our PC attack under a fixed and optimal BS transmission power is studied in Section IV. We analyze the secrecy rates of the users in massive MIMO systems in Section V. Section VI investigates the hybrid attack model. We provide numerical results in Section VII, and conclude the paper in Section VIII.

Throughout the paper, we adopt the following notation. $\mathbb{E}[\cdot]$ indicates the expectation of a random variable. Row vectors and matrices are denoted by bold lower-case and upper-case letters, respectively. $(\cdot)^*$ and $(\cdot)^T$ represent the complex conjugate transpose and transpose of a vector or matrix, respectively. Frobenius norm and the absolute value of a real or complex number are denoted by $\|\cdot\|$ and $|\cdot|$, respectively. $\mathbf{A} \in \mathbb{C}^{M\times N}$ means that $\mathbf{A}$ is an $M \times N$ complex matrix, and $\mathbf{I}_M$ is an $M \times M$ identity matrix. $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex circularly symmetric Gaussian random variable of mean μ and variance σ^2. $[x]^+$ is defined as $\max(x, 0)$. For simplicity, $\log_2(\cdot)$ is referred to as $\log(\cdot)$.

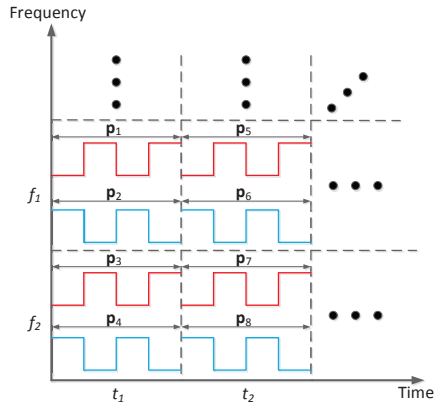

Fig. 1. Orthogonality of pilot sequences in space-time-frequency domain.

II. System Model

We consider a single-cell massive MIMO system in which the BS (Alice) uses a large array of M antenna elements to transmit/receive independent data streams to/from K single-antenna users (Bobs), where $M \gg K$. Because of the large M, the channel coherence time is not long enough to estimate the CSI of all M downlink channels at each user [3]. Therefore, TDD is used instead of FDD (in FDD, the downlink and uplink channels are estimated separately). In TDD, Alice estimates the CSI for uplink channels after receiving pilot sequences transmitted by Bobs. If these pilot symbols are not perfectly orthogonal to each other, interference among them causes erroneous channel estimates at the BS. Assuming channel reciprocity, these estimates are used in setting the precoding matrices of downlink data transmissions. There is no standardization for massive MIMO systems regarding the orthogonality of the pilot sequences. However, the authors in [4] suggested assigning an orthogonal time-frequency pilot sequence to each Bob. Orthogonal space-time block codes can also be utilized, as in 802.11ac systems, to increase the number of orthogonal pilot sequences. Fig. 1 shows an example of eight pilot sequences. Pilot sequences $\mathbf{p}_1$ and $\mathbf{p}_2$ are orthogonal space-time coded sequences. They are sent in the same time interval (t_1) over the same frequency (f_1) by two different Bobs. On the other hand, the orthogonality of $\mathbf{p}_1$ and $\mathbf{p}_5$ is guaranteed by transmitting them in different time intervals t_1 and t_2, e.g., $\mathbf{p}_1 = 0$ during t_2. Similarly, $\mathbf{p}_1$ and $\mathbf{p}_3$ are transmitted on different frequencies f_1 and f_2, respectively.

The received signal at Alice during the pilot transmission phase is given by:

$$\mathbf{Y}_A = \sum_{i=1}^{K} \sqrt{P_k}\mathbf{h}_k^T \mathbf{p}_k + \mathbf{W} \tag{1}$$

where $\mathbf{h}_k^T \in \mathbb{C}^{M\times 1}$ represents the uplink channel from Bob$_k$ (kth Bob) to Alice. The mth entry, $m \in \{1, \cdots, M\}$, of this vector is given by $h_k^{(m)} = \sqrt{\theta_k} g_k^{(m)}$, where θ_k and $g_k^{(m)} \sim \mathcal{CN}(0,1)$ represent the path-loss component (large-scale fading) and small-scale effects of the channel (Rayleigh fading), respectively. Note that θ_k is the same for all antennas, so $\mathbf{h}_k$ can be written as $\mathbf{h}_k = \sqrt{\theta_k}\mathbf{g}_k$, where $\mathbf{g}_k$ is a vector of all $g_k^{(m)}$, $m \in \{1, \cdots, M\}$. $\mathbf{p}_k \in \mathbb{C}^{1\times L}$ is the transmitted pilot sequence by Bob$_k$, where L is the number of symbols in the pilot sequence. As these pilot sequences are orthogonal to each other, $\mathbf{p}_k\mathbf{p}_l^* = 0 \ \forall \ k$ and $l \in \mathcal{K}$, where $k \neq l$ and $\mathcal{K} = \{1, \cdots, K\}$. We assume $\mathbf{p}_k$ is a unit vector (i.e., $\mathbf{p}_k\mathbf{p}_k^* = 1 \ \forall k \in \mathcal{K}$). P_k is the pilot transmission power at Bob$_k$. $\mathbf{W}$ is the additive white Gaussian noise (AWGN) matrix, whose entries are zero-mean, unit-variance normal random variables.

Without loss of generality, consider the estimation of $\mathbf{h}_i$ at Alice. Let $\hat{\mathbf{h}}_i$ represent this estimate. Under a priori knowledge of $\mathbf{p}_i$, Alice post-multiplies the received signal by $\mathbf{p}_i^*$ and divides it by $\sqrt{P_i}$ and L to obtain:

$$\begin{aligned}\hat{\mathbf{h}}_i^T &= \frac{\mathbf{Y}_A\mathbf{p}_i^*}{\sqrt{P_i}L} = \sum_{k=1}^{K} \frac{\sqrt{P_k}\mathbf{h}_k^T\mathbf{p}_k\mathbf{p}_i^*}{\sqrt{P_iL}} + \frac{\mathbf{W}\mathbf{p}_i^*}{\sqrt{P_i}L} \\ &= \mathbf{h}_i^T + \tilde{\mathbf{w}}_i^T \end{aligned} \tag{2}$$

where $\tilde{\mathbf{w}}_i^T \triangleq \frac{\mathbf{W}\mathbf{p}_i^*}{\sqrt{P_iL}} \sim \mathcal{CN}(0, \frac{1}{P_iL}\mathbf{I}_M)$.

A. PC Attack Model

We now describe the first attack model considered in this paper. The attacker aims to contaminate pilot transmissions by imposing his signal. We assume that the attacker knows the pilot sequences used by Bobs (generally, pilot sequences are publicly known). Because the number of orthogonal pilots is limited, after eavesdropping on the channels for some time, the attacker can learn the pilots assigned to various Bobs. Let $\mathbf{x}_J \in \mathbb{C}^{1\times L}$ be the signal generated by the attacker. After the attack, the received signal at Alice will be modified as follows:

$$\mathbf{Y}_A = \sum_{k=1}^{K} \sqrt{P_k}\mathbf{h}_k^T \mathbf{p}_k + \mathbf{h}_J^T\mathbf{x}_J + \mathbf{W} \tag{3}$$

where $\mathbf{h}_J^T \in \mathbb{C}^{M\times 1}$ represents the channel vector from the attacker to Alice. In the literature, $\mathbf{x}_J$ is often designed such that only a single user is targeted by the attacker [5], [9]. This user is

selected randomly without any optimization. More specifically, $\mathbf{x}_J$ is often set to $\sqrt{P_J}\mathbf{p}_k$, where P_J is the average jamming power. In contrast, in our model, we set $\mathbf{x}_J$ to:

$$\mathbf{x}_J = \sqrt{P_J}\sum_{k=1}^{K}\sqrt{\alpha_k}\mathbf{p}_k \tag{4}$$

where α_k is the ratio between the power that the attacker allocates to pilot $\mathbf{p}_k$ and the average jamming power. Note that $\sum_{k=1}^{K}\alpha_k \leq 1$. The objective of the attacker is to minimize the downlink sum-rate. Let R_k be the downlink transmission rate at Bob$_k$. The attacker's goal can be formulated as follows:

$$\underset{\{\alpha_k\ \forall k\in\mathcal{K}\}}{\text{minimize}}\sum_{k\in\mathcal{K}} R_k \tag{5}$$

subject to $\alpha_k \geq 0\ \forall k \in \mathcal{K}$ and $\sum_{k=1}^{K}\alpha_k \leq 1$.

III. Downlink Transmission Rates

In this section, we analyze the downlink sum-rate in the underlying massive MIMO system with/without the aforementioned PC attack.

A. No PC Attack Scenario

In massive MIMO systems, the BS often applies MRT precoder [1]–[3], [11]. For conventional MIMO systems, MRT gives rise to inter-user interference. However, as M tends to infinity, the channels between the BS and individual users become orthogonal to each other, and they individually reduce to single-input single-output (SISO) channels. In this case, MRT is the optimal precoder. Let s_k be the downlink information signal intended to Bob$_k$ $\forall k \in \mathcal{K}$, and let $\mathbf{v}_k^T \in \mathbb{C}^{M\times 1}$ be its normalized precoder, i.e., $\mathbf{v}_k\mathbf{v}_k^* = 1$. The received signal at Bob$_k$ in the downlink data transmission phase is given by:

$$y_k = \sum_{i=1}^{K}\sqrt{P_i^{(d)}}\mathbf{h}_k\mathbf{v}_i^T s_i + w_k^{(d)} \tag{6}$$

where $P_k^{(d)}$ and $w_k^{(d)}$ are, respectively, the allocated power to s_k at Alice and the AWGN with zero-mean and unit-variance at Bob$_k$. Employing MRT precoding, $\mathbf{v}_k^T$ is given by $\mathbf{v}_k^T = (\hat{\mathbf{h}}_k^*/\|\hat{\mathbf{h}}_k\|)$. The achievable downlink rate at Bob$_k$ becomes:

$$R_k = \log\left(1 + \frac{P_k^{(d)}|\mathbf{h}_k\mathbf{v}_k^T|^2}{\sum_{l\in\{\mathcal{K}\setminus k\}} P_l^{(d)}|\mathbf{h}_k\mathbf{v}_l^T|^2 + 1}\right),\ k \in \mathcal{K}. \tag{7}$$

Note that the precoding vectors are computed based on channel estimates.

Next, we study the asymptotic behavior of R_k as $M \to \infty$. Such asymptotic analysis is needed later on for comparison with the case under a PC attack. Consider the inter-user interference term $P_l^{(d)}|\mathbf{h}_k\mathbf{v}_l^T|^2$ in (7). Scaling this term by M and taking the limit as $M \to \infty$, we end up with:

$$\lim_{M\to\infty} \frac{P_l^{(d)}|\mathbf{h}_k\mathbf{v}_l^T|^2}{M} = 0 \tag{8}$$

$\forall k$ and $l \in \mathcal{K}$, where $k \neq l$ (see Appendix A for the proof). The underlying intuition behind this result is that entries of small-scale channel components of Bob$_k$ and Bob$_l$ are independent random variables of zero-mean and unit-variance. Hence, $\lim_{M\to\infty} \mathbf{g}_l\mathbf{g}_k^*/M = 0$. Similarly, $\lim_{M\to\infty} \mathbf{g}_l\tilde{\mathbf{w}}_k^*/M = 0$. This is a result of the channel orthogonality in massive MIMO systems. On the other hand, for the term in the numerator in (7), we have

$$\lim_{M\to\infty} \frac{P_k^{(d)}|\mathbf{h}_k\mathbf{v}_k^T|^2}{M} = \frac{P_k^{(d)}\theta_k^2}{\theta_k + \frac{1}{P_k L}} > 0 \tag{9}$$

(see Appendix B for the proof). Hence, the downlink rate at Bob$_k$ behaves asymptotically as:

$$R_k \sim \log\left(1 + \frac{P_k^{(d)}\theta_k^2}{(\theta_k + \frac{1}{P_k L})\frac{1}{M}}\right). \tag{10}$$

In our paper, we consider a finite but sufficiently large M, with $M \gg K$, so the channels are near-orthogonal. As a result, the inter-user interference can be neglected as in (8). Moreover, for a sufficiently large M, $|\mathbf{h}_k\mathbf{v}_k^T|^2/M$ approaches the result in (9) ([3], [4], [11]). (In Section VII, we numerically verify these results.) As explained before, θ_k is the large-scale channel components at Bob$_k$. Equation (10) indicates that the SINR does not depend on the small-scale fading components, as they are averaged out by the large antenna array (channel hardening). The term $(1/M)$ in the equation comes from the AWGN $w_k^{(d)}$ at Bob$_k$. For example, as $M \to \infty$, the noise term vanishes and the SINR tends to infinity. Another noise term arises due to the channel estimation error $\tilde{\mathbf{w}}_i$. For example, as the length of the pilots, L, increases, the second term in the denominator becomes smaller. This leads to an increase in the downlink rate. The same effect is also observed when the power allocated for pilots increases.

In this paper, we consider two different downlink transmit power allocation strategies at Alice: "fixed" and "optimal". Both strategies are subject to an average power constraint P_A. Under the fixed power allocation, $P_k^{(d)}$ $\forall k \in \mathcal{K}$ is known to the attacker. For example, based on a fairness criterion, these values may be determined before the pilot transmission phase (e.g.,

when Bobs are registered with the network), and Alice may convey this information to Bobs through a feedback channel. If the attacker eavesdrops on this channel, she can obtain the power allocation values. In an instance of this setup, Alice may simply allocate powers uniformly to the information signals, i.e., $P_1^{(d)} = \cdots = P_K^{(d)} = P_A/K$. On the other hand, under the "optimal" power allocation strategy, Alice relies on the well-known water-filling technique to assign powers, using $(\theta_k + (P_k L)^{-1})/(M\theta_k^2)$ as the water levels (see, e.g., [12]).

B. Presence of PC Attack

Under the attack model in (4), the following channel estimation is performed at Alice for each Bob_k:

$$\hat{\mathbf{h}}_k = \mathbf{h}_k + \sqrt{\alpha_k u_k}\mathbf{h}_J + \tilde{\mathbf{w}}_k \tag{11}$$

where u_k is the ratio between the average power at the attacker and the pilot transmission power at Bob_k, i.e., $u_k = P_J/P_k$. In the rest of the paper, we assume that u_k is known to the attacker. Previously, we assumed that the attacker learns the pilot sequences by eavesdropping on the uplink transmissions. The attacker can similarly learn the pilot transmission power. Also note that these transmission powers are fixed in the current cellular systems. Alice is not aware of the presence of the attacker, so she treats $\hat{\mathbf{h}}_k$ as the correct channel estimate. Employing MRT precoding based on this estimation, Alice computes the precoder vector of s_k as:

$$\mathbf{v}_k^T = \frac{(\mathbf{h}_k + \sqrt{\alpha_k u_k}\mathbf{h}_J + \tilde{\mathbf{w}}_k)^*}{\|\mathbf{h}_k + \sqrt{\alpha_k u_k}\mathbf{h}_J + \tilde{\mathbf{w}}_k\|}. \tag{12}$$

Substituting this precoder vector in (7), the attacker's optimization problem in (5) becomes non-convex. To obtain a tractable attack model, we analyze the asymptotic behavior of R_k as $M \to \infty$. Following the same steps as in the case of no attacker, the following expression can be obtained as $M \to \infty$:

$$R_k = \log\left(1 + \frac{P_k^{(d)} M\theta_k^2}{\theta_k + \alpha_k u_k \theta_J + \frac{1}{P_k L}}\right). \tag{13}$$

As M increases, the massive MIMO system becomes more resilient to PC attacks. However, the vulnerability of the system against such an attack can be observed in (13), which shows that the SINR decreases with an increase in the jamming power $\alpha_k u_k$, with the functional form as given therein.

As in the previous section, a fixed or "optimal" power allocation strategy can be employed to calculate each $P_k^{(d)}$. Fixed power allocation is performed exactly as before, whereas "optimal"

power allocation corresponds to the following strategy. Let $\phi_k \triangleq \theta_k + \alpha_k u_k \theta_J + \frac{1}{P_k L}$. Then, Alice tries to maximize $R_{\text{sum}} = \sum_{k=1}^{K} R_k$ to obtain the "optimal" power allocation vector:

$$\left[P_1^{(d)} \cdots P_K^{(d)}\right] = \underset{x_k, \forall k \in \mathcal{K}}{\text{argmax}} \sum_{k=1}^{K} \log\left(1 + \frac{x_k M \theta_k^2}{\phi_k}\right) \tag{14}$$

subject to $\sum_{k=1}^{K} P_k^{(d)} \leq P_A$ and $P_k^{(d)} \geq 0$, $\forall k \in \mathcal{K}$. Because Alice is unaware of the attack, she will not necessarily solve the above problem. However, our goal is to observe the effect of PC attack, even if Alice employs the least favorable power allocation scheme from the perspective of the attacker. This way, we can establish an upper-bound on the downlink sum-rate under a PC attack.

IV. Analysis of Optimal PC Attack

A. Fixed Power Allocation at Alice

In this section, we study the optimal PC attack strategy. Our analysis provides a lower bound on the downlink sum-rate under a PC attack for a given power allocation at Alice. We incorporate (13) into problem (5), considering fixed power allocation for the information signals at Alice:

$$\begin{aligned} \textbf{P1} : \underset{\{\alpha_k,\, \forall k \in \mathcal{K}\}}{\text{minimize}} & \sum_{k=1}^{K} \log\left(1 + \frac{P_k^{(d)} M \theta_k^2}{\theta_k + \alpha_k u_k \theta_J + \frac{1}{P_k L}}\right) \\ s.t. \qquad & \alpha_k \geq 0 \; \forall k \in \mathcal{K}, \quad \sum_{k=1}^{K} \alpha_k \leq 1. \end{aligned}$$

For a given $k \in \mathcal{K}$, we assume that $\theta_k = A z_k^{-\gamma}$, where A is a constant that depends on the operating frequency, transmit and receive antennas, while γ and z_k are the path-loss exponent and the distance between Alice and Bob$_k$, respectively. Similarly, z_J is the distance between Alice and the attacker. For simplicity, the antennas at Bobs and the attacker are assumed to be identical, so the same A is considered for all of them. As a result, the objective function of **P1** is converted to the following one:

$$R_{\text{sum}} = \sum_{k=1}^{K} \log\left(1 + \frac{P_k^{(d)} M A z_k^{-2\gamma}}{\alpha_k u_k z_J^{-\gamma} + z_k^{-\gamma} + \frac{1}{A P_k L}}\right) \tag{15}$$

Next, we discuss two different scenarios based on the information available to the attacker.

1) Perfect Information: Here, we assume that the attacker has perfect knowledge of the distances between Alice and individual Bobs as well as her own distance to Alice. Indeed, this is an idealized scenario (from the attacker's point of view), and is merely studied to provide a benchmark for comparison with the case of uncertainty in distances. **P1** is a convex programming problem, and we obtain the optimal solution as follows.

Theorem 1: **P1** has the following closed-form solution:

$$\alpha_k = \left[\frac{\sqrt{A_k(A_k + 4/\lambda)} - A_k - 2B_k}{2} \right]^+ \quad \forall k \in \mathcal{K} \tag{16}$$

where

$$A_k \triangleq \frac{P_k^{(d)} M A z_J^{\gamma}}{u_k z_k^{2\gamma}} \text{ and } B_k \triangleq \frac{z_J^{\gamma}}{u_k z_k^{\gamma}} + \frac{z_J^{\gamma}}{u_k A P_k L}.$$

λ is the *Karush-Kuhn-Tucker* (KKT) multiplier and is chosen such that $\sum_{k=1}^{K} \alpha_k = 1$. It can be easily computed by the *bisection* method as $\sum_{k=1}^{K} \alpha_k$ is a decreasing function of it.

Proof: See Appendix C. ■

2) Uncertainty in Distances: Suppose that the attacker does not have perfect knowledge about various distances. Let Z_k and Z_J be random variables (rvs) that correspond to the Alice-Bob$_k$ and Alice-attacker distances, respectively. In this case, the expected value of R_{sum} is given by:

$$\begin{aligned} \mathbb{E}[\, R_{\text{sum}} \,] &= \mathbb{E}\left[\sum_{k=1}^{K} \log\left(1 + \frac{P_k^{(d)} M A Z_k^{-2\gamma}}{\alpha_k u_k Z_J^{-\gamma} + Z_k^{-\gamma} + \frac{1}{A P_k L}} \right) \right] \\ &= \sum_{k=1}^{K} \mathbb{E}\left[\log\left(1 + \frac{P_k^{(d)} M A Z^{-2\gamma}}{\alpha_k u_k Z_J^{-\gamma} + Z^{-\gamma} + \frac{1}{A P_k L}} \right) \right] \end{aligned} \tag{17}$$

where Z is a generic rv that has the same distribution as Z_k for all k. In (17), the expectation is taken over Z and Z_J. The last equality follows from the assumption that the distributions of the distances between individual Bobs and Alice are identical. We further assume that Bobs and the attacker are randomly and uniformly located in a circular ring around Alice. Let $D_{\min}$ and $D_{\max}$ be the minimum and maximum possible distances between Alice and any Bob, respectively. Hence, the CDF of Z is given by $\Pr[Z \leq x] = (x^2 - D_{\min}^2)/(D_{\max}^2 - D_{\min}^2)$ where $x \in [D_{\min}, D_{\max}]$. Accordingly, the PDF of Z is given by $f_Z(x) = 2x/(D_{\max}^2 - D_{\min}^2)$, for $x \in [D_{\min}, D_{\max}]$.

Let $\Phi_k \triangleq \alpha_k u_k Z_J^{-\gamma} + Z^{-\gamma} + \frac{1}{A P_k L}$. Under fixed downlink power allocation, the optimal PC attack can be formulated by the following stochastic programming problem:

$$\textbf{P2} : \underset{\{\alpha_k \ \forall k \in \mathcal{K}\}}{\text{minimize}} \sum_{k=1}^{K} \mathbb{E}\left[\log\left(1 + \frac{P_k^{(d)} M A Z^{-2\gamma}}{\Phi_k} \right) \right]$$

$$s.t. \qquad \alpha_k \geq 0 \; \forall k \in \mathcal{K}, \quad \sum_{k=1}^{K} \alpha_k \leq 1.$$

The objective function in **P2** can be rewritten as:

$$\sum_{k=1}^{K} \int_{D_{\min}}^{D_{\max}} \int_{D_{\min}}^{D_{\max}} \frac{4xy}{(D_{\max}^2 - D_{\min}^2)^2} \log(\Psi(x,y))\, dx\, dy \tag{18}$$

where

$$\Psi(x,y) \triangleq 1 + \frac{P_k^{(d)} M A x^{-2\gamma}}{\alpha_k u_k y^{-\gamma} + x^{-\gamma} + \frac{1}{A P_k L}}$$

for $x, y \in [D_{\min}, D_{\max}]$. This is a convex programming problem, as the objective function and inequality constraints are all convex functions. The integral in (18) can be approximated by Simpson's Rule for double integrals, and can be solved efficiently by applying the interior point method. Note that **P2** need only be solved offline, so the time complexity of this solution method is not a concern. We also note that although we only study a uniform distribution for the locations of Bobs and the attacker, any arbitrary distribution can be considered. The integral operation preserves the convexity, so the same steps can be followed to solve **P2**. Our numerical results (not shown for brevity) indicate that for typical values of P_A, A, K, and $D_{\max}$, the attacker should target all Bobs by equally allocating its average power to various pilot sequences under uniform power allocation when $u_k = u_l \; \forall k, l \in \mathcal{K}$. That is, $\alpha_k = P_J/K \; \forall k \in \mathcal{K}$. This is due to the symmetry of Bobs for this special case, as will be discussed in Section VII.

3) Discussion: Let $\mathbf{z} \triangleq [z_1, \cdots, z_K]$ be the vector of distances from Alice to various Bobs (known to the attacker). Let $\boldsymbol{\alpha}^*(\mathbf{z}, z_J) \triangleq [\alpha_1^*(\mathbf{z}, z_J), \cdots, \alpha_K^*(\mathbf{z}, z_J)]$ and $\boldsymbol{\alpha}^* \triangleq [\alpha_1^*, \cdots, \alpha_K^*]$ be the optimal solutions to **P1** and **P2**, respectively. In this case, the objective function of **P2** becomes $\mathbb{E}_{\mathbf{Z}, Z_J}[R_{\text{sum}}(\boldsymbol{\alpha}^*)]$, and $\mathbb{E}_{\mathbf{Z}, Z_J}[R_{\text{sum}}(\boldsymbol{\alpha}^*(\mathbf{Z}, Z_J))]$ becomes the expectation of the optimal solution of **P1** under perfect information, where $\mathbf{Z}$ is the vector of i.i.d. distances $Z_1, \cdots, Z_K$. The expectations are taken over the random distances, as previously explained. The *expected value of perfect information* (EVPI) is defined as follows:

$$\text{EVPI} \triangleq \mathbb{E}_{\mathbf{Z}, Z_J}[R_{\text{sum}}(\boldsymbol{\alpha}^*)] - \mathbb{E}_{\mathbf{Z}, Z_J}[R_{\text{sum}}(\boldsymbol{\alpha}^*(\mathbf{Z}, Z_J))]. \tag{19}$$

Note that EVPI is always greater than or equal to zero, as the case with perfect information outperforms the one with uncertainty. If EVPI is small, the attacker does not gain much by knowing the exact distances. It can perform attacks almost as powerful as when perfect information is available. On the other hand, if EVPI is high, the attacker may try to acquire distance information by estimating Bobs' locations relative to its own. For example, a group of colluding adversaries

can employ localization techniques (e.g., RSSI and time-of-arrival) to estimate Alice-to-Bobs distances [13], [14]. This requires more complex and costly systems at the attacker. In Section VII, we study the behavior of EVPI.

B. Optimal Power Allocation

In this section, we derive the optimal PC attack strategy when Alice adopts optimal (the least favorable from the perspective of the attacker) power allocation strategy for downlink data transmissions. Note that Alice is assumed to be unaware of the attack. Therefore, she cannot customize her power allocation strategy to combat such an attacker. However, while the attacker tries to minimize the downlink sum-rate, Alice tries to maximize this rate, without knowing about the attack. This is a *min-max* problem, and its solution is found as follows. As seen from (15), R_{sum} is a function of $\mathbf{P}^{(d)} \triangleq \left[P_1^{(d)} \cdots P_K^{(d)}\right]$ and $\boldsymbol{\alpha} \triangleq [\alpha_1, \cdots, \alpha_K]$. Thus, the problem can be formulated as a *convex-concave* game; for a fixed $\mathbf{P}^{(d)}$, $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha})$ is a convex function of $\boldsymbol{\alpha}$, and for a fixed $\boldsymbol{\alpha}$, $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha})$ is a concave function of $\mathbf{P}^{(d)}$. This means that the attacker needs to solve the following game:

$$\begin{aligned}
\textbf{P3} : \underset{\{\boldsymbol{\alpha}\}}{\text{minimize}} \ & \left\{ \underset{\{\mathbf{P}^{(d)}\}}{\text{maximize}} \ R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha}) \right\} \\
s.t. \qquad & \alpha_k \geq 0 \ \forall k \in \mathcal{K}, \quad \sum_{k=1}^{K} \alpha_k \leq 1 \\
& P_k^{(d)} \geq 0 \ \forall k \in \mathcal{K}, \quad \sum_{k=1}^{K} P_k^{(d)} \leq P_A
\end{aligned}$$

Let an optimal solution of this game, or a *saddle point*, be $(\mathbf{P}^{(d)*}, \boldsymbol{\alpha}^*)$. That is (for any possible power allocation $\mathbf{P}^{(d)}$),

$$R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha}^*) \leq R_{\text{sum}}(\mathbf{P}^{(d)*}, \boldsymbol{\alpha}^*) \leq R_{\text{sum}}(\mathbf{P}^{(d)*}, \boldsymbol{\alpha}).$$

This relationship shows that an upper-bound on $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha})$ is obtained by solving **P3**. For instance, when $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$, $\mathbf{P}^{(d)*}$ maximizes $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha}^*)$. This optimal solution is obtained by a well-known water-filling technique. Specifically,

$$P_k^{(d)*} = \left[\eta - \frac{\alpha_k^* u_k z_J^{-\gamma} + z_k^{-\gamma} + \frac{1}{AP_k L}}{MAz_k^{-2\gamma}} \right]^+ \tag{20}$$

where η is a water-filling level chosen such that $\sum_{k=1}^{K} P_k^{(d)} = P_A$. η can be computed by bisection method as this summation is an increasing function of it. Similarly, when $\mathbf{P}^{(d)} = \mathbf{P}^{(d)*}$, $\boldsymbol{\alpha}^*$

minimizes $R_{\text{sum}}(\mathbf{P}^{(d)*}, \boldsymbol{\alpha})$. The optimal solution of this problem was previously given in Theorem 1. We propose to solve this game by using an iterative *Gauss-Seidel* method. To do that, we first solve $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha})$ for some initial values of α_k, e.g., $\alpha_k = 0 \ \forall k \in \mathcal{K}$ (initially, there is no PC attack). Then, the obtained $P_k^{(d)}$ values are used in $R_{\text{sum}}(\mathbf{P}^{(d)*}, \boldsymbol{\alpha})$, and this problem is solved with respect to $\alpha_k \ \forall k \in \mathcal{K}$ as explained in Theorem 1. After this step, the second iteration starts by solving $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha}^*)$ using the new values of α_k's. As the number of iterations increases, a better approximation for the saddle point is obtained. We evaluate the number of iterations required to reach the Nash equilibrium of this game, and observe that the algorithm almost always converges after 10 iterations.

Theorem 2: Gauss-Seidel iterations converge when used to solve **P3**.

Proof: See Appendix D. ■

Note that the above analysis applies to the case of perfect information where distances are known to the attacker. It can be easily extended to the case where only the probability distribution of distances is known. The same steps in Section IV-A2 are applied to account for the uncertainty. In particular, the expectation of $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha})$ over Z_J and Z_k's is considered in the objective function of **P3**. The resulting problem is still a convex-concave game that can be solved by the Gauss-Seidel method. We skip this analysis here due to space limitations.

V. Secrecy Analysis Under PC Attack

As we analyzed in the previous sections, channels between Bobs and Alice are near-orthogonal in massive MIMO systems as long as $M \gg K$. Indeed, as $M \to \infty$, inter-user interference vanishes in massive MIMO systems. The same reason also makes massive MIMO systems well-protected against passive eavesdroppers (Eve). For example, channels of Eve and Bobs are near-orthogonal as well, so the mutual information leakage at Eve is negligible. However, we showed the vulnerability of massive MIMO systems against an active attacker that contaminates the pilot transmissions. So far, we only considered the case where the attacker's objective is to minimize the downlink sum-rate. PC attack also makes Alice transmit information signals towards the attacker, as the precoding vectors are designed based on the erroneous channel estimates, which are linear combinations of CSI at Bobs as well as at the attacker. Therefore, the attacker receives the information signals intended to Bobs in the data transmission phase.

As a secrecy metric, we consider the individual secrecy rates of Bobs, which ensure that information leakage to an eavesdropper from each information message vanishes [15], [16].

Specifically, we study a problem in which the attacker, Eve, aims at minimizing the maximum of the achievable individual secrecy rates at Bobs by leveraging PC attacks. Given MRT precoding at Alice and the same attack model detailed in Sections II and III, the received signal at the attacker in the downlink data transmission phase is given by:

$$y_{\text{eve}} = \sum_{i=1}^{K} \sqrt{P_i^{(d)}} \mathbf{h}_J \mathbf{v}_i^T s_i + w_J \tag{21}$$

where w_J is the AWGN at Eve, and $\mathbf{v}_k^T = (\mathbf{h}_k + \sqrt{\alpha_k u_k}\mathbf{h}_J + \tilde{\mathbf{w}}_k)^* / \|\mathbf{h}_k + \sqrt{\alpha_k u_k}\mathbf{h}_J + \tilde{\mathbf{w}}_k\|$. The individual information leakage rate of s_k is given by:

$$R_k^e = \log\left(1 + \frac{P_k^{(d)} |\mathbf{h}_J \mathbf{v}_k^T|^2}{\sum_{l \in \{\mathcal{K} \setminus k\}} P_l^{(d)} |\mathbf{h}_J \mathbf{v}_l^T|^2 + 1}\right), \ \forall k \in \mathcal{K} \tag{22}$$

Note that R_k^e is obtained from the mutual information between s_k and y_{eve} where the all other information signals are interpreted as noise. Similar to the previous sections, we analyze the asymptotic behavior of R_k^e as $M \to \infty$. As a result of this analysis, the following limit is obtained:

$$\lim_{M \to \infty} \frac{P_k^{(d)} |\mathbf{h}_J \mathbf{v}_k^T|^2}{M} = \frac{P_k^{(d)} \alpha_k u_k \theta_J^2}{\theta_k + \alpha_k u_k \theta_J + \frac{1}{P_k L}} \tag{23}$$

(see Appendix E for the derivation of (23)). Therefore:

$$\lim_{M \to \infty} R_k^e = \log\left(1 + \frac{\dfrac{P_k^{(d)} \alpha_k u_k \theta_J^2}{\theta_k + \alpha_k u_k \theta_J + \dfrac{1}{P_k L}}}{\sum_{l \in \{\mathcal{K} \setminus k\}} \dfrac{P_l^{(d)} \alpha_l u_l \theta_J^2}{\theta_l + \alpha_l u_l \theta_J + \frac{1}{P_l L}}}\right). \tag{24}$$

Note that $\lim_{M \to \infty} R_k^e$ is independent of M. As we analyze the asymptotic behavior of the system, $\lim_{M \to \infty} R_k^e$ is referred to as R_k^e in the rest of the paper.

A. Known Distances at Attacker

The achievable individual secrecy rate for Bob$_k$ is defined by $[R_k - R_k^e]^+$ [16]. Under the fixed power allocation for the downlink signals at Alice and perfect information at the attacker, the optimal PC attack to minimize the maximum of individual secrecy rates is formulated as follows:

$$\underset{\{\alpha_k, \forall k \in \mathcal{K}\}}{\text{minimize}} \quad \max\{R_1 - R_1^e, \cdots, R_K - R_K^e, 0\}$$

$$s.t. \qquad \alpha_k \geq 0 \; \forall k \in \mathcal{K}, \quad \sum_{k=1}^{K} \alpha_k \leq 1$$

We reformulate this problem by introducing a new decision variable ν, such that $\nu \geq \max\{R_1 - R_1^e, \cdots, R_K - R_K^e, 0\}$. This is equivalent to $\nu \geq 0$ and $\nu \geq R_k - R_k^e \; \forall k \in \mathcal{K}$. The problem is now converted to the following one:

$$\begin{aligned} \textbf{P4}: \underset{\{\nu, \alpha_k, \, \forall k \in \mathcal{K}\}}{\text{minimize}} \quad & \nu \\ s.t. \qquad & R_k - R_k^e - \nu \leq 0 \; \forall k \in \mathcal{K} \\ & \nu \geq 0, \quad \alpha_k \geq 0 \; \forall k \in \mathcal{K}, \quad \sum_{k=1}^{K} \alpha_k \leq 1 \end{aligned}$$

Note that the solution of **P4** provides the tightest upper bound on the achievable individual secrecy rate that can be achieved by any Bob in a given massive MIMO system under an optimal PC attack. However, due to the interference of the information signals at Eve, the first constraint function in **P4** is not convex. This makes the problem intractable. Let $G_k \triangleq P_k^{(d)} \theta_J$. Therefore, R_k and R_k^e are given as follows:

$$R_k = \log\left(1 + \frac{A_k}{\alpha_k + B_k}\right), \tag{25}$$

$$R_k^e = \log\left(1 + \frac{\dfrac{G_k \alpha_k}{\alpha_k + B_k}}{\sum_{l \neq k}^{K} \dfrac{G_l \alpha_l}{\alpha_l + B_l}}\right). \tag{26}$$

Let U_k be an upper bound on $R_k - R_k^e$ such that:

$$U_k \triangleq R_k - \log\left(1 + \frac{G_k \alpha_k}{(\alpha_k + B_k) I_k}\right) \tag{27}$$

where $I_k \triangleq (\sum_{l \neq k}^{K} G_l / B_l)$. Note that the function $R_k - R_k^e$ is a monotonically increasing function with respect to $\alpha_l \; \forall l \in \mathcal{K}$, $l \neq k$. An upper bound of this function is obtained when $\alpha_l = 1$ $\forall l \in \mathcal{K}$, $l \neq k$. Replacing the first constraint in **P4** by $U_k - \nu \leq 0 \; \forall k \in \mathcal{K}$ makes the problem tractable, and its solution still provides an upper bound on the achievable individual secrecy rate for any Bob. Furthermore, the logarithm function can be removed by defining $\hat{\nu} \triangleq 2^{\nu}$. Then, **P4** becomes:

$$\begin{aligned} \textbf{P5}: \underset{\{\hat{\nu}, \alpha_k, \, \forall k \in \mathcal{K}\}}{\text{minimize}} \quad & \hat{\nu} \\ s.t. \qquad & \frac{I_k(\alpha_k + A_k + B_k)}{\alpha_k(I_k + G_k) + B_k I_k} - \hat{\nu} \leq 0 \; \forall k \in \mathcal{K} \end{aligned}$$

$$\hat{\nu} \geq 1, \quad \alpha_k \geq 0 \; \forall k \in \mathcal{K}, \quad \sum_{k=1}^{K} \alpha_k \leq 1$$

Let $f_k(\alpha_k)$ denote $I_k(\alpha_k + A_k + B_k)/(\alpha_k(I_k + G_k) + B_k I_k)$. Then, $f_k(\alpha_k)$ is a monotonically decreasing function with respect to α_k. Thus, we propose the following iterative method to numerically solve **P5**. Initially, $\max\{f_1(\alpha_1), \cdots, f_K(\alpha_K)\}$ is found where $\alpha_k = 0 \; \forall k \in \mathcal{K}$. WLOG, consider that $f_i(\alpha_i)$ is the maximum. Then, $\alpha_i \leftarrow \alpha_i + \delta$ where δ is a positive real number. After that, the same process is repeated with the new values of α_k's as long as $\sum_{k=1}^{K} \alpha_k \leq 1$ and $\hat{\nu} \geq 1$. In Section VII, we numerically compare two upper bounds that are obtained by **P4** and **P5**.

B. Unknown Distances at Attacker

If the location information of Bobs is not available at the attacker, she cannot guarantee any upper bound on the individual secrecy rates. Therefore, in this case, we replace the first constraint of **P4** by a chance constraint, as follows:

$$\Pr\{R_k - R_k^e \geq \nu\} \leq \epsilon \; \forall k \in \mathcal{K} \tag{28}$$

where $\epsilon \in [0, 1]$ is a given parameter. Note that the randomness in (28) comes from the distances between Bobs and Alice, $Z_k \; \forall k \in \mathcal{K}$. (We assume that the attacker knows her distance to Alice, which is a stationary massive MIMO BS.) This constraint guarantees that the probability of achieving an individual secrecy rate that is higher than or equal to ν is less than or equal to ϵ at Bobs. That is, only ϵ fraction of Bobs can achieve an individual secrecy rate above ν. As (28) does not have a closed-form expression, **P4** is intractable for this case as well. Therefore, we use a similar bounding method as in the known distances case to make the problem tractable. Let $\hat{U}_k$ be an upper bound on $R_k - R_k^e$ such that:

$$\hat{U}_k \triangleq R_k - \log\left(1 + \frac{G_k \alpha_k}{(\alpha_k + B_k)\hat{I}_k}\right) \tag{29}$$

where

$$\hat{I}_k \triangleq \sum_{l \neq k}^{K} \frac{P_l^{(d)} A u_l z_J^{-2\gamma}}{u_l z_J^{-\gamma} + D_{\max}^{-\gamma} + \frac{1}{A P_l L}}. \tag{30}$$

Note that the function $R_k - R_k^e$ is a monotonically increasing function with respect to both α_l and $z_l \; \forall l \in \mathcal{K}$, $l \neq k$. An upper bound of this function is obtained when $\alpha_l = 1$ and $z_l = D_{\max}$ $\forall l \in \mathcal{K}$, $l \neq k$. Thus, the following inequalities are obtained:

$$\Pr\{R_k - R_k^e \geq \nu\} \leq \Pr\{\hat{U}_k \geq \nu\} \tag{31}$$

$$= \Pr\{P_k^{(d)} MA\hat{I}_k Z_k^{-2\gamma} - (\hat{\nu}-1)\hat{I}_k Z_k^{-\gamma} \geq$$

$$\hat{\nu} P_k^{(d)} \alpha_k u_k A z_J^{-2\gamma} + (\hat{\nu}-1) I_k (\alpha_k u_k z_J^{-\gamma} + (AP_k L)^{-1})\} \tag{32}$$

$$\leq \Pr\{P_k^{(d)} MA\hat{I}_k Z_k^{-2\gamma} - (\hat{\nu}-1)\hat{I}_k Z_k^{-2\gamma} \geq$$

$$\hat{\nu} P_k^{(d)} \alpha_k u_k A z_J^{-2\gamma} + (\hat{\nu}-1) I_k (\alpha_k u_k z_J^{-\gamma} + (AP_k L)^{-1})\} \tag{33}$$

$$= \Pr\left\{ Z_k \leq \sqrt[2\gamma]{\frac{P_k^{(d)} MA\hat{I}_k - (\hat{\nu}-1)\hat{I}_k}{J_k}} \right\} \tag{34}$$

where $J_k \triangleq \hat{\nu} P_k^{(d)} \alpha_k u_k A z_J^{-2\gamma} + (\hat{\nu}-1) I_k (\alpha_k u_k z_J^{-\gamma} + (AP_k L)^{-1})$. To analyze the chance constraint, we exploit (34), which is the CDF of Z_k. As we stated before, $\Pr[Z_k \leq x] = (x^2 - D_{\min}^2)/(D_{\max}^2 - D_{\min}^2)$ where $x \in [D_{\min}, D_{\max}]$. Hence, the chance constraint (28) is converted to:

$$\frac{P_k^{(d)} MA\hat{I}_k - (\hat{\nu}-1)\hat{I}_k}{J_k} \leq (\epsilon(D_{\max}^2 - D_{\min}^2) + D_{\min}^2)^{\gamma} \tag{35}$$

$\forall k \in \mathcal{K}$. This is equivalent to:

$$\frac{\hat{I}_k (P_k^{(d)} MA + 1 + Q(\alpha_k u_k z_J^{-\gamma} + (AP_k L)^{-1}))}{\hat{I}_k + Q(P_k^{(d)} \alpha_k u_k A z_J^{-2\gamma} + \hat{I}_k (\alpha_k u_k z_J^{-\gamma} + (AP_k L)^{-1}))} \leq \hat{\nu} \tag{36}$$

$\forall k \in \mathcal{K}$ where $Q = (\epsilon(D_{\max}^2 - D_{\min}^2) + D_{\min}^2)^{\gamma}$. To find the minimum $\hat{\nu}$ for a given ϵ, the same problem as **P5** is considered at the attacker after replacing the first constraint by (36). Note that the constraint function in (36) is a monotonically decreasing function with respect to α_k. Therefore, the method that we propose for solving **P5** in the previous subsection can be used here as well.

In this paper, we study the problem of minimizing the maximum of individual secrecy rates. The problem in which the attacker aims at minimizing the sum of the individual secrecy rates could be also solved by following similar steps. Particularly, the problem would be similarly reformulated, and the new problem would be a convex optimization problem as well. Due to space limitations, we omit the results here.

VI. Hybrid Full-duplex Attack

So far, we have considered jamming the pilot transmission phase. However, if the attacker is equipped with a full-duplex (FD) radio that allows it to transmit and receive signals simultaneously over the same frequency, a more sophisticated attack can be launched. Further, a stronger attack can also be launched with a multi-antenna (MIMO) FD-based attacker. In particular, consider an attacker with an average power constraint over the whole transmission phase (pilot

(a) (b)

Fig. 2. (a) Attacker contaminates the CSI estimation at Alice while overhearing the pilots from Bobs, (b) attacker generates the jamming signals to reduce the signal strength at Bobs during data transmission.

and downlink data phases). Using an FD radio, the attacker can generate jamming signals during both phases. For instance, the attacker may contaminate the CSI estimation process at Alice, as in Fig. 2(a), without knowing the channels between itself and Bobs. At the same time, the attacker can overhear the pilots (dashed lines in Fig. 2(a)) from Bobs using the FD radio, and exploit this knowledge to transmit jamming signals during the downlink transmission phase, as shown in Fig. 2(b). We call this attack a *hybrid* attack, as it combines the PC attack and conventional data-jamming attack. Notice that even though the hybrid attack performs at least as good as the PC attack, it requires an additional hardware capability (FD radio) at the attacker.

Even though the attacker needs one antenna to generate a jamming signal in data transmission phase, we study a more general scenario where she is equipped with $N+1$ antennas, where $N>0$. Our goal is to find an optimal strategy for the attacker to minimize the downlink sum-rate, exploiting its multiple antennas. One of these antennas is reserved for the PC attack, while the others receive the pilot signals from Bobs. The attacker estimates $\mathbf{h}_{Jk} \in \mathbb{C}^{1\times N}$, the channel between Bob_k and herself, during the pilot transmission phase. The self-interference signal at the receiving antennas of the attacker is canceled by employing FD radio design techniques in [17], [18]. For example, the self-interference channel is obtained by transmitting a pilot from the antenna that jams the pilot signal. Then, the self-interference signal is extracted from the received signals using this information. Let n_i be the ith jamming signal in the downlink transmission phase, $i \in \mathcal{N} = \{1, \cdots, N\}$. Let $h_{Jk}^{(i)} = g_{Jk}^{(i)}\sqrt{Az_{Jk}^{-\gamma}}$, $i \in \mathcal{N}$ and $k \in \mathcal{K}$, be the channel gain between the ith antenna of the attacker and Bob_k, where z_{Jk} denotes the distance between the attacker and Bob_k and $g_{Jk}^{(i)}$ is the small-scale fading. β_i $\forall i \in \mathcal{N}$ denotes the ratio between the

allocated power for n_i and P_J. By using the same PC attack model in Section II-A and MRT precoding at Alice, the received signal at Bob$_k$ during the downlink data transmission phase is given by:

$$y_k = \sum_{i=1}^{K} \sqrt{P_i^{(d)}} \mathbf{h}_k \frac{\hat{\mathbf{h}}_i^*}{\|\hat{\mathbf{h}}_i\|} s_i + \sum_{i=1}^{N} \sqrt{\beta_i P_J} h_{Jk}^{(i)} n_i + w_k^{(d)}. \tag{37}$$

Adding the jamming term to (10), the following downlink sum-rate is obtained:

$$R_{\text{sum}} = \sum_{k=1}^{K} \log \left(1 + \frac{C_k}{D_k \left(\sum_{i=1}^{N} \beta_i P_J |g_{Jk}^{(i)}|^2 A z_{Jk}^{-e} + 1 \right)} \right) \tag{38}$$

where

$$C_k \triangleq P_k^{(d)} M A z_k^{-2\gamma} \text{ and } D_k \triangleq \alpha_k u_k z_J^{-\gamma} + z_k^{-\gamma} + \frac{1}{A P_k L}.$$

In this section, we do not analyze the secrecy, and focus on the attack that is studied in Section IV. Given the setup above, we formulate a two-stage stochastic optimization problem to find the optimal attacking strategy that minimizes the downlink sum-rate at Bobs. This problem can be solved for various scenarios (e.g., perfect information, uncertainity in the distances and channels, etc.) by utilizing the techniques in Section IV and the ones presented in this section. The solutions of these problems are discussed in Section VII. For now, we explain our solution approach for one of these scenarios. Specifically, we assume that the distances, powers of information signals, and other constants in (38) are known to the attacker. In the first stage of the problem, the attacker finds the optimal values of α_k $\forall k \in \mathcal{K}$ without knowing any $g_{Jk}^{(i)}$ $\forall k \in \mathcal{K}$ and $\forall i \in \mathcal{N}$. In the second stage (after learning $g_{Jk}^{(i)}$ $\forall k \in \mathcal{K}$ and $\forall i \in \mathcal{N}$ during the pilot transmission phase), the attacker optimally allocates the remaining power to the N jamming signals in the data transmission phase, i.e., β_i $\forall i \in \mathcal{N}$. Let ω represent a certain realization of the channel, $g_{Jk}^{(i)}$, and let Ω be the set of all realizations. (Note that $g_{Jk}^{(i)}$ and β_i are functions of these realizations.) Let t_p and t_d be the duration of pilot and data transmission phases, respectively. The two-stage stochastic problem can be formulated as follows:

$$\textbf{P6}: \underset{\substack{\{\alpha_k \ \forall k \in \mathcal{K}\} \\ \{\beta_i(\omega) \ \forall i \in \mathcal{N}, \ \forall \omega \in \Omega\}}}{\text{minimize}} \quad \mathbb{E}_\omega \left[\sum_{k=1}^{K} \log \left(1 + \frac{C_k}{D_k \left(E_k + 1 \right)} \right) \right]$$

$$\begin{aligned} s.t. \quad & \alpha_k \geq 0 \ \forall k \in \mathcal{K} \\ & \beta_i(\omega) \geq 0 \ \forall i \in \mathcal{N}, \ \forall \omega \in \Omega \\ & \frac{F_k}{t_p + t_d} \leq 1 \ \forall \omega \in \Omega \end{aligned}$$

where $F_k \triangleq t_p \sum_{k=1}^{K} \alpha_k + t_d \sum_{i=1}^{N} \beta_i(\omega)$ and $E_k \triangleq \sum_{i=1}^{N} \beta_i(\omega) |g_{Jk}^{(i)}(\omega)|^2 P_J A z_{Jk}^{-e}$. Note that $g_{Jk}^{(i)}$ is a continuous random variable. **P6** can be approximately solved by creating T realizations, e.g., Ω has a cardinality of T. In particular, we replace the expectation in **P6** by the sum of these equiprobable T realizations. Therefore, we end up with K first-stage decision variables, namely α_k $\forall k \in \mathcal{K}$, and NT second-stage decision variables, namely $\beta_i(\omega)$ $\forall i \in \mathcal{N}$ and $\forall \omega \in \Omega$. The underlying problem is a convex programming problem, and can be solved by the interior point method. When T is large (for better approximation), the complexity of solving the problem increases. However, as the problem is solved offline, the time complexity is not a concern.

VII. Numerical Results and Discussion

We model the channel gain from each transmit antenna to each receive antenna as $h = g\sqrt{Ad^{-3.522}}$, where $g \sim \mathcal{CN}(0,1)$ and $A = 3.0682 \times 10^{-5}$. The path-loss is modeled using the COST-Hata Model with center frequency is 2 GHz [19]. The average transmit powers at Alice, Bob$_k$, and the attacker are 46, 20, and 30 dBm, respectively. The durations of the pilot and data transmission phases are set to be equal [4]. We consider a 20 MHz channel with noise floor of -101 dBm. Bobs and the attacker are uniformly and randomly distributed within a circular ring whose center is Alice and whose outer radius is $D_{\max}$ and $D_{\max,\mathrm{J}}$, respectively. We set $D_{\max}$ to 750 meters and $D_{\min}$ to 10 meters. Our results are averaged over 10^5 different network realizations.

We set the number of users $K = 10$. In Fig. 3(a), we consider uniform power allocation for both the information signals at Alice and the jamming signals at the attacker. The figure depicts the downlink sum-rate vs. M. It shows that (10) and (13) are good approximations for the downlink rates in (7). Note that the approximation-based sum-rate is slightly higher than the exact values, as the inter-user interference does not perfectly vanish at a finite M. In our subsequent results, we set M to 1000.

We observe the effect of the maximum distance between Alice and the attacker ($D_{\max,\mathrm{J}}$) in Figs. 3(b) and 3(c). In the case of a single-user PC attack, only one randomly selected Bob is targeted by the attacker. This attack can also be interpreted as an unintentional interference from a user in an adjacent cell. It does not have a big impact on the sum-rate. PC with uncertainty (PC-unc) and PC with perfect information (PC-pi) were explained in Section IV-A, and optimal PC-pi was studied in Section IV-B. Note that optimal PC-pi gives an upper-bound on the sum-rate of a massive MIMO system under an optimal PC attack. As the attacker moves farther

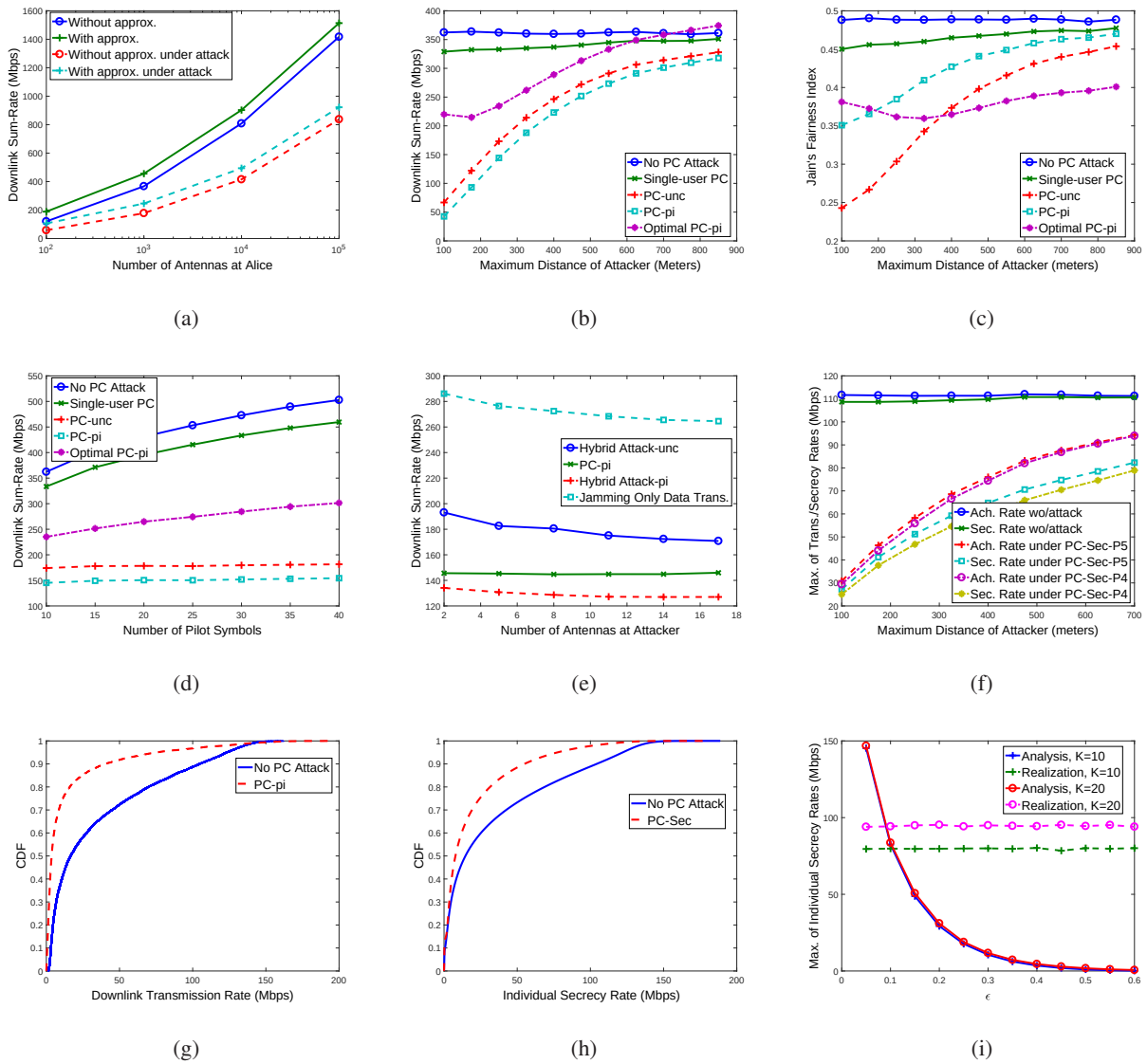

Fig. 3. (a) Downlink sum-rate vs. M under uniform power allocation at both Alice and the attacker, (b) downlink sum-rate vs. $D_{\max,\mathrm{J}}$, (c) Jain's fairness index vs. $D_{\max,\mathrm{J}}$, (d) downlink sum-rate vs. number of pilot symbols, (e) downlink sum-rate vs. number of antennas at the attacker, (f) maximum of individual secrecy rates vs. $D_{\max,\mathrm{J}}$, (g) CDF vs. downlink transmission rate, (h) CDF vs. individual secrecy rate, (i) maximum of individual secrecy rates vs. ϵ.

from Alice, the sum-rate increases in all attack schemes. In Fig. 3(b), EVPI is around 20 Mbps. This says that when the attacker knows the distribution of Bobs, it can launch attacks that are almost as powerful as when the attacker has complete CSI. We also observe that the downlink sum-rate without an attack (no PC attack case) is less than the one with the optimal PC-pi if $D_{\max,\mathrm{J}}$ is higher than 700 meters. The reason is that Alice uniformly allocates downlink transmission powers in no PC attack scheme, whereas she employs optimal power allocation

in the optimal PC-pi. In Fig. 3(c), we depict Jain's fairness index for different schemes. Jain's fairness index ranges from $1/K$ to 1 for the worst and best cases, respectively (if all users have the same downlink rate, the fairness index is 1). The figure shows that fairness among Bobs is significantly reduced when PC attacks take place. PC-unc decreases the fairness more than PC-pi. The reason behind this phenomena is that when the attacker is close to Alice and knows the distances, Bobs with higher downlink rates are targeted. Therefore, Bobs are forced to have closer downlink rates, which increases the fairness index. Note that even though PC-pi makes the fairness higher compared to PC-unc, the sum-rate is lower in PC-pi.

In Fig 3(d), we set $D_{\max,\mathrm{J}}$ to 250 meters, and study the effect of the number of pilot symbols L. As L increases, the sum-rate increases as well in no PC, single-user PC, and optimal PC-pi attacks. The reason is that the error in MRT precoding vectors due to erroneous channel estimates decreases, and the signal strength at Bobs increases. On the other hand, the sum-rate does not increase under the PC-unc and PC-pi attacks. Note that in these cases, a fixed power is allocated for the information signals at Alice, and she does not exploit the decrease in channel estimation errors.

In Fig. 3(e), we compare hybrid and PC attacks under a similar average jamming power constraint. We observe that the hybrid attacks outperform PC attacks with respect to the sum-rate. Moreover, as the number of antennas at the attacker increases, the sum-rate slightly decreases for the hybrid attack. Note that the hybrid attacks utilizes multiple antennas, whereas PC attacks use a single-antenna. EVPI for the hybrid attacks is around 60 Mbps, which is much higher than the one for PC attacks. The reason is that the hybrid attack includes one more source of uncertainty due to the channels between Bobs and the attacker. Another important result is that attacking only downlink data transmissions (no jamming during pilot transmission phase) does not have as a great of an impact on performance as the impact of the PC attack.

We evaluate the effect of PC attack on individual secrecy rates in Fig. 3(f). Specifically, we compare the schemes where there is no PC attack and PC attacks whose objective is to minimize the maximum of the individual secrecy rates (PC-Sec). PC-Sec-P4 and PC-Sec-P5 denote the results of the problems **P4** and **P5**, respectively, with known distances. Note that even though **P4** is not a tractable problem, we obtain its results with a brute force method. For each scheme, we show the results of both individual secrecy rates and transmission rates between Alice and Bobs. It is observed that the massive MIMO systems are resilient to passive eavesdroppers, as the maximum of transmission/secrecy rates are almost the same without a PC attack. On the

other hand, PC attack decreases the maximum of individual secrecy rates from nearly 110 Mbps to 55 Mbps when $D_{\max,\mathrm{J}}$ is 325 meters. Moreover, we observe that when the attacker moves farther from Alice, PC attack still reduces the maximum of individual secrecy rates by almost 30%. It is also noted that the solution of **P5** is very close to the solution of **P4**, which provides the tightest upper bound.

The empirical CDF of downlink transmission rate and individual secrecy rate under various schemes are shown in Figs. 3(g) and 3(h), respectively. 90% of Bobs achieve a transmission rate less than 40 Mbps under PC-pi. In the absence of a PC attack, nearly 33% of Bobs achieve a transmission rate higher than 40 Mbps. In Fig. 3(h), we observe that 13% of Bobs have a zero individual secrecy rate under PC-Sec, whereas only 7% fraction of Bobs have a zero individual secrecy rate when there is no PC attack. Moreover, only 5% of Bobs have a secrecy rate above 75 Mbps.

In Fig. 3(i), we evaluate our secrecy analysis with unknown distances at the attacker. We observe the effect of the designed parameter ϵ on the maximum of individual secrecy rates for both cases where $K = 10$ and $K = 20$. Based on our analysis, 0.1 fraction of Bobs may achieve an individual secrecy rate higher than 83 Mbps. When $K = 10$, the maximum of individual secrecy rates is just below this threshold value on average. On the other hand, when $K = 20$, this threshold value is exceeded almost always as expected. Note that when $\epsilon = 0.6$, the attacker guarantees that at least 0.4 fraction of Bobs have zero individual secrecy rate, which emphasizes the vulnerability of a massive MIMO system against a PC attack.

VIII. Conclusion

We considered a single-cell massive MIMO system with several mobile users, and demonstrated vulnerabilities of uplink pilot transmissions against jamming attacks. Specifically, the attacker generates pilot sequences similar to those of users and contaminates the pilot transmissions to distort channel estimation at the BS. This PC attack reduces the downlink transmission rates, as the beamforming techniques utilized by the BS heavily depend on accurate CSI estimates. We formulated an optimization problem from the standpoint of the attacker to minimize the downlink sum-rate. Both cases when the attacker knows or does not know the distances between the BS and users were considered. Using (stochastic) optimization and game theory, we derived the optimal attacking strategies when the BS employs either fixed or optimal power allocation for downlink transmissions. We also analyzed the secrecy rates of the users in massive MIMO

systems. In particular, we showed that even though such systems are robust against a passive eavesdropper, the PC attack significantly reduces the maximum of the individual secrecy rates. Numerical results showed that the downlink sum-rate is reduced by more than 50% if the average distance between the attacker and the BS is less than the one of the users. We also observed that even if the attacker does not know the channels and the locations of the users, it can launch powerful attacks as if it has the perfect information. In this work, we assumed that the BS and users are not aware of the attacker. An interesting future work is to develop counter algorithms to prevent PC attacks.

APPENDIX A
PROOF OF EQUATION (8)

$$\lim_{M\to\infty} \frac{P_l^{(d)}|\mathbf{h}_k\mathbf{v}_l^T|^2}{M} = \lim_{M\to\infty} \frac{P_l^{(d)}|\frac{\mathbf{h}_k\hat{\mathbf{h}}_l^*}{M}|^2}{\frac{\|\hat{\mathbf{h}}_l\|^2}{M}} \tag{39}$$

Let us evaluate the limit of the numerator and denominator separately. The limit of the denominator is given by:

$$\lim_{M\to\infty} \frac{\|\hat{\mathbf{h}}_l\|^2}{M} = \theta_l + \frac{1}{P_l L} \tag{40}$$

The equality is due to the fact that given a vector $\mathbf{x} \in \mathbb{C}^{1\times M}$ with a distribution $\mathcal{CN}(\mathbf{0}, c\mathbf{I})$, $\lim_{M\to\infty} \mathbf{x}\mathbf{x}^*/M = c$ [11, Lemma 1]. We analyze the limit of the numerator as follows:

$$\lim_{M\to\infty} \frac{\mathbf{h}_k\hat{\mathbf{h}}_l^*}{M} = \lim_{M\to\infty} \frac{\sqrt{\theta_k}\mathbf{g}_k(\sqrt{\theta_l}\mathbf{g}_l + \tilde{\mathbf{w}}_l)^*}{M} = 0 \tag{41}$$

$\mathbf{g}_k$, $\mathbf{g}_l$, and $\tilde{\mathbf{w}}_l$ are independent vectors, and the result follows from [11, Lemma 1]. The expression in the numerator of (39) is a continuous function of $\mathbf{h}_k\hat{\mathbf{h}}_l^*/M$. Therefore, using the Continuous Mapping Theorem, we have the following result:

$$\lim_{M\to\infty} P_l^{(d)}|\frac{\mathbf{h}_k\hat{\mathbf{h}}_l^*}{M}|^2 = 0 \tag{42}$$

It proves the equation (8).

APPENDIX B

PROOF OF EQUATION (9)

$$\lim_{M\to\infty} \frac{P_k^{(d)}|\mathbf{h}_k\mathbf{v}_k^T|^2}{M} = \lim_{M\to\infty} \frac{P_k^{(d)}|\frac{\mathbf{h}_k\hat{\mathbf{h}}_k^*}{M}|^2}{\frac{\|\hat{\mathbf{h}}_k\|^2}{M}} \tag{43}$$

In this proof, we follow the same steps as in Appendix A. Therefore, $\lim_{M\to\infty}\|\hat{\mathbf{h}}_k\|^2/M = \theta_k + 1/(P_k L)$. We exploit the Continuous Mapping Theorem to evaluate the limit of the numerator as follows:

$$\lim_{M\to\infty} \frac{\mathbf{h}_k\hat{\mathbf{h}}_k^*}{M} = \lim_{M\to\infty} \frac{\sqrt{\theta_k}\mathbf{g}_k(\sqrt{\theta_k}\mathbf{g}_k + \tilde{\mathbf{w}}_k)^*}{M} \tag{44}$$

$$= \lim_{M\to\infty} \frac{\theta_k\mathbf{g}_k\mathbf{g}_k^*}{M} = \theta_k \tag{45}$$

Hence,

$$\lim_{M\to\infty} \frac{P_k^{(d)}|\mathbf{h}_k\mathbf{v}_k^T|^2}{M} = \frac{P_k^{(d)}\theta_k^2}{\theta_k + \frac{1}{P_k L}} \tag{46}$$

APPENDIX C

PROOF OF THEOREM 2

Let us define

$$A_k = \frac{P_k^{(d)} M A z_J^\gamma}{u_k z_k^{2\gamma}} \text{ and } B_k = \frac{z_J^\gamma}{u_k z_k^\gamma} + \frac{z_J^\gamma}{u_k A P_k L}$$

$\forall k \in \mathcal{K}$. Therefore, the objective of **P1** can be written by

$$R_{\text{sum}} = \sum_{k=1}^{K} \log\left(1 + \frac{A_k}{\alpha_k + B_k}\right) \tag{47}$$

Hence, the Lagrangian function of this problem is given by

$$L(\boldsymbol{\alpha}) = \sum_{k=1}^{K} \log(1 + \frac{A_k}{\alpha_k + B_k}) + \lambda(\sum_{k=1}^{K} \alpha_k - 1). \tag{48}$$

Its first derivative with respect to α_k becomes

$$\frac{\partial L(\boldsymbol{\alpha})}{\partial \alpha_k} = \frac{-A_k}{(\alpha_k + B_k)(\alpha_k + A_k + B_k)} + \lambda. \tag{49}$$

Let α_k^* $\forall k \in \mathcal{K}$ be the optimal value that minimizes the objective function of **P1**. These values are also the roots of the polynomial functions where the equation (49) is equal to zero. Also,

note that $\alpha_k^* \ \forall k \in \mathcal{K}$ is a nonnegative number, and their summation is equal to 1 due to the complementary slackness. Therefore,

$$\alpha_k^* = \left[\frac{\sqrt{A_k(A_k + 4/\lambda)} - A_k - 2B_k}{2} \right]^+ \tag{50}$$

where λ is chosen such that $\sum_{k=1}^{K} \alpha_k^* = 1$.

APPENDIX D
PROOF OF THEOREM 3

The players of the game described in **P3** are Alice and the attacker. In this game, the utility function of Alice is $R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha})$, and her strategy is to choose the optimal power allocation for the downlink transmissions. Similarly, $-R_{\text{sum}}(\mathbf{P}^{(d)}, \boldsymbol{\alpha})$ is the attacker's utility, and her strategy is to find the optimal $\boldsymbol{\alpha}$ to maximize this utility. The strategy sets of both players are non-empty, compact, and convex subsets of real numbers (the constraints in **P3** are linear functions). Furthermore, their utility functions are continuous and diagonally strictly concave. As a result, the existence and uniqueness of NE is proved for this game, and Gauss-Seidel method converges to this point [20].

APPENDIX E
PROOF OF EQUATION (23)

$$\lim_{M\to\infty} \frac{P_k^{(d)} |\mathbf{h}_J \mathbf{v}_k^T|^2}{M} = \lim_{M\to\infty} \frac{P_k^{(d)} |\frac{\mathbf{h}_J \hat{\mathbf{h}}_k^*}{M}|^2}{\frac{\|\hat{\mathbf{h}}_k\|^2}{M}} \tag{51}$$

Similar to the analysis in Appendices A and B, the limit of the denominator is $\theta_k + \alpha_k u_k \theta_J + \frac{1}{P_k L}$. We again use the Continuous Mapping Theorem to find the limit of the numerator. In particular,

$$\lim_{M\to\infty} \frac{\mathbf{h}_J \hat{\mathbf{h}}_k^*}{M} = \lim_{M\to\infty} \frac{\sqrt{\alpha_k u_k} \theta_J \mathbf{g}_J \mathbf{g}_J^*}{M} \tag{52}$$

$$= \sqrt{\alpha_k u_k} \theta_k \tag{53}$$

It proves the equation (23).

REFERENCES

[1] E. Bj, E. G. Larsson, T. L. Marzetta *et al.*, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

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

[3] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

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

[5] X. Zhou, B. Maham, and A. Hjorungnes, "Pilot contamination for active eavesdropping," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 903–907, March 2012.

[6] J. K. Tugnait, "Self-contamination for detection of pilot contamination attack in multiple antenna systems," *IEEE Wireless Communications Letters*, vol. 4, no. 5, pp. 525–528, Oct. 2015.

[7] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 21–27, June 2015.

[8] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission in the presence of an active eavesdropper," in *Proc. of the IEEE International Conference on Communications (ICC)*, London, UK, June 2015, pp. 1434–1440.

[9] Y. O. Basciftci, C. E. Koksal, and A. Ashikhmin. (Mar. 2015) "Physical layer security in massive MIMO". [Online]. Available: http://arxiv.org/abs/1505.00396

[10] B. Akgun, M. Krunz, and O. O. Koyluoglu, "Pilot contamination attacks in massive MIMO systems," in *Proc. of the IEEE CNS Conference*, Las Vegas, Oct. 2017.

[11] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[12] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[13] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. F. Abdelzaher, "Range-free localization and its impact on large scale sensor networks," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 4, no. 4, pp. 877–906, Nov. 2005.

[14] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 24–40, July 2005.

[15] B. Akgun, O. O. Koyluoglu, and M. Krunz, "Exploiting full-duplex receivers for achieving secret communications in multiuser MISO networks," *IEEE Transactions on Communications*, vol. 65, no. 2, pp. 956–968, Feb. 2017.

[16] Y. Chen, O. O. Koyluoglu, and A. Sezgin, "Individual secrecy for broadcast channels with receiver side information," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4687–4708, July 2017.

[17] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 375–386, Aug. 2013.

[18] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proc. of the 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, Seattle, April 2014, pp. 359–372.

[19] V. Abhayawardhana, I. Wassell, D. Crosby, M. Sellars, and M. Brown, "Comparison of empirical propagation path loss models for fixed wireless access systems," in *Proc. of the IEEE VTC'05*, Stockholm, Sweden, May 2005, pp. 73–77.

[20] R. Cominetti, F. Facchinei, and J. B. Lasserre, *Modern optimization modelling techniques*. Springer Science & Business Media, 2012.