Secure Massive MIMO with the Artificial Noise-Aided Downlink Training

Nguyen, P., Ngo, H. Q., Duong, T. Q., Tuan, H. D., & Tourki, K. (2018). Secure Massive MIMO with the Artificial Noise-Aided Downlink Training. *IEEE Journal on Selected Areas in Communications*, *36*(4), 802-816. Advance online publication. https://doi.org/10.1109/JSAC.2018.2825140

Published in:
IEEE Journal on Selected Areas in Communications

Secure Massive MIMO with the Artificial Noise-Aided Downlink Training

Nam-Phong Nguyen, Hien Quoc Ngo, Trung Q. Duong, Hoang Duong Tuan, and Kamel Tourki

***Abstract*—This paper considers a massive MIMO network that includes one multiple-antenna base station, one multiple-antenna eavesdropper, and K single-antenna users. The eavesdropper operates in passive mode and tries to overhear the confidential information from one of the users in the down-link transmission. In order to secure the confidential information, two artificial noise (AN)-aiding schemes are proposed. In the first scheme, AN is injected into the downlink training signals to prevent the eavesdropper from obtaining the correct channel state information of the eavesdropping link. In the second scheme, AN is deployed in both downlink training phase and payload data transmission phase to further degrade the eavesdropping channel. Analytical expressions and tight approximations of the achievable secrecy rate of the considered systems are derived with taking imperfect channel estimation and two types of precoding, i.e., maximum-ratio-transmission and zero-forcing, into consideration. Optimization algorithms for power allocation are proposed to enhance the secrecy performance of the proposed AN-aiding schemes. The results reveal that deploying AN in the downlink training phase of massive MIMO networks does not affect the downlink channel estimation process at users while enabling the system to suppress the downlink channel estimation process at eavesdropper. As a consequence, the proposed AN-aided schemes improve the system performance significantly. Furthermore, implementing AN in both phases allows the considered system having a flexible solution to maximize its secrecy performance at the price of higher complexity.**

***Index Terms*—Physical layer security, massive MIMO, artificial noise.**

I. Introduction

The booming of wireless communication demands huge efforts in securing information. Exploiting the broadcast nature of wireless channels, adversary can easily intercept the confidential messages. The conventional method for security is to implement cryptographic encryption in the application layers. However, this approach is potentially vulnerable to malicious attack because it is based on assumptions of computational complexity [1]. Recently, physical layer security (PLS) [2] has attracted a broad attention from the research community as a complement to the traditional cryptographic encryption. PLS takes advantage of the randomness of wireless channels to enhance the secrecy performance of wireless communication.

PLS concept was first outlined in [3] in which a network consisting of an information source, an intended receiver, and a passive eavesdropper was considered. It has shown that the perfect secured transmission between Alice and Bob can be achieved as long as the condition of the legitimate channel is better than that of the eavesdropping channel. Usually, eavesdroppers operate in the passive mode to prevent the legitimate side from obtaining their channel state information (CSI) [4]. There have been methods to improve the condition of the legitimate channels and degrade the eavesdropping channel's quality, e.g., using multiple antennas and/or using AN-aiding source [5]. Combining multiple-antennas and AN can further enhance the secrecy performance when precoding is implemented to cancel the interference at the legitimate users while suppressing the illegitimate channels [5]–[7].

Recently, massive multiple-input multiple-output (MIMO) network has attracted a lot of attention from the research community and has become the key candidate for next generation wireless networks [8]–[10]. By using a very large number of antennas to serve several users simultaneously in the same frequency band, massive MIMO network offers great power efficiency and spectral efficiency [11]–[13]. In massive MIMO, with time-duplex division (TDD), the number of antennas at the base station (BS) does not affect the resources needed for the channel estimation. Therefore, TDD is more preferable than frequency-duplex division (FDD). In TDD operation, the BS and users exchange their CSI over uplink and dowlink training phases. In uplink training, the users send their pilots to the BS. The BS estimates the channel based on the knowledge of the pilots and then creates the precoding matrix. In dowlink training, the BS processes beamforming pilot to users and then users estimate the effective channels. These effective channels are used in decoding information in the payload data transmission phase. However, these processes enable the illegitimate side to gain CSI of the eavesdropping channels. As a consequence, the eavesdropper can take advantage on this knowledge to successfully decode the transmitted information from the base station in the downlink transmission. Therefore, protecting the training phases from exposing to the illegitimate side is crucial [14], where deploying AN in protecting downlink training phase of massive MIMO network is promising.

Nam-Phong Nguyen, Hien Quoc Ngo, and Trung Q. Duong are with Queen's University Belfast, U.K. (e-mail: {pnguyen04, hien.ngo, trung.q.duong}@qub.ac.uk).

Hoang Duong Tuan is with University of Technology Sydney, Australia (e-mail: tuan.hoang@uts.edu.au).

Kamel Tourki is with France Research Center, Huawei Technologies Co. (e-mail: kamel.tourki@gmail.com).

This work was supported in part by the U.K. Royal Academy of Engineering Research Fellowship under Grant RF1415/14/22 and by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/P019374/1. In addition, this work was supported by a Research Environment Links grant, ID 339568416, under the Newton Programme Vietnam partnership. The grant is funded by the UK Department of Business, Energy and Industrial Strategy (BEIS) and delivered by the British Council. For further information, please visit www.newtonfund.ac.uk.

A. Related Works

There have been several studies on PLS in massive MIMO network using AN. In [15], a hybrid spatial and temporal AN scheme to protect the confidential information in a masive MIMO OFDM system was proposed. In this work, the spatial AN is injected to the precoded signal and also transmitted into the orthogonal direction of the information vector. The paper [16] investigated the secrecy performance of a massive MIMO system in the presence of one multi-antennas eavesdropper and hardware impairment effects. A generalized null-space AN in payload data transmission phase was proposed to deal with the effect of hardware impairment. In [17], the authors proposed a low complexity joint data and AN precoding scheme in a massive MIMO network with limited number of RF chains. Various AN schemes for secrecy enhancement in massive MIMO systems with distributed antennas were proposed in [18]. The authors in [19] suggested a symbol phase rotated scheme to protect the massive MIMO system from a massive MIMO eavesdropper. The results have shown that the proposed scheme can prevent the massive MIMO eavesdropper from recovering most of the transmitted symbols. In [20], various data precoders and AN precoders were proposed to secure the multi-cell massive MIMO system when the eavesdropper's CSI is unavailable. In [21], a phase-only zero-forcing (ZF) AN scheme was studied to reduce the complexity in computing the conventional ZF AN for securing the massive MIMO network. In [22], the authors investigated the performance of an AN jamming-aided scheme at the transmitter over Rician channels. Furthermore, the authors in [23] used null-space and random AN generated from the spare antennas at the BS for securing the multi-cell massive MIMO system. However, to the best of the authors' knowledge, there is no work investigating the AN-based scheme in the downlink training phase of massive MIMO.

B. Contributions

In this paper, we consider a massive MIMO downlink in the presence of a multiple-antenna eavesdropper. The main contributions of this paper are as follows:

- We propose two AN-aiding schemes, i.e., downlink training phase AN-aiding scheme and both phases (downlink training phase and payload data transmission phase) AN-aiding scheme to secure the massive MIMO network in the presence of a multiple-antenna eavesdropper. Furthermore, the effect of imperfect channel estimation is also considered.
- In order to study the behaviors of the proposed AN-aiding schemes on the massive MIMO network, we develop analytical expressions and tight approximations of the achievable secrecy rate of the considered system in two cases where the system uses maximum ratio transmission (MRT) precoding and implements ZF precoding while taking into consideration the influence of imperfect channel estimation.
- Optimization algorithms of extremely low-complexity for power allocation are also developed to enhance the secrecy performance of the two proposed AN-aiding schemes. By applying these algorithms, we show that the secrecy performance of the considered system is significantly enhanced. The results show that deploying AN in the downlink training phase of massive MIMO networks can confuse the eavesdropper's channel estimation while staying harmless to the channel estimation process at the users. In addition, using AN in both phases offers the system a flexible solution to maximize its secrecy performance at the price of a higher complexity in the power allocation process.

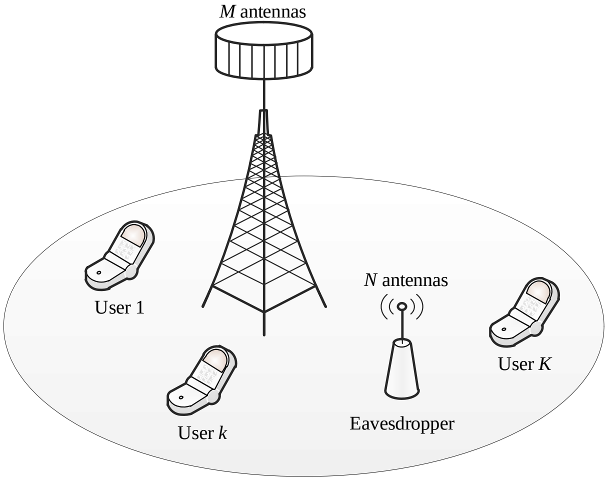


Fig. 1: System model.

The rest of this paper is organized as follows. The system and channel models are described in Section II. The analytical expressions and tight approximations for the achievable secrecy rate of the considered system with the two proposed AN-aiding schemes are developed in Section III and Section IV. In Section V, optimization algorithms for power allocation are proposed. The numerical results based on Monte-Carlo methods and discussions are presented in Section VI. Finally, we conclude our paper in Section VII.

Notations: Subscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ stand for the transpose, the conjugate, and the conjugate transpose, respectively. The expectation operation, variance operation, and Euclidean norm are denoted by $(\cdot)^*$, $\mathbb{E}\{\cdot\}$, $\mathrm{var}(\cdot)$, and $\|\cdot\|$, respectively. $\mathbb{C}^{M \times N}$ represents the space of all $M \times N$ matrices with complex-valued elements. $\boldsymbol{I}_N$ denotes the N-dimensional identity matrix. $\mathcal{CN}(\mu, \sigma^2)$ indicates complex normal distribution with μ mean and σ^2 variance.

II. System and Channel Models

Consider a massive MIMO network which includes one M-antenna BS, K single-antenna end users (EUs), and one N-antenna eavesdropper as described in Fig. 1.

In this system, the BS transmits information to the k-th EU, U_k, over channel vector $\sqrt{\beta_k}\boldsymbol{h}_k$, where $\boldsymbol{h}_k \in \mathbb{C}^{M \times 1}$ is the small-scale fading vector, $\boldsymbol{h}_k \sim \mathcal{CN}(0, \boldsymbol{I}_M)$, and β_k models the large-scale fading. We assume that all the K users share the same time-frequency resource.

The eavesdropper tries to eavesdrop the confidential information from the BS through the channel matrix $\sqrt{\beta_{\mathsf{E}}}\boldsymbol{G} \in$

$\mathbb{C}^{M\times N}$, where the elements of $\boldsymbol{G}$ follow i.i.d. $\mathcal{CN}(0,1)$, and β_{E} models the large-scale fading.

A. Uplink training

All K EUs simultaneously send their orthogonal pilots to the BS for CSI acquisition. The received signal at the BS is

$$\boldsymbol{Y}_{\mathsf{BS}} = \sum_{k}^{K} \sqrt{\rho_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k}\boldsymbol{h}_k\boldsymbol{\omega}_k^H + \boldsymbol{N}_{\mathsf{BS}}, \tag{1}$$

where $\boldsymbol{\omega}_k \in \mathbb{C}^{\tau_{\mathsf{u}}\times 1}$ are the orthogonal pilot sequences of the length τ_{u}, i.e., $\boldsymbol{\omega}_k^H\boldsymbol{\omega}_k = 1$, $\boldsymbol{\omega}_j^H\boldsymbol{\omega}_k = 0$ when $j \neq k$, ρ_{u} is the transmit power at EU, and $\boldsymbol{N}_{\mathsf{BS}} \in \mathbb{C}^{M\times\tau_{\mathsf{u}}}$ is the AWGN at the BS with its elements following $\mathcal{CN}(0,\sigma_0^2)$.

Being known at the BS, $\boldsymbol{\omega}_k$ is used to get $\check{\boldsymbol{y}}_{\mathsf{BS}}$ such that

$$\underbrace{\boldsymbol{Y}_{\mathsf{BS}}\boldsymbol{\omega}_k}_{\check{\boldsymbol{y}}_{\mathsf{BS}}} = \sqrt{\rho_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k}\boldsymbol{h}_k + \underbrace{\boldsymbol{N}_{\mathsf{BS}}\boldsymbol{\omega}_k}_{\boldsymbol{n}_{\mathsf{BS}}}, \tag{2}$$

where $\boldsymbol{n}_{\mathsf{BS}} \sim \mathcal{CN}(0,\sigma_0^2\boldsymbol{I}_{\tau_{\mathsf{u}}})$.

From (2), the BS deploys the MMSE technique to estimate the channel between the BS and U_k. The channel estimate $\hat{\boldsymbol{h}}_k$ is given by

$$\begin{aligned}\hat{\boldsymbol{h}}_k &= \Lambda_{\boldsymbol{h}_k,\check{\boldsymbol{y}}_{\mathsf{BS}}}\Lambda^{-1}_{\check{\boldsymbol{y}}_{\mathsf{BS}},\check{\boldsymbol{y}}_{\mathsf{BS}}}\check{\boldsymbol{y}}_{\mathsf{BS}}\\ &= \frac{\rho_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k}{\rho_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k+\sigma_0^2}\boldsymbol{h}_k + \frac{\sqrt{\rho_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k}}{\rho_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k+\sigma_0^2}\boldsymbol{n}_{\mathsf{BS}},\end{aligned} \tag{3}$$

where $\Lambda_{\boldsymbol{x},\boldsymbol{y}}$ is the covariance matrix of random vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. The estimation error is defined as

$$\tilde{\boldsymbol{h}}_k = \boldsymbol{h}_k - \hat{\boldsymbol{h}}_k. \tag{4}$$

From the MMSE's properties, $\tilde{\boldsymbol{h}}_k$ and $\hat{\boldsymbol{h}}_k$ are mutually independent. Besides, we also have $\tilde{\boldsymbol{h}}_k \sim \mathcal{CN}\left(0,\boldsymbol{I}_M(1-\sigma^2_{\mathsf{u,k}})\right)$, $\hat{\boldsymbol{h}}_k \sim \mathcal{CN}\left(0,\sigma^2_{\mathsf{u,k}}\boldsymbol{I}_M\right)$, and $\sigma_{\mathsf{u,k}} = \sqrt{\frac{\gamma_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k}{\gamma_{\mathsf{u}}\tau_{\mathsf{u}}\beta_k+1}}$ where $\gamma_{\mathsf{u}} = \frac{\rho_{\mathsf{u}}}{\sigma_0^2}$.

B. Downlink Data Transmission

After performing the channel estimation, BS uses the estimated channels to precode the symbols intended for all K users. To enhance the secrecy performance, AN is injected to the transmit data as in [1]. Consequently, the transmitted signal vector at the BS is given by

$$\boldsymbol{S}_{\mathsf{BS}} = \sqrt{\rho_{\mathsf{d}}}\left(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{W}\boldsymbol{x} + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{J}\boldsymbol{\lambda}\right), \tag{5}$$

where $\alpha_{\mathsf{p}}^{\mathsf{IT}}$ is the transmit power ratio for data, $\alpha_{\mathsf{n}}^{\mathsf{IT}}$ is the transmit power ratio for AN, $\boldsymbol{W} = [\boldsymbol{w}_1 ... \boldsymbol{w}_K] \in \mathbb{C}^{M\times K}$ is the precoding matrix, $\boldsymbol{J} = [\boldsymbol{j}_1 ... \boldsymbol{j}_K] \in \mathbb{C}^{M\times K}$ is the AN matrix, $\|\boldsymbol{j}_k\| = 1$, $\boldsymbol{x} = [x_1 ... x_K]$, where $\mathbb{E}\left\{|x_k|^2\right\} = 1$, is the vector of K symbols intended for the K users, and $\boldsymbol{\lambda} = [\lambda_1 ... \lambda_K]$, where $\mathbb{E}\left\{|\lambda_k|^2\right\} = 1$, is the AN vector which is independent of x. The AN matrix lies in the null-space of the estimated legitimate channels, i.e., $\hat{\boldsymbol{H}}^T\boldsymbol{J} = \boldsymbol{0}$.

From (5), the transmit power condition for downlink is $\mathbb{E}\left\{\|\boldsymbol{S}_{\mathsf{BS}}\|^2\right\} \leq \rho_{\mathsf{d}}$. As a result,

$$\begin{aligned}&\alpha_{\mathsf{p}}^2\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{W}\boldsymbol{W}^H)\right\} + \alpha_{\mathsf{n}}^2\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{J}\boldsymbol{J}^H)\right\}\\ &+ \alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{W}\boldsymbol{J}^H)\right\} + \alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{J}\boldsymbol{W}^H)\right\} \leq 1.\end{aligned} \tag{6}$$

Note that $\alpha_{\mathsf{n}}^{\mathsf{IT}} = 0$ corresponds to the case where no AN is added in the data transmission phase.

1) At End Users: The received signals at U_k is

$$\begin{aligned}y_k &= \sqrt{\beta_k}\boldsymbol{h}_k^T\boldsymbol{S}_{\mathsf{BS}} + n_k\\ &= \sqrt{\rho_{\mathsf{d}}\beta_k}\boldsymbol{h}_k^T\left(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{W}\boldsymbol{x} + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{J}\boldsymbol{\lambda}\right) + n_k\\ &= \sqrt{\rho_{\mathsf{d}}\beta_k}\boldsymbol{h}_k^T(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{w}_k x_k + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{j}_k\lambda_k)\\ &+ \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}\beta_k}\boldsymbol{h}_k^T(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{w}_l x_l + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{j}_l\lambda_l) + n_k\\ &= \sqrt{\rho_{\mathsf{d}}\beta_k}(\alpha_{\mathsf{p}}^{\mathsf{IT}}a_{kk}x_k + \alpha_{\mathsf{n}}^{\mathsf{IT}}e_{kk}\lambda_k)\\ &+ \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}\beta_k}(\alpha_{\mathsf{p}}^{\mathsf{IT}}a_{kl}x_l + \alpha_{\mathsf{n}}^{\mathsf{IT}}e_{kl}\lambda_l) + n_k,\end{aligned} \tag{7}$$

where $a_{kk} = \boldsymbol{h}_k^T\boldsymbol{w}_k$, $a_{kl} = \boldsymbol{h}_k^T\boldsymbol{w}_l$, $e_{kk} = \tilde{\boldsymbol{h}}_k^T\boldsymbol{j}_k$, $e_{kl} = \tilde{\boldsymbol{h}}_k^T\boldsymbol{j}_l$, and $n_k \sim \mathcal{CN}(0,\sigma_0^2)$ is AWGN at U_k. Since $\|j_l\|^2 = 1$, e_{kk} and e_{kl} follow $\mathcal{CN}(0,1-\sigma^2_{\mathsf{u,k}})$.

2) At the Eavesdroppers: The received signal at the eavesdropper is

$$\begin{aligned}\boldsymbol{y}_{\mathsf{E}} &= \sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\boldsymbol{G}^T\boldsymbol{S}_{\mathsf{BS}} + \boldsymbol{n}_{\mathsf{E}}\\ &= \sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\boldsymbol{G}^T(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{w}_k x_k + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{j}_k\lambda_k)\\ &+ \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\boldsymbol{G}^T(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{w}_l x_l + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{j}_l\lambda_l) + \boldsymbol{n}_{\mathsf{E}}\\ &= \sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{c}_k x_k + \sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{u}_k\lambda_k\\ &+ \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{c}_l x_l + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{u}_l\lambda_l) + \boldsymbol{n}_{\mathsf{E}},\end{aligned} \tag{8}$$

where $\boldsymbol{c}_k = \boldsymbol{G}^T\boldsymbol{w}_k$, $\boldsymbol{c}_l = \boldsymbol{G}^T\boldsymbol{w}_l$, $\boldsymbol{u}_k = \boldsymbol{G}^T\boldsymbol{j}_k$, $\boldsymbol{u}_l = \boldsymbol{G}^T\boldsymbol{j}_l$, and $\boldsymbol{n}_{\mathsf{E}} \sim \mathcal{CN}(0,\sigma_{\mathsf{E}}^2\boldsymbol{I}_N)$ is the noise at the eavesdropper. Since $\|j_l\|^2 = 1$, $\boldsymbol{u}_k$ and $\boldsymbol{u}_l$ follow $\mathcal{CN}(0,\boldsymbol{I}_N)$.

C. Downlink Training

To detect the desired signal, the k-th UE needs to estimate the effective channel gain a_{kk}. Typically, the BS beamforms the pilots using the precoding matrix in Section II-B. Then from the received pilot signals, each EU will estimate its corresponding effective channel gain [12]. During this phase, the eavesdropper can also estimate its desired effective channel (between the BS and the eavesdropper) in order to improve its ability of eavesdropping the confidential information from the BS. In this work, we propose to use AN during the training phase to enhance the secrecy performance. With our proposed scheme, AN is added into the training signal to contaminate the channel estimation at the eavesdropper. More precisely, the pilot signal beamformed from the BS is given by

$$\boldsymbol{X}_{\mathsf{BS}} = \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}}(\alpha_{\mathsf{p}}\boldsymbol{W} + \alpha_{\mathsf{n}}\boldsymbol{Z})\boldsymbol{\Phi}^H, \tag{9}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{\tau_{\mathsf{d}}\times K}$ represents pilot sequences for K EUs, each of them has length τ_{d}, $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1 ... \boldsymbol{\phi}_K]$, $\boldsymbol{\Phi}^H\boldsymbol{\Phi} = \boldsymbol{I}_K$, $\boldsymbol{Z} = [\boldsymbol{z}_1 ... \boldsymbol{z}_K] \in \mathbb{C}^{M\times K}$, where $\|\boldsymbol{z}_k\|^2 = 1$, is the AN matrix, α_{p} is the transmit power ratio for pilot, α_{n} is the transmit power ratio for AN, and ρ_{d} is the transmit power at the BS. The

AN matrix is implemented in the null-space of the estimated downlink channels, where $\boldsymbol{z}_k$ is chosen to satisfy

$$\hat{\boldsymbol{h}}_k^T \boldsymbol{z}_k = 0. \tag{10}$$

Using (9), the transmit power condition for downlink pilot transmission is $\mathbb{E}\left\{\|\boldsymbol{X}_{\mathsf{BS}}\|^2\right\} \leq \tau_{\mathsf{d}}\rho_{\mathsf{d}}$. As a result,

$$\begin{aligned}&\alpha_{\mathsf{p}}^2\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{W}\boldsymbol{W}^H)\right\} + \alpha_{\mathsf{n}}^2\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{Z}\boldsymbol{Z}^H)\right\}\\ &+ \alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{W}\boldsymbol{Z}^H)\right\} + \alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{Z}\boldsymbol{W}^H)\right\} \leq 1.\end{aligned} \tag{11}$$

Note that $\alpha_{\mathsf{n}} = 0$ corresponds to the conventional case where no AN is added [12].

1) At End Users: The received signal at U_k is

$$\begin{aligned}\boldsymbol{y}_{\mathsf{d},k}^T &= \sqrt{\beta_k}\boldsymbol{h}_k^T\boldsymbol{X}_{\mathsf{BS}} + \boldsymbol{n}_k^T\\ &= \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\boldsymbol{h}_k^T(\alpha_{\mathsf{p}}\boldsymbol{W} + \alpha_{\mathsf{n}}\boldsymbol{Z})\boldsymbol{\Phi}^H + \boldsymbol{n}_k^T,\end{aligned} \tag{12}$$

where $\boldsymbol{n}_k \sim \mathcal{CN}(0, \sigma_0^2\boldsymbol{I}_{\tau_{\mathsf{d}}})$ is the noise at U_k.

Since $\boldsymbol{\phi}_k$ is known at U_k, U_k can process the received signal as follows:

$$\begin{aligned}\underbrace{\boldsymbol{y}_{\mathsf{d},k}^T\boldsymbol{\phi}_k}_{\check{y}_{\mathsf{d},k}} &= \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\boldsymbol{h}_k^T(\alpha_{\mathsf{p}}\boldsymbol{W} + \alpha_{\mathsf{n}}\boldsymbol{Z})\boldsymbol{\Phi}^H\boldsymbol{\phi}_k + \underbrace{\boldsymbol{n}_k^T\boldsymbol{\phi}_k}_{\tilde{n}_k}\\ &= \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\alpha_{\mathsf{p}}\boldsymbol{h}_k^T\boldsymbol{w}_k + \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\alpha_{\mathsf{n}}\boldsymbol{h}_k^T\boldsymbol{z}_k + \tilde{n}_k\\ &= \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\alpha_{\mathsf{p}}a_{kk} + \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\alpha_{\mathsf{n}}b_{kk} + \tilde{n}_k,\end{aligned} \tag{13}$$

where $b_{kk} = \boldsymbol{h}_k^T\boldsymbol{z}_k$, b_{kk} follows $\mathcal{CN}(0, 1-\sigma_{\mathsf{u,k}}^2)$. U_k estimates the effective channel a_{kk} for decoding confidential messages in IT phase. EUs deploys the MMSE to estimate a_{kk}. The estimated effective channel, i.e., $\hat{a}_{kk}$ is given as [24]

$$\hat{a}_{kk} = \mathbb{E}\left\{a_{kk}\right\} + \Lambda_{a_{kk},\check{y}_{\mathsf{d},k}}\Lambda_{\check{y}_{\mathsf{d},k},\check{y}_{\mathsf{d},k}}^{-1}(\check{y}_{\mathsf{d},k} - \mathbb{E}\left\{\check{y}_{\mathsf{d},k}\right\}), \tag{14}$$

where

$$\begin{aligned}\Lambda_{a_{kk},\check{y}_{\mathsf{d},k}} = &\sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\left[\alpha_{\mathsf{p}}\,\mathrm{var}(a_{kk}) + \alpha_{\mathsf{n}}\mathbb{E}\left\{a_{kk}b_{kk}^*\right\}\right.\\ &\left.-\alpha_{\mathsf{n}}\mathbb{E}\left\{a_{kk}\right\}\mathbb{E}\left\{b_{kk}^*\right\}\right],\end{aligned} \tag{15}$$

and

$$\begin{aligned}&\Lambda_{\check{y}_{\mathsf{d},k},\check{y}_{\mathsf{d},k}} = \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right]\\ &+ \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\left(\mathbb{E}\left\{a_{kk}b_{kk}^*\right\} + \mathbb{E}\left\{b_{kk}a_{kk}^*\right\}\right)\\ &- \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\left(\mathbb{E}\left\{a_{kk}\right\}\mathbb{E}\left\{b_{kk}^*\right\} + \mathbb{E}\left\{b_{kk}\right\}\mathbb{E}\left\{a_{kk}^*\right\}\right) + \sigma_0^2.\end{aligned} \tag{16}$$

2) At Eavesdropper: During the downlink training, the eavesdropper intercepts the training information. The received signal at the eavesdropper is given as

$$\begin{aligned}\boldsymbol{Y}_{\mathsf{d,E}} &= \boldsymbol{G}^T\boldsymbol{X}_{\mathsf{BS}} + \boldsymbol{N}_{\mathsf{E}}\\ &= \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\boldsymbol{G}^T(\alpha_{\mathsf{p}}\boldsymbol{W} + \alpha_{\mathsf{n}}\boldsymbol{Z})\boldsymbol{\Phi}^H + \boldsymbol{N}_{\mathsf{E}},\end{aligned} \tag{17}$$

where $\boldsymbol{N}_{\mathsf{E}} \in \mathbb{C}^{N\times\tau_{\mathsf{d}}}$ is the noise at the eavesdropper, elements of $\boldsymbol{N}_{\mathsf{E}}$ follow $\mathcal{CN}(0, \sigma_{\mathsf{E}}^2)$.

Since the eavesdropper can easily have the knowledge of $\boldsymbol{\Phi}$, it can process the received signal as follows:

$$\underbrace{\boldsymbol{Y}_{\mathsf{d,E}}\boldsymbol{\Phi}}_{\check{\boldsymbol{Y}}_{\mathsf{d,E}}} = \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\boldsymbol{G}^T(\alpha_{\mathsf{p}}\boldsymbol{W} + \alpha_{\mathsf{n}}\boldsymbol{Z}) + \underbrace{\boldsymbol{N}_{\mathsf{E}}\boldsymbol{\Phi}}_{\check{\boldsymbol{N}}_{\mathsf{E}}}. \tag{18}$$

The eavesdropped information related to U_k can be demonstrated as

$$\begin{aligned}\check{\boldsymbol{y}}_{\mathsf{d,E},k} &= \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\boldsymbol{G}^T(\alpha_{\mathsf{p}}\boldsymbol{w}_k + \alpha_{\mathsf{n}}\boldsymbol{z}_k) + \check{\boldsymbol{n}}_{\mathsf{E},k}\\ &= \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{p}}\boldsymbol{c}_k + \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{n}}\boldsymbol{d}_k + \check{\boldsymbol{n}}_{\mathsf{E},k}.\end{aligned} \tag{19}$$

where $\boldsymbol{d}_k = \boldsymbol{G}^T\boldsymbol{z}_k$.

The eavesdropper uses the MMSE to estimate $\boldsymbol{c}_k$ for decoding confidential messages in the payload data transmission phase. The estimated effective channel at the eavesdropper is

$$\widehat{\boldsymbol{c}_k} = \mathbb{E}\left\{\boldsymbol{c}_k\right\} + \Lambda_{\boldsymbol{c}_k,\check{\boldsymbol{y}}_{\mathsf{d,E},k}}\Lambda_{\check{\boldsymbol{y}}_{\mathsf{d,E},k},\check{\boldsymbol{y}}_{\mathsf{d,E},k}}^{-1}\left(\check{\boldsymbol{y}}_{\mathsf{d,E},k} - \mathbb{E}\left\{\check{\boldsymbol{y}}_{\mathsf{d,E},k}\right\}\right), \tag{20}$$

where

$$\begin{aligned}\Lambda_{\boldsymbol{c}_k,\check{\boldsymbol{y}}_{\mathsf{d,E},k}} = &\sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\left(\alpha_{\mathsf{p}}\Lambda_{\boldsymbol{c}_k,\boldsymbol{c}_k} + \alpha_{\mathsf{n}}\mathbb{E}\left\{\boldsymbol{c}_k\boldsymbol{d}_k^H\right\}\right.\\ &\left.-\alpha_{\mathsf{n}}\mathbb{E}\left\{\boldsymbol{c}_k\right\}\mathbb{E}\left\{\boldsymbol{d}_k^H\right\}\right),\end{aligned} \tag{21}$$

and

$$\begin{aligned}&\Lambda_{\check{\boldsymbol{y}}_{\mathsf{d,E},k},\check{\boldsymbol{y}}_{\mathsf{d,E},k}} = \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}\left(\alpha_{\mathsf{p}}^2\Lambda_{\boldsymbol{c}_k,\boldsymbol{c}_k} + \alpha_{\mathsf{n}}^2\Lambda_{\boldsymbol{d}_k,\boldsymbol{d}_k}\right)\\ &+ \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}\alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\left(\mathbb{E}\left\{\boldsymbol{c}_k\boldsymbol{d}_k^H\right\} + \mathbb{E}\left\{\boldsymbol{d}_k\boldsymbol{c}_k^H\right\}\right)\\ &- \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}\alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\left(\mathbb{E}\left\{\boldsymbol{c}_k\right\}\mathbb{E}\left\{\boldsymbol{d}_k^H\right\} + \mathbb{E}\left\{\boldsymbol{d}_k\right\}\mathbb{E}\left\{\boldsymbol{c}_k^H\right\}\right) + \sigma_{\mathsf{E}}^2\boldsymbol{I}_N.\end{aligned} \tag{22}$$

D. Choosing Beamforming Matrix

In this paper, we consider two well-known precoding methods, i.e., MRT and ZF, which are widely implemented in massive MIMO systems. Being simple for implementation, MRT suffers from intra-interference, i.e., the undesired signals among users. Meanwhile, ZF precoding can effectively eliminate undesired signals among users at the price of high complexity.

1) MRT: For MRT, the precoding matrix is

$$\boldsymbol{w}_k^{\mathsf{MRT}} = \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}. \tag{23}$$

Following Appendix A, we have

$$\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}\right\} = \sigma_{\mathsf{u,k}}\frac{\Gamma\left(\frac{2M+1}{2}\right)}{\Gamma(M)}, \tag{24}$$

$$\mathbb{E}\left\{|a_{kk}^{\mathsf{MRT}}|^2\right\} = (M-1)\sigma_{\mathsf{u,k}}^2 + 1, \text{ and} \tag{25}$$

$$\mathrm{var}(a_{kk}^{\mathsf{MRT}}) = (M-1)\sigma_{\mathsf{u,k}}^2 + 1 - \sigma_{\mathsf{u,k}}^2\left(\frac{\Gamma\left(\frac{2M+1}{2}\right)}{\Gamma(M)}\right)^2. \tag{26}$$

2) ZF: For ZF, the precoding matrix is chosen to null out the undesired signals from multi-users with the knowledge of the channel estimated during the training phase. The ZF precoding matrix is

$$\boldsymbol{w}_k = \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}, \tag{27}$$

where $\boldsymbol{v}_k$ is the k-th column of $\boldsymbol{V}$:

$$\boldsymbol{V} = \hat{\boldsymbol{H}}^*(\hat{\boldsymbol{H}}^T\hat{\boldsymbol{H}}^*)^{-1}. \tag{28}$$

Because $\hat{\boldsymbol{H}}^T\boldsymbol{V} = \boldsymbol{I}$, we have:

$$\begin{cases} \hat{\boldsymbol{h}}_k^T \boldsymbol{v}_k = 1, \\ \hat{\boldsymbol{h}}_k^T \boldsymbol{v}_l = 0, \quad l \neq k. \end{cases} \tag{29}$$

Therefore,

$$\mathbb{E}\left\{a_{kk}^{\mathsf{ZF}}\right\} = \frac{\Gamma(M-K+\frac{3}{2})}{\sigma_{\mathsf{u,k}}\Gamma(M-K+1)}, \tag{30}$$

$$\mathbb{E}\left\{|a_{kk}^{\mathsf{ZF}}|^2\right\} = \frac{M+1-K}{\sigma_{\mathsf{u,k}}^2} + 1 - \sigma_{\mathsf{u,k}}^2, \text{ and} \tag{31}$$

$$\mathrm{var}(a_{kk}^{\mathsf{ZF}}) = \frac{M+1-K}{\sigma_{\mathsf{u,k}}^2} + 1 - \sigma_{\mathsf{u,k}}^2 - \left(\frac{\Gamma(M-K+\frac{3}{2})}{\sigma_{\mathsf{u,k}}\Gamma(M-K+1)}\right)^2. \tag{32}$$

The detailed proof of (32) is shown in Appendix A.

Remark 1: From (26) and (32), it is observed that

$$\frac{\Gamma(M+\frac{1}{2})}{\Gamma(M)} \overset{M\to\infty}{\to} \sqrt{M}, \tag{33}$$

$$\frac{\Gamma(M-K+\frac{3}{2})}{\Gamma(M-K+1)} \overset{M\to\infty}{\to} \sqrt{M-K+1}. \tag{34}$$

Therefore, $\mathrm{var}(a_{kk}^{\mathsf{MRT,ZF}}) \overset{M\to\infty}{\approx} 1 - \sigma_{\mathsf{u,k}}^2$.

The following conditions hold true for both MRT and ZF. From (11), we have the condition for α_{p} and α_{n} as

$$\begin{aligned} &\alpha_{\mathsf{p}}^2 \underbrace{\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{W}\boldsymbol{W}^H)\right\}}_{=K} + \alpha_{\mathsf{n}}^2 \underbrace{\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{Z}\boldsymbol{Z}^H)\right\}}_{=K} \\ &+ \alpha_{\mathsf{p}}\alpha_{\mathsf{n}} \underbrace{\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{W}\boldsymbol{Z}^H)\right\}}_{=0} + \alpha_{\mathsf{p}}\alpha_{\mathsf{n}} \underbrace{\mathbb{E}\left\{\mathrm{tr}(\boldsymbol{Z}\boldsymbol{W}^H)\right\}}_{=0} = 1 \\ &\to \alpha_{\mathsf{p}}^2 K + \alpha_{\mathsf{n}}^2 K = 1 \to \alpha_{\mathsf{n}}^2 = K^{-1} - \alpha_{\mathsf{p}}^2. \end{aligned} \tag{35}$$

Similarly, from (6), we have the condition for $\alpha_{\mathsf{p}}^{\mathsf{IT}}$ and $\alpha_{\mathsf{n}}^{\mathsf{IT}}$ as

$$(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2 + (\alpha_{\mathsf{n}}^{\mathsf{IT}})^2 = K^{-1}. \tag{36}$$

E. Estimated Effective Channel at End User

1) MRT: Following Appendix B, we obtain $\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}} b_{kk}^*\right\} = 0$ and $\mathbb{E}\left\{b_{kk}(a_{kk}^{\mathsf{MRT}})^*\right\} = 0$. Plugging $\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}} b_{kk}^*\right\}$ and $\mathbb{E}\left\{b_{kk}(a_{kk}^{\mathsf{MRT}})^*\right\}$ into (15) and (16), we have

$$\Lambda_{a_{kk}^{\mathsf{MRT}}, \tilde{y}_{\mathsf{d},k}} = \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\,\alpha_{\mathsf{p}}\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}), \tag{37}$$

and

$$\Lambda_{\tilde{y}_{\mathsf{d},k}, \tilde{y}_{\mathsf{d},k}} = \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right] + \sigma_0^2. \tag{38}$$

The estimate of the effective channel at U_k is

$$\begin{aligned} \hat{a}_{kk}^{\mathsf{MRT}} = &\frac{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}})a_{kk}^{\mathsf{MRT}}}{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right] + \sigma_0^2} \\ &+ \frac{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\alpha_{\mathsf{p}}\alpha_{\mathsf{n}}\,\mathrm{var}(a_{kk}^{\mathsf{MRT}})b_{kk}}{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right] + \sigma_0^2} \\ &+ \frac{\sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k}\,\alpha_{\mathsf{p}}\,\mathrm{var}(a_{kk}^{\mathsf{MRT}})\tilde{n}_k}{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right] + \sigma_0^2} \\ &+ \mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}\right\} - \frac{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}})\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}\right\}}{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_k\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right] + \sigma_0^2}. \end{aligned} \tag{39}$$

From (39), the mean and variance of $\hat{a}_{kk}^{\mathsf{MRT}}$ are

$$\mathbb{E}\left\{\hat{a}_{kk}^{\mathsf{MRT}}\right\} = \mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}\right\}, \tag{40}$$

and

$$\mathrm{var}(\hat{a}_{kk}^{\mathsf{MRT}}) = \frac{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}})^2}{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right] + 1}, \tag{41}$$

where $\gamma_{\mathsf{d}} = \frac{\rho_{\mathsf{d}}\beta_k}{\sigma_0^2}$.

The downlink channel estimation error is defined as

$$\tilde{a}_{kk}^{\mathsf{MRT}} = a_{kk}^{\mathsf{MRT}} - \hat{a}_{kk}^{\mathsf{MRT}}. \tag{42}$$

From (42), the mean and variance of $a_{kk}^{\widetilde{\mathsf{MRT}}}$ are

$$\mathbb{E}\left\{\tilde{a}_{kk}^{\mathsf{MRT}}\right\} = 0, \tag{43}$$

and

$$\mathrm{var}(\tilde{a}_{kk}^{\mathsf{MRT}}) = \frac{\mathrm{var}(a_{kk}^{\mathsf{MRT}})(\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk}) + 1)}{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{MRT}}) + \gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk}) + 1}. \tag{44}$$

2) ZF : Following Appendix B, we have $\mathbb{E}\left\{a_{kk}^{\mathsf{ZF}} b_{kk}^*\right\} = 0$ and $\mathbb{E}\left\{b_{kk}(a_{kk}^{\mathsf{ZF}})^*\right\} = 0$. Similar to the MRT case, we have $\mathbb{E}\left\{\hat{a}_{kk}^{\mathsf{ZF}}\right\} = \mathbb{E}\left\{a_{kk}\right\}$, $\mathbb{E}\left\{\tilde{a}_{kk}^{\mathsf{ZF}}\right\} = 0$,

$$\mathrm{var}(\hat{a}_{kk}^{\mathsf{ZF}}) = \frac{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{ZF}})^2}{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\left[\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{ZF}}) + \alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk})\right] + 1}, \tag{45}$$

and

$$\mathrm{var}(\tilde{a}_{kk}^{\mathsf{ZF}}) = \frac{\mathrm{var}(a_{kk}^{\mathsf{ZF}})(\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk}) + 1)}{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{p}}^2\,\mathrm{var}(a_{kk}^{\mathsf{ZF}}) + \gamma_{\mathsf{d}}\tau_{\mathsf{d}}\alpha_{\mathsf{n}}^2\,\mathrm{var}(b_{kk}) + 1}. \tag{46}$$

Remark 2: From (44) and (46), it is observed that $\mathrm{var}\left(\tilde{a}_{kk}^{\mathsf{MRT,ZF}}\right) \overset{M\to\infty}{\approx} \frac{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}(1-\sigma_{\mathsf{u,k}}^2)+1}{\gamma_{\mathsf{d}}\tau_{\mathsf{d}}(\alpha_{\mathsf{p}}^2+\alpha_{\mathsf{n}}^2)+1}$. From (35), we have condition for AN noise power parameters as $\alpha_{\mathsf{p}}^2 + \alpha_{\mathsf{n}}^2 = 1/K$. It means that the estimation error of the downlink at the legitimate users is independent to AN power parameters when the number of transmit antennas at the BS is large. Besides, $\mathrm{var}\left(\tilde{a}_{kk}^{\mathsf{MRT,ZF}}\right)$ can be reduced by improving the uplink training process and decreasing the number of users.

F. Estimated Effective Channel at the Eavesdropper

Following Appendix C, for both MRT and ZF, we have $\mathbb{E}\left\{\boldsymbol{c}_k\right\} = 0$, $\mathbb{E}\left\{\boldsymbol{d}_k\right\} = 0$, $\Lambda_{\boldsymbol{c}_k,\boldsymbol{c}_k} = \boldsymbol{I}_N$, $\Lambda_{\boldsymbol{d}_k,\boldsymbol{d}_k} = \boldsymbol{I}_N$, $\mathbb{E}\left\{\boldsymbol{c}_k\boldsymbol{d}_k^H\right\} = 0$, and $\mathbb{E}\left\{\boldsymbol{d}_k\boldsymbol{c}_k^H\right\} = 0$.

As a consequence, we have

$$\Lambda_{\boldsymbol{c}_k,\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k}} = \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{p}}\Lambda_{\boldsymbol{c}_k,\boldsymbol{c}_k} = \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{p}}\boldsymbol{I}_N, \tag{47}$$

$$\Lambda_{\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k},\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k}} = \rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}\left(\alpha_{\mathsf{p}}^2 + \alpha_{\mathsf{n}}^2\right)\boldsymbol{I}_N + \sigma_{\mathsf{E}}^2\boldsymbol{I}_N. \tag{48}$$

estimate of the effective channel at the eavesdropper is

$$\begin{aligned}\hat{\boldsymbol{c}_k} &= \mathbb{E}\left\{\boldsymbol{c}_k\right\} + \Lambda_{\boldsymbol{c}_k,\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k}}\Lambda^{-1}_{\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k},\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k}}\left(\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k} - \mathbb{E}\left\{\check{\boldsymbol{y}}_{\mathsf{d},\mathsf{E},k}\right\}\right)\\ &= \frac{\sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{p}}}{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}\left[\alpha_{\mathsf{p}}^2 + \alpha_{\mathsf{n}}^2\right] + \sigma_{\mathsf{E}}^2}\Big(\sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{p}}\boldsymbol{c}_k\\ &\quad + \sqrt{\rho_{\mathsf{d}}\tau_{\mathsf{d}}\beta_{\mathsf{E}}}\alpha_{\mathsf{n}}\boldsymbol{d}_k + \check{\boldsymbol{n}}_{\mathsf{E},k}\Big),\end{aligned} \tag{49}$$

where $\gamma_{\mathsf{E}} = \frac{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}{\sigma_{\mathsf{E}}^2}$. We have $\mathbb{E}\left\{\hat{\boldsymbol{c}_k}\right\} = 0$ and

$$\Lambda_{\hat{\boldsymbol{c}}_k,\hat{\boldsymbol{c}}_k} = \frac{\gamma_{\mathsf{E}}\tau_{\mathsf{d}}\alpha_{\mathsf{p}}^2}{(\gamma_{\mathsf{E}}\tau_{\mathsf{d}}[\alpha_{\mathsf{p}}^2 + \alpha_{\mathsf{n}}^2] + 1)} = \sigma^2_{\hat{\boldsymbol{c}}_k}\boldsymbol{I}_N. \tag{50}$$

Estimation error at the eavesdropper is

$$\tilde{\boldsymbol{c}}_k = \boldsymbol{c}_k - \hat{\boldsymbol{c}}_k, \tag{51}$$

with $\mathbb{E}\left\{\tilde{\boldsymbol{c}}_k\right\} = 0$ and

$$\Lambda_{\tilde{\boldsymbol{c}}_k,\tilde{\boldsymbol{c}}_k} = \frac{(\gamma_{\mathsf{E}}\tau_{\mathsf{d}}\alpha_{\mathsf{n}}^2 + 1)}{(\gamma_{\mathsf{E}}\tau_{\mathsf{d}}[\alpha_{\mathsf{p}}^2 + \alpha_{\mathsf{n}}^2] + 1)} = \sigma^2_{\tilde{\boldsymbol{c}}_k}\boldsymbol{I}_N. \tag{52}$$

Remark 3: The estimation error of the downlink at the eavesdropper depends on the AN power parameters which means that we can adjust the AN power parameter to maximize the advantage for the legitimate side.
The downlink payload data transmission phases of the two proposed AN schemes are described in details in Section III and Section IV.

III. Achievable Secrecy Rate in Downlink Training Artificial Noise Aiding Scheme

In this section, we develop the tight approximations for the achievable secrecy rate of the considered system when the downlink training phase AN-aiding scheme is applied.

A. Achievable Legitimate Rate

1) MRT: In the payload data transmission phase the BS transmits information to the users without AN. By plugging $\boldsymbol{w}_k = \frac{\hat{\boldsymbol{h}}_k^*}{\|\hat{\boldsymbol{h}}_k^*\|}$ into (7) and setting $\alpha_{\mathsf{n}}^{\mathsf{IT}} = 0$, the received signal at U_k is

$$\begin{aligned}y_k &= \sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}\boldsymbol{h}_k^T\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}x_k + \sum_{l\neq k}^{K}\sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}\boldsymbol{h}_k^T\frac{\hat{\boldsymbol{h}}_l^*}{\left\|\hat{\boldsymbol{h}}_l^*\right\|}x_l + n_k\\ &= \sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}a_{kk}^{\mathsf{MRT}}x_k + \sum_{l\neq k}^{K}\sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}a_{kl}^{\mathsf{MRT}}x_l + n_k,\end{aligned} \tag{53}$$

where x_k and x_l are the desired confidential information for U_k and U_l, respectively. Obviously, $a_{kl}^{\mathsf{MRT}} \sim \mathcal{CN}(0,1)$ for $k \neq l$. U_k only has knowledge of $\hat{a}_{kk}^{\mathsf{MRT}}$. The achievable legitimate rate at U_k is formulated in (54) on the top of the next page [25, Eq. (2.46)],

2) ZF: By plugging $\boldsymbol{w}_k = \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}$ into (7) and setting $\alpha_{\mathsf{n}}^{\mathsf{IT}} = 0$, the received signal at U_k is

$$\begin{aligned}y_k &= \sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}\boldsymbol{h}_k^T\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}x_k + \sum_{l\neq k}^{K}\sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}\tilde{\boldsymbol{h}}_k^T\frac{\boldsymbol{v}_l}{\|\boldsymbol{v}_l\|}x_l + n_k\\ &= \sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}a_{kk}^{\mathsf{ZF}}x_k + \sum_{l\neq k}^{K}\sqrt{\frac{\rho_{\mathsf{d}}}{K}\beta_k}a_{kl}^{\mathsf{ZF}}x_l + n_k.\end{aligned} \tag{56}$$

Obviously, $a_{kl}^{\mathsf{ZF}} \sim \mathcal{CN}(0,(1-\sigma^2_{\mathsf{u,k}}))$ for $k \neq l$. Similar to the case of using MRT, the achievable legitimate rate at U_k is expressed in (57) on the next page.

B. Achievable Eavesdropping Rate

From (7), setting $\alpha_{\mathsf{n}}^{\mathsf{IT}} = 0$, the received signal at the eavesdropper in both cases of using MRT and ZF is

$$\boldsymbol{y}_{\mathsf{E}} = \sqrt{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}}\boldsymbol{c}_k x_k + \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}}\boldsymbol{c}_l x_l + \boldsymbol{n}_{\mathsf{E}}. \tag{58}$$

The eavesdropper performs MRC

$$\begin{aligned}y_{\mathsf{E}}^{\mathsf{MRC}} &= \sqrt{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}}\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_k x_k\\ &\quad + \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}}\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_l x_l + \frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{n}_{\mathsf{E}}.\end{aligned} \tag{59}$$

The eavesdropper only has knowledge of $\hat{\mathbf{c}}_k$. Therefore, the achievable eavesdropping rate is described in (60) on the next page.

C. Achievable Secrecy Rate

From (54), (57), and (59), the following lemma is given.

Lemma 1: When downlink training phase AN-aiding scheme is applied, the achievable secrecy rate of the considered system is [1]

$$R_{\mathsf{s,DT}}^{\mathcal{A}} = [R_k^{\mathcal{A}} - R_{\mathsf{E}}]^+, \tag{61}$$

where $\mathcal{A} = \{\mathsf{MRT}, \mathsf{ZF}\}$, $R_k^{\mathcal{A}}$ and R_{E} are given in (54), (57), and (60), respectively.

IV. Achievable Secrecy Rate in Both Phases Artificial Noise Aiding Scheme

In this section, we develop a tight approximation for the achievable secrecy rate of the considered system when the both phases AN-aiding scheme is applied. In this scheme, the AN is added in the downlink training phase as in Section III. In the payload data transmission phase, AN is injected into the null-space of the legitimate channels to confuse the eavesdropper. The AN matrix is chosen based on the knowledge of the estimated legitimate channels, i.e., $\hat{\boldsymbol{H}}^T\boldsymbol{J} = \boldsymbol{0}$, where $\boldsymbol{J} = [\boldsymbol{j}_1 \cdots \boldsymbol{j}_K] \in \mathbb{C}^{M\times K}$, where $\|\boldsymbol{j}_k\|^2 = 1$, is the AN matrix.

[1] This is a lower bound of the real secrecy rate $\mathbb{E}\{[R_k^{\mathcal{A}} - R_{\mathsf{E}}]^+\}$

$$
\begin{aligned}
&R_k^{\mathsf{MRT}} = \\
&\mathbb{E}\left\{\log_2\left(1+\frac{\rho_{\mathsf{d}}\beta_k K^{-1}|\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}} \mid \hat{a}_{kk}^{\mathsf{MRT}}\right\}|^2}{\rho_{\mathsf{d}}\beta_k K^{-1}\operatorname{var}(a_{kk}^{\mathsf{MRT}} \mid \hat{a}_{kk}^{\mathsf{MRT}}) + \rho_{\mathsf{d}}\beta_k K^{-1}\sum\limits_{l\neq k}^{K}\mathbb{E}\left\{|a_{kl}^{\mathsf{MRT}} x_l|^2 \mid \hat{a}_{kk}^{\mathsf{MRT}}\right\} + \mathbb{E}\left\{|n_k|^2\right\}}\right)\right\} \\
&\overset{(a)}{\approx} \mathbb{E}\left\{\log_2\left(1+\frac{\rho_{\mathsf{d}}\beta_k K^{-1}|\hat{a}_{kk}^{\mathsf{MRT}}|^2}{\rho_{\mathsf{d}}\beta_k K^{-1}\operatorname{var}(\tilde{a}_{kk}^{\mathsf{MRT}}) + \rho_{\mathsf{d}}\beta_k K^{-1}\sum\limits_{l\neq k}^{K}\mathbb{E}\left\{|a_{kl}^{\mathsf{MRT}}|^2\right\} + \sigma_0^2}\right)\right\} \\
&\overset{(b)}{\approx} \log_2\left(1+\frac{\gamma_{\mathsf{d}} K^{-1} M \sigma_{\mathsf{u,k}}^2}{\gamma_{\mathsf{d}} K^{-1}\operatorname{var}(\tilde{a}_{kk}^{\mathsf{MRT}}) + \gamma_{\mathsf{d}} K^{-1}(K-1)+1}\right), \qquad (54)
\end{aligned}
$$

where step (b) is obtained by using the law of large number

$$
\frac{1}{M}\|\boldsymbol{v}\|^2 \overset{M\to\infty}{\to} \frac{1}{M}\mathbb{E}\left\{\|\boldsymbol{v}\|^2\right\}, \qquad (55)
$$

where $\boldsymbol{v} \sim \mathcal{CN}(0, \boldsymbol{I}_M)$. The details of step (a) is described in Appendix D.

$$
\begin{aligned}
&R_k^{\mathsf{ZF}} = \\
&\mathbb{E}\left\{\log_2\left(1+\frac{\rho_{\mathsf{d}} K^{-1}\beta_k|\mathbb{E}\left\{a_{kk}^{\mathsf{ZF}} \mid \hat{a}_{kk}^{\mathsf{ZF}}\right\}|^2}{\rho_{\mathsf{d}} K^{-1}\beta_k\operatorname{var}(a_{kk}^{\mathsf{ZF}} \mid \hat{a}_{kk}^{\mathsf{ZF}}) + \rho_{\mathsf{d}} K^{-1}\beta_k\sum\limits_{l\neq k}^{K}\mathbb{E}\left\{|a_{kl}^{\mathsf{ZF}} x_l|^2 \mid \hat{a}_{kk}^{\mathsf{ZF}}\right\} + \mathbb{E}\left\{|n_k|^2\right\}}\right)\right\} \\
&\overset{(a)}{\approx} \mathbb{E}\left\{\log_2\left(1+\frac{\rho_{\mathsf{d}} K^{-1}\beta_k|\hat{a}_{kk}^{\mathsf{ZF}}|^2}{\rho_{\mathsf{d}} K^{-1}\beta_k\operatorname{var}(\tilde{a}_{kk}^{\mathsf{ZF}}) + \rho_{\mathsf{d}} K^{-1}\beta_k\sum\limits_{l\neq k}^{K}\mathbb{E}\left\{|a_{kl}^{\mathsf{ZF}}|^2\right\} + \sigma_0^2}\right)\right\} \\
&\approx \log_2\left(1+\frac{\gamma_{\mathsf{d}} K^{-1}\frac{M-K}{\sigma_{\mathsf{u,k}}^2}}{\gamma_{\mathsf{d}} K^{-1}\operatorname{var}(\tilde{a}_{kk}^{\mathsf{ZF}}) + \gamma_{\mathsf{d}} K^{-1}(K-1)(1-\sigma_{\mathsf{u,k}}^2)+1}\right). \qquad (57)
\end{aligned}
$$

The details of step (a) is demonstrated in Appendix E.

A. Achievable Legitimate Rate

1) MRT: By plugging $\boldsymbol{w}_k = \frac{\hat{\boldsymbol{h}}_k^*}{\|\hat{\boldsymbol{h}}_k^*\|}$ into (7), the received signal at U_k is

$$
\begin{aligned}
y_k &= \sqrt{\rho_{\mathsf{d}}\beta_k}\alpha_{\mathsf{p}}^{\mathsf{IT}} a_{kk}^{\mathsf{MRT}} x_k + \sqrt{\rho_{\mathsf{d}}\beta_k}\alpha_{\mathsf{n}}^{\mathsf{IT}} e_{kk}\lambda_k \\
&\quad + \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}\beta_k}(\alpha_{\mathsf{p}}^{\mathsf{IT}} a_{kl}^{\mathsf{MRT}} x_l + \alpha_{\mathsf{n}}^{\mathsf{IT}} e_{kl}\lambda_l) + n_k. \qquad (62)
\end{aligned}
$$

The achievable rate at U_k is formulated in (63) on the next page.

2) ZF: By plugging $\boldsymbol{w}_k = \frac{\boldsymbol{v}_l}{\|\boldsymbol{v}_l\|}$ into (7), the received signal at U_k is

$$
\begin{aligned}
y_k &= \sqrt{\rho_{\mathsf{d}}\beta_k}\alpha_{\mathsf{p}}^{\mathsf{IT}} a_{kk}^{\mathsf{ZF}} x_k + \sqrt{\rho_{\mathsf{d}}\beta_k}\alpha_{\mathsf{n}}^{\mathsf{IT}} b_{kk}\lambda_k \\
&\quad + \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}\beta_k}(\alpha_{\mathsf{p}}^{\mathsf{IT}} a_{kl}^{\mathsf{ZF}} x_l + \alpha_{\mathsf{n}}^{\mathsf{IT}} b_{kl}\lambda_l) + n_k. \qquad (64)
\end{aligned}
$$

The achievable rate at U_k is in (65) on the next page.

B. Achievable Eavesdropping Rate

The eavesdropper uses MRC to process the received signal in (8). After performing MRC, the processed signal at the eavesdropper is in (66) on the next page.

The achievable eavesdropping rate at the eavesdropper is in (67) on the next page.

C. Achievable Secrecy Rate

From (63), (65), and (67), the following lemma is given.

Lemma 2: When downlink training phase AN-aiding scheme is applied, the achievable rate of the considered system is

$$
R_{\mathsf{s,BP}}^{\mathcal{A}} = [R_k^{\mathcal{A}} - R_{\mathsf{E}}]^+, \qquad (68)
$$

where $[x]^+ = \max(0, x)$, $\mathcal{A} = \{\mathsf{MRT}, \mathsf{ZF}\}$, $R_k^{\mathcal{A}}$ and R_{E} are given in (63), (65), and (67), respectively.

$$\begin{aligned} R_{\mathsf{E}} &= \mathbb{E}\left\{\log_2\left(1+\frac{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}\left|\mathbb{E}\left\{\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_k \mid \hat{\boldsymbol{c}}_k\right\}\right|^2}{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}\operatorname{var}\left(\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_k \mid \hat{\boldsymbol{c}}_k\right)+\sum_{k\neq l}^{K}\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}\mathbb{E}\left\{\left|\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_l\right|^2 \mid \hat{\boldsymbol{c}}_k\right\}+\sigma_{\mathsf{E}}^2}\right)\right\} \\ &= \mathbb{E}\left\{\log_2\left(1+\frac{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}\left\|\hat{\boldsymbol{c}}_k\right\|^2}{\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}\sigma_{\tilde{\boldsymbol{c}}_k}^2+\sum_{k\neq l}^{K}\rho_{\mathsf{d}}K^{-1}\beta_{\mathsf{E}}\mathbb{E}\left\{\left|\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_l\right|^2\right\}+\sigma_{\mathsf{E}}^2}\right)\right\} \\ &= \mathbb{E}\left\{\log_2\left(1+\frac{\gamma_{\mathsf{E}}K^{-1}\left\|\hat{\boldsymbol{c}}_k\right\|^2}{\gamma_{\mathsf{E}}K^{-1}\sigma_{\tilde{\boldsymbol{c}}_k}^2+\gamma_{\mathsf{E}}K^{-1}(K-1)+1}\right)\right\} \\ &\overset{(a)}{\approx} \log_2\left(1+\frac{\gamma_{\mathsf{E}}K^{-1}N\sigma_{\hat{\boldsymbol{c}}_k}^2}{\gamma_{\mathsf{E}}K^{-1}\sigma_{\tilde{\boldsymbol{c}}_k}^2+\gamma_{\mathsf{E}}K^{-1}(K-1)+1}\right), \end{aligned} \tag{60}$$

where step (a) uses the identity (55).

$$R_k^{\mathsf{MRT}} = \log_2\left(1+\frac{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2 M\sigma_{\mathsf{u,k}}^2}{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\operatorname{var}(\tilde{a}_{kk}^{\mathsf{MRT}})+\gamma_{\mathsf{d}}(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2(1-\sigma_{\mathsf{u,k}}^2)+\gamma_{\mathsf{d}}(K-1)[\frac{1}{K}-(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2\sigma_{\mathsf{u,k}}^2]+1}\right), \tag{63}$$

$$R_k^{\mathsf{ZF}} = \log_2\left(1+\frac{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\frac{M-K}{\sigma_{\mathsf{u,k}}^2}}{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\operatorname{var}(\tilde{a}_{kk}^{\mathsf{ZF}})+\gamma_{\mathsf{d}}(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2(1-\sigma_{\mathsf{u,k}}^2)+\gamma_{\mathsf{d}}\frac{(K-1)}{K}(1-\sigma_{\mathsf{u,k}}^2)+1}\right), \tag{65}$$

$$y_{\mathsf{E}}^{\mathsf{MRC}} = \sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{c}_k x_k + \underbrace{\sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{u}_k z_k + \sum_{l\neq k}^{K}\sqrt{\rho_{\mathsf{d}}\beta_{\mathsf{E}}}\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}(\alpha_{\mathsf{p}}^{\mathsf{IT}}\boldsymbol{c}_l x_l + \alpha_{\mathsf{n}}^{\mathsf{IT}}\boldsymbol{u}_l z_l)}_{\Theta} + \frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{n}_{\mathsf{E}}, \tag{66}$$

$$\begin{aligned} R_{\mathsf{E}} &= \mathbb{E}\left\{\log_2\left(1+\frac{\rho_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\beta_{\mathsf{E}}\left|\mathbb{E}\left\{\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_k \mid \hat{\boldsymbol{c}}_k\right\}\right|^2}{\rho_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\beta_{\mathsf{E}}\operatorname{var}\left(\frac{\hat{\boldsymbol{c}}_k^H}{\|\hat{\boldsymbol{c}}_k\|}\boldsymbol{c}_k \mid \hat{\boldsymbol{c}}_k\right)+\mathbb{E}\left\{\left|\Theta\right|^2 \mid \hat{\boldsymbol{c}}_k\right\}+\sigma_{\mathsf{E}}^2}\right)\right\} \\ &= \log_2\left(1+\frac{\gamma_{\mathsf{E}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2 N\sigma_{\hat{\boldsymbol{c}}_k}^2}{\gamma_{\mathsf{E}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\sigma_{\tilde{\boldsymbol{c}}_k}^2+\gamma_{\mathsf{E}}(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2+\gamma_{\mathsf{E}}\frac{(K-1)}{K}+1}\right), \end{aligned} \tag{67}$$

V. Resource Allocation

In this section, optimization algorithms for power allocation are proposed to maximize the achievable secrecy rate of the considered system.

A. AN in Training Phase

1) MRT: Plugging (44) and (35) into (54), the achievable legitimate rate of the considered system when using AN in downlink training phase with MRT can be rewritten as

$$\begin{aligned} R_k^{\mathsf{MRT}} &= \log_2\left(1+\frac{\gamma_{\mathsf{d}}K^{-1}\sigma_{\mathsf{u,k}}^2 M}{\gamma_{\mathsf{d}}K^{-1}\operatorname{var}(\tilde{a}_{kk})+\gamma_{\mathsf{d}}K^{-1}(K-1)+1}\right) \\ &= \log_2\left(1+\frac{A_1\alpha_{\mathsf{p}}^2+A_2}{A_3\alpha_{\mathsf{p}}^2+A_4}\right), \end{aligned} \tag{69}$$

where $A_1 = \gamma_{\mathsf{d}}^2\tau_{\mathsf{d}}K^{-1}M\sigma_{\mathsf{u,k}}^2(\operatorname{var}(a_{kk}^{\mathsf{MRT}}) - \operatorname{var}(b_{kk}))$, $A_2 = (\gamma_{\mathsf{d}}^2 M\sigma_{\mathsf{u,k}}^2\tau_{\mathsf{d}}\operatorname{var}(b_{kk})K^{-2} + \gamma_{\mathsf{d}}K^{-1}M\sigma_{\mathsf{u,k}}^2)$, $A_3 = [\gamma_{\mathsf{d}}K^{-1}(K - 1) + 1]\gamma_{\mathsf{d}}\tau_{\mathsf{d}}(\operatorname{var}(a_{kk}^{\mathsf{MRT}}) - \operatorname{var}(b_{kk})) - \gamma_{\mathsf{d}}^2K^{-1}\tau_{\mathsf{d}}\operatorname{var}(a_{kk}^{\mathsf{MRT}})\operatorname{var}(b_{kk})$, and $A_4 = (\gamma_{\mathsf{d}}\tau_{\mathsf{d}}K^{-1}\operatorname{var}(b_{kk}) + 1)(\gamma_{\mathsf{d}}K^{-1}(K - 1) + \gamma_{\mathsf{d}}K^{-1}\operatorname{var}(a_{kk}^{\mathsf{MRT}}) + 1)$.

$$
\begin{aligned}
R_k^{\mathsf{ZF}} &= \log_2\left(1+\frac{\gamma_{\mathsf{d}}K^{-1}\frac{M-K}{\sigma^2_{\mathsf{u,k}}}}{\gamma_{\mathsf{d}}K^{-1}\operatorname{var}(\tilde{a}_{kk}^{\mathsf{ZF}})+\gamma_{\mathsf{d}}K^{-1}(K-1)(1-\sigma^2_{\mathsf{u,k}})+1}\right)\\
&= \log_2\left(1+\frac{A_1\alpha_{\mathsf{p}}^2+A_2}{A_3\alpha_{\mathsf{p}}^2+A_4}\right), \qquad (70)
\end{aligned}
$$

where $A_1 = \gamma_{\mathsf{d}}^2K^{-1}\tau_{\mathsf{d}}\frac{M-K}{\sigma^2_{\mathsf{u,k}}}(\operatorname{var}(a_{kk}^{\mathsf{ZF}})-\operatorname{var}(b_{kk}))$, $A_2 = \gamma_{\mathsf{d}}K^{-1}\frac{M-K}{\sigma^2_{\mathsf{u,k}}}(\gamma_{\mathsf{d}}\tau_{\mathsf{d}}K^{-1}\operatorname{var}(b_{kk})+1)$, $A_3 = [\gamma_{\mathsf{d}}(K-1)K^{-1}(1-\sigma^2_{\mathsf{u,k}})+1]\gamma_{\mathsf{d}}\tau_{\mathsf{d}}(\operatorname{var}(a_{kk}^{\mathsf{ZF}})-\operatorname{var}(b_{kk}))-\gamma_{\mathsf{d}}^2\tau_{\mathsf{d}}K^{-1}\operatorname{var}(a_{kk}^{\mathsf{ZF}})\operatorname{var}(b_{kk})$, and $A_4 = (\gamma_{\mathsf{d}}\tau_{\mathsf{d}}K^{-1}\operatorname{var}(b_{kk})+1)(\gamma_{\mathsf{d}}K^{-1}\operatorname{var}(a_{kk}^{\mathsf{ZF}})+\gamma_{\mathsf{d}}K^{-1}(K-1)(1-\sigma^2_{\mathsf{u,k}})+1)$.

2) ZF: Plugging (46) and (35) into (57), the achievable legitimate rate of the considered system when using AN in downlink training phase with ZF can be rewritten in (70) on the next page.

3) Achievable Eavesdropping Rate: The achievable eavesdropping rate in (60) is re-arranged as in (71) on the next page.

4) Proposed Optimization Algorithm: The achievable secrecy rate of the considered system when donwlink training AN scheme is applied can be re-formulated as

$$
\begin{aligned}
&R_{\mathsf{s,DT}}^{\mathcal{A}} = [R_k^{\mathcal{A}} - R_{\mathsf{E}}]^+\\
&= \left[\log_2\left(1+\frac{A_1\alpha_{\mathsf{p}}^2+A_2}{A_3\alpha_{\mathsf{p}}^2+A_4}\right)-\log_2\left(1+\frac{A_5\alpha_{\mathsf{p}}^2}{A_6\alpha_{\mathsf{p}}^2+A_7}\right)\right]^+, \qquad (72)
\end{aligned}
$$

where $\mathcal{A} = \{\mathsf{MRT},\mathsf{ZF}\}$. To maximize the secrecy rate, the transmit power ratio for pilot/AN transmission needs to be optimized. The proposed problem is

$$
\max_{\alpha_{\mathsf{p}}}\ R_{\mathsf{s,DT}}^{\mathcal{A}} \quad \text{s. t.} \quad 0 \le \alpha_{\mathsf{p}}^2 \le \frac{1}{K}. \qquad (73)
$$

Making the variable change $\alpha_{\mathsf{p}}^2 \rightarrow x$, problem (73) is equivalent to the following problem:

$$
\begin{aligned}
\max_{x} \quad & f(x) := \ln(a_1+a_2x)+\ln(a_3+a_4x)\\
& -\ln(a_5+a_6x)-\ln(a_7+a_8x) \qquad (74a)\\
\text{s. t.} \quad & 0 \le x \le \frac{1}{K}, \qquad (74b)
\end{aligned}
$$

where $a_1 = A_2+A_4$, $a_2 = A_1+A_3$, $a_3 = A_7$, $a_4 = A_5+A_6$, $a_5 = A_4$, $a_6 = A_3$, $a_7 = A_7$ and $a_8 = A_6$. As its objective function in (74a) is difference of two concave (d.c.) functions $\ln(a_1+a_2x)+\ln(a_3+a_4x)$ and $\ln(a_5+a_6x)+\ln(a_7+a_8x)$, (74) is a simple d.c. optimization problem, which can be solved by d.c. iterations [26]. We now develop a simple path-following procedure of extremely low complexity as a feasible point found in each iteration is available in closed-form. We will use the following inequalities for all $x > 0$ and $\bar{x} > 0$:

$$
-\ln(a+bx) \ge -\ln(a+b\bar{x})+\frac{b\bar{x}}{a+b\bar{x}}-\frac{b}{a+b\bar{x}}x, \qquad (75)
$$

$$
\ln(a+bx) \ge \ln(a+b\bar{x})+\frac{b\bar{x}}{a+b\bar{x}}-\frac{b\bar{x}^2}{a+b\bar{x}}\frac{1}{x}, \qquad (76)
$$

where (75) follows from the concavity of $\ln(1+bx)$ while (76) follows from the convexity of $\ln(1+1/x)$ [27, Theorem 6].

Suppose $x^{(\kappa)}$ is a feasible point found from $(\kappa-1)$th iteration. Then applying (75) and (76) yields

$$
f(x) \ge f^{(\kappa)}(x) := a^{(\kappa)} - b^{(\kappa)}x - \frac{c^{(\kappa)}}{x}
$$

for

$$
\begin{aligned}
0 < a^{(\kappa)} &= f(x^{(\kappa)}) + \sum_{n=1}^{4}\frac{a_{2n}x^{(\kappa)}}{a_{2n-1}+a_{2n}x^{(\kappa)}},\\
0 < b^{(\kappa)} &= \sum_{n=3}^{4}\frac{a_{2n}x^{(\kappa)}}{a_{2n-1}+a_{2n}x^{(\kappa)}}, \qquad (77)\\
0 < c^{(\kappa)} &= \sum_{n=1}^{2}\frac{a_{2n}(x^{(\kappa)})^2}{a_{2n-1}+a_{2n}x^{(\kappa)}}.
\end{aligned}
$$

We solve the following lower bounding maximization problem to generate the next feasible point $x^{(\kappa+1)}$

$$
\max_{x}\ [a^{(\kappa)} - b^{(\kappa)}x - \frac{c^{(\kappa)}}{x}] \quad \text{s. t.} \quad 0 < x \le \frac{1}{K}, \qquad (78)
$$

which admits a closed-form solution

$$
x^{(\kappa+1)} = \begin{cases} \sqrt{c^{(\kappa)}/b^{(\kappa)}} & \text{if} \quad \sqrt{c^{(\kappa)}/b^{(\kappa)}} < 1/K,\\ 1/K & \text{otherwise.}\end{cases} \qquad (79)
$$

Note that $f(x^{(\kappa)}) = f^{(\kappa)}(x^{(\kappa)})$. Since $x^{(\kappa+1)}$ and $x^{(\kappa)}$ are the optimal solution and a feasible point for (78), it is true that

$$
f(x^{(\kappa+1)}) \ge f^{(\kappa)}(x^{(\kappa+1)}) > f^{(\kappa)}(x^{(\kappa)}) = f(x^{(\kappa)}),
$$

i.e. $x^{(\kappa+1)}$ is a better feasible point than $x^{(\kappa)}$ for (74). Then, it can be easily shown that (see e.g. [26] and [27]) the sequence $\{x^{(\kappa)}\}$ of improved feasible points converges to a solution satisfying the Karush-Kuh-Tucker condition for (74). Such iterative procedure is formalized by Algorithm 1.

Algorithm 1 : An algorithm for solving problem (74)

```
1: **Initialization**: Choose a feasible point $x^{(0)}$ for (74). Set $\kappa := 0$.
2: **Repeat**
3: Iterate $x^{(\kappa+1)}$ by (79).
4: Set $\kappa := \kappa + 1$.
5: **Until** convergence of the objective in (74).
```

$$
\begin{aligned}
R_{\mathsf{E}} &= \log_2\left(1+\frac{\gamma_{\mathsf{E}}^2 K^{-1}\tau_{\mathsf{d}} N\alpha_{\mathsf{p}}^2}{\gamma_{\mathsf{E}}K^{-1}(\gamma_{\mathsf{E}}\tau_{\mathsf{d}}\alpha_{\mathsf{n}}^2+1)+(\gamma_{\mathsf{E}}K^{-1}(K-1)+1)(\gamma_{\mathsf{E}}\tau_{\mathsf{d}}[\alpha_{\mathsf{p}}^2+\alpha_{\mathsf{n}}^2]+1)}\right)\\
&= \log_2\left(1+\frac{A_5\alpha_{\mathsf{p}}^2}{A_6\alpha_{\mathsf{p}}^2+A_7}\right),
\end{aligned}
\tag{71}
$$

where $A_5=\gamma_{\mathsf{E}}^2K^{-1}\tau_{\mathsf{d}}N$, $A_6=-\gamma_{\mathsf{E}}^2\tau_{\mathsf{d}}K^{-1}$, and $A_7=(\gamma_{\mathsf{E}}+1)(\gamma_{\mathsf{E}}\tau_{\mathsf{d}}K^{-1}+1)$.

$$
\begin{aligned}
&R_k^{\mathsf{MRT}} =\\
&\log_2\left(1+\frac{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2M\sigma_{\mathsf{u,k}}^2}{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\operatorname{var}(\tilde{a}_{kk}^{\mathsf{MRT}})+\gamma_{\mathsf{d}}(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2(1-\sigma_{\mathsf{u,k}}^2)+\gamma_{\mathsf{d}}(K-1)[\frac{1}{K}-(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2\sigma_{\mathsf{u,k}}^2]+1}\right)\\
&= \log_2\left(1+\frac{A_1(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2(A_2\alpha_{\mathsf{p}}^2+A_3)}{A_4(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2(A_5\alpha_{\mathsf{p}}^2+A_3)+(A_2\alpha_{\mathsf{p}}^2+A_3)(A_6(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2+A_7)}\right),
\end{aligned}
\tag{80}
$$

with $A_1=\gamma_{\mathsf{d}}M\sigma_{\mathsf{u,k}}^2$, $A_2=\gamma_{\mathsf{d}}\tau_{\mathsf{d}}(\operatorname{var}(a_{kk}^{\mathsf{MRT}})-\operatorname{var}(b_{kk}))$, $A_3=\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\operatorname{var}(b_{kk})K^{-1}+1$, $A_4=\gamma_{\mathsf{d}}\operatorname{var}(a_{kk}^{\mathsf{MRT}})$, $A_5=-\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\operatorname{var}(b_{kk})$, $A_6=-\gamma_{\mathsf{d}}K(1-\sigma_{\mathsf{u,k}}^2)+\gamma_{\mathsf{d}}(K-1)$, $A_7=\gamma_{\mathsf{d}}(1-\sigma_{\mathsf{u,k}}^2)+1$.

B. AN in Both Downlink Training Phase and Information Transmission Phase

1) MRT: Plugging (44) into (63), the achievable legitimate rate of the considered system when using AN in both downlink training phase and payload data transmission phase with MRT is re-formulated as in (80) on the next page.

2) ZF: Plugging (46) into (65), the achievable legitimate rate of the considered system when using AN in both downlink training phase and payload data transmission phase with ZF is re-formulated as in (81).

3) Achievable Eavesdropping Rate:

$$
\begin{aligned}
R_{\mathsf{E}} &= \log_2\left(1+\frac{\gamma_{\mathsf{E}}^2\tau_{\mathsf{d}}N\sigma_{\hat{\mathsf{c}}_k}^2(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2}{\gamma_{\mathsf{E}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\sigma_{\hat{\mathsf{c}}_k}^2+\gamma_{\mathsf{E}}(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2+\gamma_{\mathsf{E}}\frac{(K-1)}{K}+1}\right)\\
&= \log_2\left(1+\frac{A_8\alpha_{\mathsf{p}}^2(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2}{A_9\alpha_{\mathsf{p}}^2(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2+A_{10}}\right),
\end{aligned}
\tag{82}
$$

where $A_8=\gamma_{\mathsf{E}}^2\tau_{\mathsf{d}}N$, $A_9=-\gamma_{\mathsf{E}}^2\tau_{\mathsf{d}}$, and $A_{10}=(\gamma_{\mathsf{E}}+1)(\gamma_{\mathsf{E}}\tau_{\mathsf{d}}K^{-1}+1)$.

4) Proposed Optimization Algorithm: The achievable secrecy rate is re-arranged as (83) on the next page.

We seek the values for pilot transmit power ratio in downlink phase, α_{p}, and data transmit power ratio in payload data transmission phase, $\alpha_{\mathsf{p}}^{\mathsf{IT}}$, to maximize the achievable rate of the considered system. The proposed optimization problem is

$$
\max_{\alpha_{\mathsf{p}}^{\mathsf{IT}},\alpha_{\mathsf{p}}} \; R_{\mathsf{s,BP}} \quad \text{s. t.} \quad 0\le(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\le\frac{1}{K}, 0\le\alpha_{\mathsf{p}}^2\le\frac{1}{K}. \tag{84}
$$

Make the variable changes $\alpha_{\mathsf{p}}^2\to x$ and $(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2\to y$ to transform (84) to

$$
\begin{aligned}
\max_{x,y}\; F(x,y) \;:=\; &\ln(\bar{a}_1+a_{11}xy+\bar{a}_2x+b_2y)\\
&+\ln(\bar{a}_3+a_{12}xy)\\
&-\ln(\bar{a}_5+a_{13}xy+\bar{a}_6x+b_6y)\\
&-\ln(\bar{a}_7+a_{14}xy) \qquad \text{(85a)}\\
\text{s. t.} \quad & 0\le x\le\frac{1}{K}, 0\le y\le\frac{1}{K}, \qquad \text{(85b)}
\end{aligned}
$$

where $\bar{a}_1=A_5A_7$, $a_{11}=A_4A_5+A_2A_6+A_1A_2$, $\bar{a}_2=A_2A_7$, $b_2=A_4A_3+A_3A_6+A_1A_3$, $\bar{a}_3=A_{10}$, $a_{12}=A_9$, $\bar{a}_5=A_3A_7$, $a_{13}=A_4A_5+A_2A_6$, $\bar{a}_6=A_2A_7$, $b_6=A_4A_3+A_3A_6$, $\bar{a}_7=A_{10}$, $a_{14}=A_9+A_8$.

Suppose $(x^{(\kappa)},y^{(\kappa)})$ is a feasible point found from the $(\kappa-1)$th iteration. In iterating $x^{(\kappa+1)}$ consider (74) for

$$
\begin{aligned}
&a_1=\bar{a}_1+b_2y^{(\kappa)}, a_2=a_{11}y^{(\kappa)}+\bar{a}_2, a_3=\bar{a}_3, a_4=a_{12}y^{(\kappa)},\\
&a_5=\bar{a}_5+b_6y^{(\kappa)}, a_6=\bar{a}_6+a_{13}y^{(\kappa)}, a_7=\bar{a}_7, a_8=a_{14}y^{(\kappa)}.
\end{aligned}
\tag{86}
$$

Recalling definition (77), $x^{(\kappa+1)}$ defined by (79) is a better feasible point than $x^{(\kappa)}$:

$$
F(x^{(\kappa+1)},y^{(\kappa)})>F(x^{(\kappa)},y^{(\kappa)}). \tag{87}
$$

Similarly, in iterating $y^{(\kappa+1)}$ consider (74) for

$$
\begin{aligned}
&a_1=\bar{a}_1+\bar{a}_2x^{(\kappa+1)}, a_2=a_{11}x^{(\kappa+1)}+b_2, a_3=\bar{a}_3,\\
&a_4=a_{12}x^{(\kappa+1)}, a_5=\bar{a}_5+\bar{a}_6x^{(\kappa+1)}, a_6=b_6+a_{13}x^{(\kappa+1)},\\
&a_7=\bar{a}_7, a_8=a_{14}x^{(\kappa+1)}.
\end{aligned}
\tag{88}
$$

With definition (77), according to (79),

$$
y^{(\kappa+1)}=\begin{cases}\sqrt{c^{(\kappa)}/b^{(\kappa)}} & \text{if} \quad \sqrt{c^{(\kappa)}/b^{(\kappa)}}<1/K,\\ 1/K & \text{otherwise.}\end{cases} \tag{89}
$$

is a better feasible point than $y^{(\kappa)}$:

$$
F(x^{(\kappa+1)},y^{(\kappa+1)})>F(x^{(\kappa+1)},y^{(\kappa)}). \tag{90}
$$

Thus, in Algorithm 2 we propose another path-following computational procedure for solving (85).

$$R_k^{\mathsf{ZF}} = \log_2\left(1 + \frac{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2 \frac{M-K}{\sigma_{\mathsf{u,k}}^2}}{\gamma_{\mathsf{d}}(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2 \operatorname{var}(\tilde{a}_{kk}^{\mathsf{ZF}}) + \gamma_{\mathsf{d}}(\alpha_{\mathsf{n}}^{\mathsf{IT}})^2(1-\sigma_{\mathsf{u,k}}^2) + \gamma_{\mathsf{d}}\frac{(K-1)}{K}(1-\sigma_{\mathsf{u,k}}^2)+1}\right)$$
$$= \log_2\left(1 + \frac{A_1(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2(A_2\alpha_{\mathsf{p}}^2 + A_3)}{A_4(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2(A_5\alpha_{\mathsf{p}}^2+A_3) + (A_2\alpha_{\mathsf{p}}^2+A_3)(A_6(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2+A_7)}\right), \tag{81}$$

where $A_1 = \gamma_{\mathsf{d}}\frac{M-K}{\sigma_{\mathsf{u,k}}^2}$, $A_2 = \gamma_{\mathsf{d}}\tau_{\mathsf{d}}(\operatorname{var}(a_{kk}^{\mathsf{ZF}}) - \operatorname{var}(b_{kk}))$, $A_3 = \gamma_{\mathsf{d}}\tau_{\mathsf{d}}\operatorname{var}(b_{kk})K^{-1}+1$, $A_4 = \gamma_{\mathsf{d}}\operatorname{var}(a_{kk}^{\mathsf{ZF}})$, $A_5 = -\gamma_{\mathsf{d}}\tau_{\mathsf{d}}\operatorname{var}(b_{kk})$, $A_6 = -\gamma_{\mathsf{d}}(1-\sigma_{\mathsf{u,k}}^2)$, $A_7 = \gamma_{\mathsf{d}}(1-\sigma_{\mathsf{u,k}}^2)K^{-1} + \gamma_{\mathsf{d}}(K-1)K^{-1}(1-\sigma_{\mathsf{u,k}}^2)+1$.

$$R_{\mathsf{s,BP}} = [R_k^{\mathsf{A}} - R_{\mathsf{E}}]^+$$
$$= \left[\log_2\left(1 + \frac{A_1(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2(A_2\alpha_{\mathsf{p}}^2 + A_3)}{A_4(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2(A_5\alpha_{\mathsf{p}}^2+A_3) + (A_2\alpha_{\mathsf{p}}^2+A_3)(A_6(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2+A_7)}\right) - \log_2\left(1 + \frac{A_8\alpha_{\mathsf{p}}^2(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2}{A_9\alpha_{\mathsf{p}}^2(\alpha_{\mathsf{p}}^{\mathsf{IT}})^2 + A_{10}}\right)\right]^+. \tag{83}$$

Algorithm 2 : An algorithm for solving problem (85)

```
Initialization: Take a feasible point $x^{(0)}$ and $y^{(0)}$ for (85).
    Set $\kappa := 0$.
Repeat
Iterate $x^{(\kappa+1)}$ according to (79).
Iterate $y^{(\kappa+1)}$ according to (89).
Set $\kappa := \kappa + 1$.
Until convergence of the objectives in (85).
```

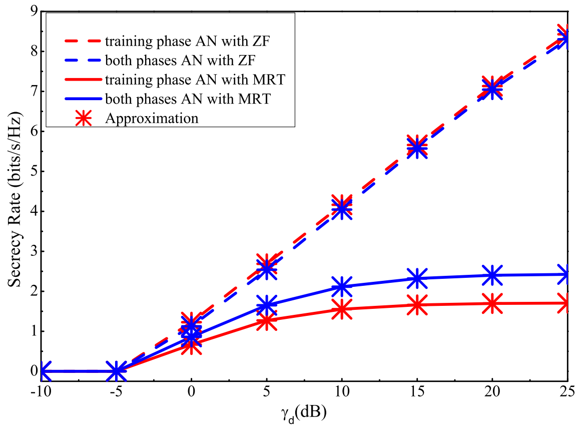


Fig. 2: Approximation tightness.

VI. Numerical Results

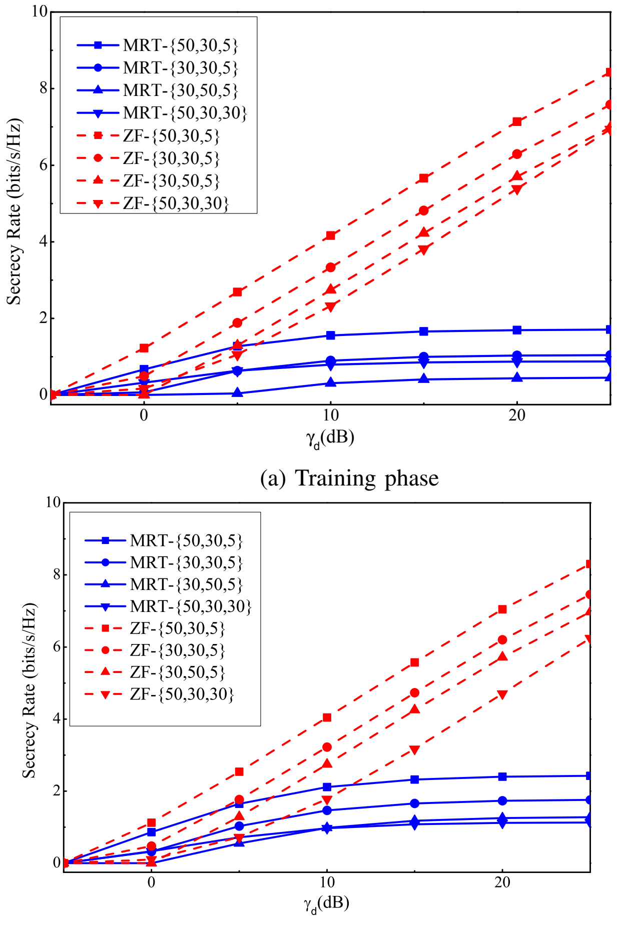
In this section, numerical results based on Monte-Carlo simulation are presented to prove the correctness of our analysis. Without loss of generality, the following parameters are fixed throughout this section: $\gamma_{\mathsf{u}} = 20$ dB, $\gamma_{\mathsf{E}} = 15$ dB, $M = 50$, $N = 30$, $K = 5$, $\beta_{\mathsf{E}} = \beta_k = 1$.

In Fig. 2, the tightness of our approximations is demonstrated. In this setup, an equal resource allocation scheme is applied, i.e., $\alpha_{\mathsf{p}} = \alpha_{\mathsf{n}} = \sqrt{\frac{1}{2K}}$. We can see that the approximations are very tight. Therefore, these approximations will be used in the following numerical work. It is witnessed that ZF scheme shows better performance than MRT scheme. It is because the undesired signals in ZF scheme is much smaller than that in the MRT scheme. In the MRT case, using AN in both phases has better performance than that of using AN in downlink training. The reason is that in the MRT case, the undesired signals from other users is the dominant part of interference. By using AN in payload data transmission phase, a part of transmit power that causes the interference turns into AN for confusing the eavesdropper. In the ZF case, the secrecy performance of using AN in downlink training scheme is slightly better than that of using AN in both phases scheme. It is because the user of using AN in both phases scheme suffers from the leakage of AN. Meanwhile, in using AN in downlink training scheme, the user is free from that interference.

Fig. 3 shows the effect of the number of antennas at the BS, M, the number of antennas at the eavesdropper, N, and the number of users, K, on the secrecy performance of the two proposed AN schemes. In this setup, an equal power allocation is deployed, i.e., $\alpha_{\mathsf{p}} = \alpha_{\mathsf{n}} = \sqrt{\frac{1}{2K}}$. As increasing the number of antennas at the BS and/or reducing the number of users, the secrecy performance in the two proposed AN schemes enhances. It is because each user will receive less undesired signals of the other users and the transmit power of the desired signals is also increased. Besides, increasing the number of antennas at the eavesdropper will enhance the performance of the illegitimate side. As a result, the achievable secrecy rate of the considered system decreases.

Fig. 4 and Fig. 5 demonstrate the convergence of Algorithm 1 and Algorithm 2, respectively. The parameters are set as $x^{(0)} = \frac{1}{10K}$ and $y^{(0)} = \frac{1}{10K}$. After a few iterations the object function of achievable secrecy rate converges to a maximum value.

Fig. 6 provides the performance comparison between the optimal AN transmit power scheme and the equal power allocation scheme. In this setup, $M = 30$ and $N = 50$ By applying the optimal power allocation, the secrecy performance of the two proposed AN schemes is better than that by the equal power allocation. Particularly, the performance of the two proposed AN schemes is similar when the optimal power



(a) Training phase

(b) Both phases

Fig. 3: AN schemes with different value sets of {M,N,K}.

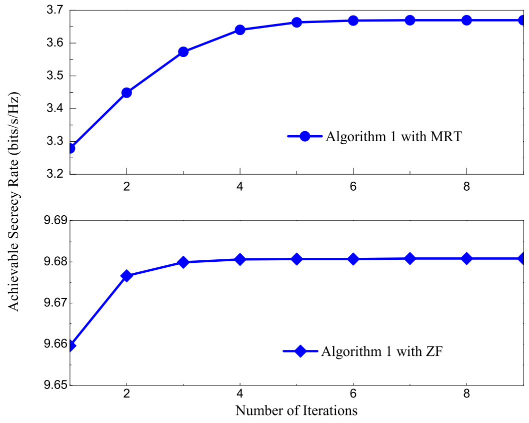


Fig. 4: Convergence of optimization algorithm 1.

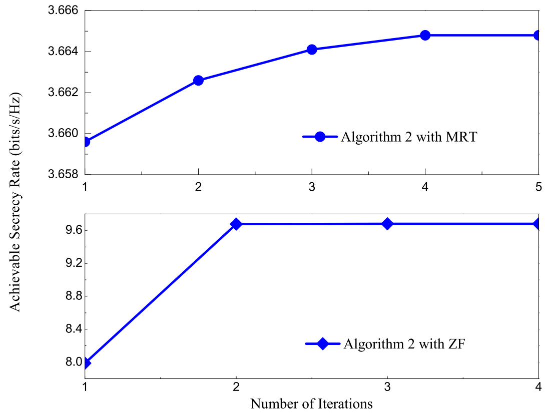


Fig. 5: Convergence of optimization algorithm 2.

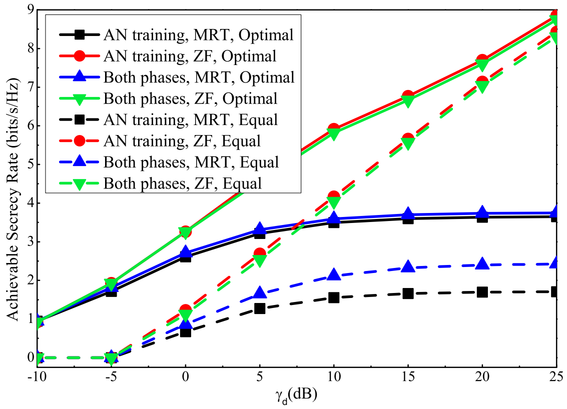


Fig. 6: Comparison between the optimal AN transmit power scheme and the equal power allocation scheme.

allocation is applied. Besides, as the transmit power increases, the improvement in MRT scheme is more significant than that in the ZF scheme. In MRT scheme, the legitimate users suffer from a larger amount of interference than in the ZF scheme. Therefore, the optimization scheme shows a bigger gain.

In Fig. 7, a comparison among AN-aiding schemes is presented. In this simulation, we compare the secrecy perfor-

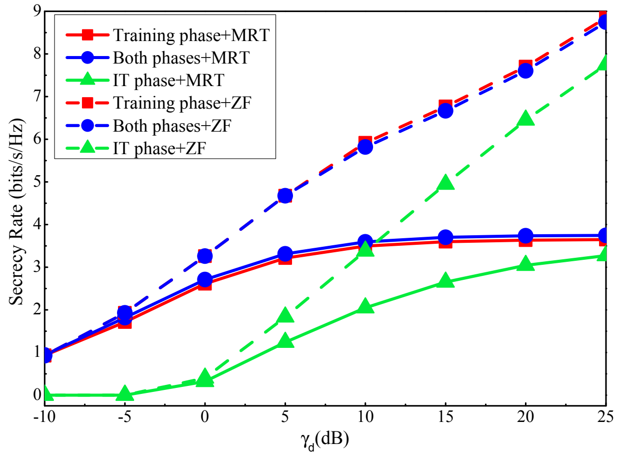


Fig. 7: Comparison among AN schemes.

mance of the two proposed AN schemes with the conventional payload data transmission phase AN-aiding scheme in which AN is placed in the null space of the estimated down-link channel. The results show that the two proposed schemes outperform the conventional AN in payload data transmission phase. The reason is that in the conventional payload data transmission phase AN-aiding scheme, both users and eavesdropper has their best estimated CSI in the training phase. In addition, in payload data transmission phase, a significant part of transmit power of the BS will be used as jamming power to degrade the performance of the eavesdropper. Consequently, the achievable rate at the user is reduced. Meanwhile, in training phase AN-aiding scheme, the outcome of power allocation process is the smaller channel estimation error at the users. As a consequence, using more power to transmit information in the payload data transmission phase leads to a better achievable rate at the users while the achievable rate at the eavesdropper is kept sufficiently low. Using AN in both phases offers the considered system a flexible solution to maximize its secrecy rate at the price of a higher complexity in the power allocation process.

VII. Conclusion

In this paper, two AN-aiding schemes have been proposed to enhance the secrecy performance of a massive MIMO network in the presence of a multiple-antenna eavesdropper. Analytical expressions and tight approximations for the achievable secrecy rate of the considered system have been developed to investigate the performance of the two proposed AN-aiding schemes in the presence of imperfect channel estimation. The results have shown that using AN in the downlink training phase of the massive MIMO networks can effectively suppress the eavesdropping side. Besides, deploying AN in both downlink training phase and payload data transmission phase facilitates the system a flexible solution to enhance the secrecy performance at the price of higher complexity.

Appendix A

The term $\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}\right\}$ can be calculated as

$$\begin{aligned}\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}\right\} &= \mathbb{E}\left\{(\hat{\boldsymbol{h}}_k^T + \tilde{\boldsymbol{h}}_k^T)\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}\right\} \\ &= \mathbb{E}\left\{\left\|\hat{\boldsymbol{h}}_k\right\|\right\} = \frac{\sigma_{\mathsf{u,k}}}{\sqrt{2}}\mathbb{E}\left\{\sqrt{X}\right\} = \sigma_{\mathsf{u,k}}\frac{\Gamma\left(\frac{2M+1}{2}\right)}{\Gamma(M)}, \end{aligned} \tag{A.1}$$

where X is Chi-squared distributed with $2M$ degrees of freedom.

The expression of $\mathbb{E}\left\{|a_{kk}^{\mathsf{MRT}}|^2\right\}$ is

$$\begin{aligned}\mathbb{E}\left\{|a_{kk}^{\mathsf{MRT}}|^2\right\} &= \mathbb{E}\left\{\left\|\hat{\boldsymbol{h}}_k\right\|^2\right\} + \mathbb{E}\left\{\left|\tilde{\boldsymbol{h}}_k^T\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}\right|^2\right\} \\ &= (M-1)\sigma_{\mathsf{u,k}}^2 + 1. \end{aligned} \tag{A.2}$$

$\mathbb{E}\left\{a_{kk}^{\mathsf{ZF}}\right\}$ can be computed as follows:

$$\begin{aligned}\mathbb{E}\left\{a_{kk}^{\mathsf{ZF}}\right\} &= \mathbb{E}\left\{(\hat{\boldsymbol{h}}_k^T + \tilde{\boldsymbol{h}}_k^T)\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\right\} \\ &= \mathbb{E}\left\{\frac{1}{\|\boldsymbol{v}_k\|}\right\} + \mathbb{E}\left\{\frac{\tilde{\boldsymbol{h}}_k^T\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\right\} = \mathbb{E}\left\{\frac{1}{\|\boldsymbol{v}_k\|}\right\} \\ &= \frac{1}{\sqrt{2}\sigma_{\mathsf{u,k}}}\mathbb{E}\left\{\sqrt{x}\right\} \\ &= \int_0^\infty \frac{x^{M-K+\frac{1}{2}}\exp\left(\frac{-x}{2}\right)}{2^{M+\frac{3}{2}-K}\sigma_{\mathsf{u,k}}\Gamma(M+1-K)}dx \\ &= \frac{\Gamma\left(M-K+\frac{3}{2}\right)}{\sigma_{\mathsf{u,k}}\Gamma(M-K+1)}, \end{aligned} \tag{A.3}$$

where x follows Chi-squared distribution with $2(M+1-K)$ degrees of freedom.

Appendix B

The term $\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}b_{kk}^*\right\}$ is formulated as

$$\begin{aligned}\mathbb{E}\left\{a_{kk}^{\mathsf{MRT}}b_{kk}^*\right\} &= \mathbb{E}\left\{(\hat{\boldsymbol{h}}_k^T + \tilde{\boldsymbol{h}}_k^T)\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}\boldsymbol{z}_k^H(\hat{\boldsymbol{h}}_k^* + \tilde{\boldsymbol{h}}_k^*)\right\} \\ &= \mathbb{E}\left\{(\hat{\boldsymbol{h}}_k^T + \tilde{\boldsymbol{h}}_k^T)\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}\underbrace{\boldsymbol{z}_k^H\hat{\boldsymbol{h}}_k^*}_{=0}\right\} + \mathbb{E}\left\{\hat{\boldsymbol{h}}_k^T\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}\boldsymbol{z}_k^H\tilde{\boldsymbol{h}}_k^*\right\} \\ &+ \mathbb{E}\left\{\boldsymbol{z}_k^H\tilde{\boldsymbol{h}}_k^*\tilde{\boldsymbol{h}}_k^T\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}\right\} \\ &= \mathbb{E}\left\{\hat{\boldsymbol{h}}_k^T\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}\boldsymbol{z}_k^H\right\}\underbrace{\mathbb{E}\left\{\tilde{\boldsymbol{h}}_k^*\right\}}_{=0} + \mathbb{E}\left\{\underbrace{\boldsymbol{z}_k^H\frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|}}_{=0}\right\}(1-\sigma_{\mathsf{u,k}}^2) \\ &= 0. \end{aligned} \tag{B.1}$$

The term $\mathbb{E}\left\{a_{kk}^{\mathsf{ZF}}b_{kk}^*\right\}$ is expressed as

$$\begin{aligned}\mathbb{E}\left\{a_{kk}^{\mathsf{ZF}}b_{kk}^*\right\} &= \mathbb{E}\left\{(\hat{\boldsymbol{h}}_k^T + \tilde{\boldsymbol{h}}_k^T)\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\boldsymbol{z}_k^H(\hat{\boldsymbol{h}}_k^* + \tilde{\boldsymbol{h}}_k^*)\right\} \\ &= \mathbb{E}\left\{(\hat{\boldsymbol{h}}_k^T + \tilde{\boldsymbol{h}}_k^T)\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\underbrace{\boldsymbol{z}_k^H\hat{\boldsymbol{h}}_k^*}_{=0}\right\} + \mathbb{E}\left\{\hat{\boldsymbol{h}}_k^T\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\boldsymbol{z}_k^H\tilde{\boldsymbol{h}}_k^*\right\} \\ &+ \mathbb{E}\left\{\boldsymbol{z}_k^H\tilde{\boldsymbol{h}}_k^*\tilde{\boldsymbol{h}}_k^T\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\right\} \\ &\overset{(a)}{=} \mathbb{E}\left\{\hat{\boldsymbol{h}}_k^T\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\boldsymbol{z}_k^H\right\}\underbrace{\mathbb{E}\left\{\tilde{\boldsymbol{h}}_k^*\right\}}_{=0} + \mathbb{E}\left\{\boldsymbol{z}_k^H\frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\right\}(1-\sigma_{\mathsf{u,k}}^2) \\ &= 0, \end{aligned} \tag{B.2}$$

where step (a) is obtained by choosing $\boldsymbol{z}_k$ satisfy $\boldsymbol{z}_k^H\boldsymbol{v}_k = 0$. Similarly, $\mathbb{E}\left\{b_{kk}(a_{kk}^{\mathsf{MRT}})^*\right\} = 0$ and $\mathbb{E}\left\{b_{kk}(a_{kk}^{\mathsf{ZF}})^*\right\} = 0$.

Appendix C
Proof of $\mathbb{E}\left\{\boldsymbol{c}_k \boldsymbol{d}_k^H\right\}$

Consider

$$\mathbb{E}\left\{ \boldsymbol{g}_m^T \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|} \boldsymbol{z}^H \boldsymbol{g}_n^* \right\} = \mathbb{E}\left\{ \boldsymbol{z}^H \boldsymbol{g}_n^* \boldsymbol{g}_m^T \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|} \right\} \tag{C.1}$$

. If $m = n$,

$$\mathbb{E}\left\{ \boldsymbol{z}^H \boldsymbol{g}_m^* \boldsymbol{g}_m^T \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|} \right\} = \mathbb{E}\left\{ \boldsymbol{z}^H \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|} \right\} = 0 \tag{C.2}$$

. If $m \neq n$, $\mathbb{E}\left\{ \boldsymbol{z}^H \boldsymbol{g}_n^* \boldsymbol{g}_m^T \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|} \right\} = 0$ because, $\boldsymbol{g}_n$ and $\boldsymbol{g}_m$ are mutually independent and independent on $\hat{\boldsymbol{h}}_k$ and $\boldsymbol{z}$.

Appendix D
Proof of a_{kk}^{MRT} approximation

The approximation of a_{kk}^{MRT} can be derived as follows:

$$\begin{aligned} a_{kk}^{\mathsf{MRT}} &= \boldsymbol{h}_k^T \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|} = \left\|\hat{\boldsymbol{h}}_k\right\| + \tilde{\boldsymbol{h}}_k^T \frac{\hat{\boldsymbol{h}}_k^*}{\left\|\hat{\boldsymbol{h}}_k^*\right\|} \approx \left\|\hat{\boldsymbol{h}}_k\right\| \\ &= \frac{\sigma_{\mathsf{u,k}}}{\sqrt{2}} X = \frac{\sigma_{\mathsf{u,k}}}{\sqrt{2}} \left(\sigma \underbrace{\frac{X-\mu}{\sigma}}_{Y} + \mu \right) \\ &\overset{M \to \infty}{\sim} \frac{\sigma_{\mathsf{u,k}}}{\sqrt{2}} \left[\sigma \mathcal{N}(0,1) + \mu \right] \\ &\sim \mathcal{N}\left(\sigma_{\mathsf{u,k}} \frac{\Gamma\left(\frac{2M+1}{2}\right)}{\Gamma(M)}, \sigma_{\mathsf{u,k}}^2 M - \sigma_{\mathsf{u,k}}^2 \left[\frac{\Gamma\left(\frac{2M+1}{2}\right)}{\Gamma(M)} \right]^2 \right) \end{aligned} \tag{D.1}$$

where X follows Chi distribution with $2M$ degrees of freedom. Mean value and variance of X are $\mu = \sqrt{2}\frac{\Gamma\left(\frac{2M+1}{2}\right)}{\Gamma(M)}$ and $\sigma^2 = 2M - \mu^2$. Therefore, $\hat{a}_{kk}^{\mathsf{MRT}}$ and $\tilde{a}_{kk}^{\mathsf{MRT}}$ are mutually independent. As a consequence, a_{kl}^{MRT} and $\hat{a}_{kk}^{\mathsf{MRT}}$ are independent. $\hat{a}_{kk}^{\mathsf{MRT}}$ can be approximated as a Gaussian RV, i.e., $\hat{a}_{kk}^{\mathsf{MRT}} \sim \mathcal{N}(\mathbb{E}\left\{\hat{a}_{kk}^{\mathsf{MRT}}\right\}, \mathrm{var}(\hat{a}_{kk}^{\mathsf{MRT}}))$.

Appendix E
Proof of a_{kk}^{ZF} approximation

Formulation of a_{kk}^{ZF} is

$$\begin{aligned} a_{kk}^{\mathsf{ZF}} &= \boldsymbol{h}_k^T \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|} = (\hat{\boldsymbol{h}}_k^T + \tilde{\boldsymbol{h}}_k^T) \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|} = \frac{1}{\|\boldsymbol{v}_k\|} + \tilde{\boldsymbol{h}}_k^T \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|} \\ &= \frac{1}{\sqrt{\left[(\hat{\boldsymbol{H}}^T \hat{\boldsymbol{H}}^*)^{-1}\right]_{k,k}}} + \tilde{\boldsymbol{h}}_k^T \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|} \\ &\approx \frac{\sqrt{M-K}}{\sigma_{\mathsf{u,k}}} + \tilde{\boldsymbol{h}}_k^T \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}. \end{aligned} \tag{E.1}$$

Therefore, $a_{kk}^{\mathsf{ZF}} \sim \mathcal{CN}(\frac{\sqrt{M-K}}{\sigma_{\mathsf{u,k}}}, 1 - \sigma_{\mathsf{u,k}}^2)$. As a consequence, $\hat{a}_{kk}^{\mathsf{ZF}}$ and $\tilde{a}_{kk}^{\mathsf{ZF}}$ are mutually independent.

References

[1] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, Feb. 2014.

[2] H. V. Poor, "Information and inference in the wireless physical layer," *IEEE Commun. Mag.*, pp. 40–47, Feb. 2012.

[3] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Jan. 1975.

[4] M. Pei, J. Wei, K.-K. Wong, and X. Wang, "Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 544–549, Feb. 2012.

[5] A. Mukherjee and A. L. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, Jan. 2011.

[6] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.

[7] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

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

[9] Y.-H. Nam, B. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.

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

[11] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Sep. 2012.

[12] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. Allerton Conf. on Commun., Control, and Comput.*, Monticello, IL, Oct. 2013, pp. 293–298.

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

[14] N.-P. Nguyen, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and D. B. da Costa, "Full-duplex cyber-weapon with massive arrays," *IEEE Trans. Commun.*, vol. 65, no.12, pp. 5544–5558, Aug. 2017.

[15] A. Shafie, Z. Ding, and N. Al-Dhahir, "Hybrid spatio-temporal artificial noise design for secure MIMOME-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 3871–3886, May 2017.

[16] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Analysis and design of secure massive MIMO systems in the presence of hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 2001–2016, Mar. 2017.

[17] J. Zhu, W. Xu, and N. Wang, "Secure massive MIMO systems with limited RF chains," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5455–5460, Jun. 2017.

[18] K. Guo, Y. Guo, and G. Ascheid, "Security-constrained power allocation in MU-massive-MIMO with distributed antennas," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8139–8153, Dec. 2016.

[19] B. Chen, C. Zhu, W. Li, J. Wei, V. C. M. Leung, and L. T. Yang, "Original symbol phase rotated secure transmission against powerful massive MIMO eavesdropper," *IEEE Access*, vol. 4, pp. 3016–3025, Jun. 2016.

[20] J. Zhu, R. Schober, and V. K. Bhargava, "Linear precoding of data and artificial noise in secure massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2245–2261, Mar. 2016.

[21] W. Zhao, S.-H. Lee, and A. Khisti, "Phase-only zero forcing for secure communication with multiple antennas," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 8, pp. 1334–1345, Dec. 2016.

[22] J. Wang, J. Lee, F. Wang, and T. Q. S. Quek, "Jamming-aided secure communication in massive MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6854–6868, Dec. 2015.

[23] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sep. 2014.

[24] S. M. Kay, *Fundamentals of statistical signal processing: Estimation theory.* Englewood Cliffs, NJ: Prentice Hall, 1993.

[25] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge, United Kingdom: Cambridge University Press, 2016.

[26] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local D.C. programming," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.

[27] H. H. M. Tam, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor, "Joint load balancing and interference management for small-cell heterogeneous networks with limited backhaul capacity," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 2, pp. 872–884, Feb. 2017.

Nam-Phong Nguyen (S'16) was born in Hanoi, Vietnam. He received the B.S. degree in electronics and telecommunication engineering and the M.S. degree in electronics engineering from the Hanoi University of Science and Technology, Vietnam, in 2012 and 2014, respectively. He received the Ph.D. degree from Queen's University Belfast in 2018. He is currently a postdoctoral fellow at Memorial University, Canada. His research interests include physical layer security, cognitive relay networks, energy harvesting communications, and massive MIMO.

Hien Quoc Ngo received the B.S. degree in electrical engineering from the Ho Chi Minh City University of Technology, Vietnam, in 2007, the M.S. degree in electronics and radio engineering from Kyung Hee University, South Korea, in 2010, and the Ph.D. degree in communication systems from Linköping University (LiU), Sweden, in 2015. In 2014, he visited the Nokia Bell Labs, Murray Hill, New Jersey, USA. From January 2016 to April 2017, Hien Quoc Ngo was a VR researcher at the Department of Electrical Engineering (ISY), LiU. He was also a Visiting Research Fellow at the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, UK, funded by the Swedish Research Council.

Hien Quoc Ngo is currently a Lecturer at Queen's University Belfast, UK. His main research interests include massive (large-scale) MIMO systems, cell-free massive MIMO, physical layer security, and cooperative communications. He has co-authored many research papers in wireless communications and co-authored the Cambridge University Press textbook *Fundamentals of Massive MIMO* (2016).

Dr. Hien Quoc Ngo received the IEEE ComSoc Stephen O. Rice Prize in Communications Theory in 2015 and the IEEE ComSoc Leonard G. Abraham Prize in 2017. He also received the IEEE Sweden VT-COM-IT Joint Chapter Best Student Journal Paper Award in 2015. He was an *IEEE Communications Letters* exemplary reviewer for 2014, an *IEEE Transactions on Communications* exemplary reviewer for 2015, and an *IEEE Wireless Communications Letters* exemplary reviewer for 2016. He was a Guest Editor of IET Communications, special issue on "Recent Advances on 5G Communications" and a Guest Editor of IEEE Access, special issue on "Modelling, Analysis, and Design of 5G Ultra-Dense Networks", in 2017. He has been a member of Technical Program Committees for several IEEE conferences such as ICC, Globecom, WCNC, VTC, WCSP, ISWCS, ATC, ComManTel.

Trung Q. Duong (S'05, M'12, SM'13) received his Ph.D. degree in Telecommunications Systems from Blekinge Institute of Technology (BTH), Sweden in 2012. Currently, he is with Queen's University Belfast (UK), where he was a Lecturer (Assistant Professor) from 2013 to 2017 and a Reader (Associate Professor) from 2018. His current research interests include Internet of Things (IoT), wireless communications, molecular communications, and signal processing. He is the author or co-author of more than 280 technical papers published in scientific journals (160 articles) and presented at international conferences (125 papers).

Dr. Duong currently serves as an Editor for the IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, IET Communications, and a Lead Senior Editor for IEEE Communications Letters. He was awarded the Best Paper Award at the IEEE Vehicular Technology Conference (VTC-Spring) in 2013, IEEE International Conference on Communications (ICC) 2014, IEEE Global Communications Conference (GLOBECOM) 2016, and IEEE Digital Signal Processing Conference (DSP) 2017. He is the recipient of prestigious Royal Academy of Engineering Research Fellowship (2016-2021) and has won a prestigious Newton Prize 2017.

Hoang Duong Tuan received the Diploma (Hons.) and Ph.D. degrees in applied mathematics from Odessa State University, Ukraine, in 1987 and 1991, respectively. He spent nine academic years in Japan as an Assistant Professor in the Department of Electronic-Mechanical Engineering, Nagoya University, from 1994 to 1999, and then as an Associate Professor in the Department of Electrical and Computer Engineering, Toyota Technological Institute, Nagoya, from 1999 to 2003. He was a Professor with the School of Electrical Engineering and Telecommunications, University of New South Wales, from 2003 to 2011. He is currently a Professor with the Centre for Health Technologies, University of Technology Sydney. He has been involved in research with the areas of optimization, control, signal processing, wireless communication, and biomedical engineering for more than 20 years.

Kamel Tourki (S'05, M'08, SM'13)was born in Antibes, France. He received the engineering degree in telecommunications in 2003 from the National School of Engineers of Tunis (Tunisia). He received his Master and PhD degrees from the University of Nice Sophia-Antipolis (France) in 2004 and 2008, respectively. He has been with Texas A&M University at Qatar (TAMUQ) from August 2008 through June 2014 as Senior Researcher, and with Ooredoo-Qatar University as Consultant. He joined Huawei France Research Center in December 2014 where he is currently Senior Research Engineer.

Dr. Tourki currently serves as Senior Editor of IEEE COMMUNICATIONS LETTERS and Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS, and he has been appointed as an Expert Scientist by the French National Agency of Research (ANR). He received twice the Research Fellow Excellence Award from TAMUQ (Apr. 2011 & Apr. 2014), Best poster award in IEEE DysPan 2012 Conference, the Outstanding Young Researcher Award from IEEE Communications Society for Europe - Middle East - Africa (EMEA) region in June 2013. Dr Tourki is IEEE Senior member.

His current research interests lie in the fields of 5G & Beyond wireless communication systems, green cooperative and cognitive systems, PHY security and energy harvesting.