Concurrent Spoofing-Jamming Attack in Massive MIMO Systems with a Full-Duplex Multi-Antenna Eavesdropper

Mahmoud Alageli, Aissa Ikhlef, *Senior Member, IEEE*, and Jonathon Chambers, *Fellow, IEEE*

***Abstract*—In this paper, we evaluate the ultimate severity of a concurrent spoofing-jamming un-detected attack from a full-duplex (FD) multi-antenna eavesdropper (EV) on a legitimate user equipment (UE) in a massive multiple-input multiple-output (MIMO) system. The FD EV can concurrently exploit its antennas to spoof one UE pilot and to estimate its link to the same UE in the uplink, and then to eavesdrop the UE's data and perform directional jamming in the downlink. From the perspective of the EV, we derive an expression for the ergodic rate difference, which is general for any possible overlap between spoofing and jamming antenna subsets. Residual spoofing and jamming self-interferences and their statistical dependencies are accounted for in the derived expression. The EV optimizes the trade-off between both spoofing-jamming powers, and antenna subsets to minimize the ergodic rate difference. Numerical results show that the EV is capable of destroying the security of the legitimate communication with a small number of antennas and a power budget equal to that of the attacked UE. The severity of the attack depends on the EV's knowledge of the power allocation strategies used at the base station (BS).**

***Index Terms*—Full-duplex, active eavesdropping, physical-layer security, spoofing, jamming, massive MIMO.**

I. Introduction

Recently, massive multiple-input multiple-output (MIMO) systems have proved effectiveness in providing high data rates to a large number of users [1]. However, restricting channel estimation to the uplink phase – due to the large antenna array at the transmitter – makes massive MIMO systems vulnerable to hacking by active spoofing attacks. Active attacks can target both the uplink and downlink phases by spoofing the attacked information user via transmitting a signal identical to its training sequences, and jamming data transmission intended for the attacked user, respectively. A considerable body of research has considered improving the security against active attacks by implementing information-theoretic and physical-layer designs [2]–[4]. The authors in [2] proposed suppressing jamming in the uplink transmission by utilizing spatial correlation of the jammer's channel. In [3], joint power control for artificial noise and the data signal were used to enhance the secrecy rate in massive MIMO networks under active spoofing attack of an energy harvester. Different types of artificial noise precoders were investigated in [4] to enhance the secrecy rate of the attacked user.

The aforementioned works considered the active attack from the service provider's point of view; however, considering the problem from the attacker's point of view is also important to quantify how severe the active malicious attack can be. In this paper, we investigate the impact of a full-duplex (FD) multi-antenna active EV whose antennas have simultaneous transmit and receive (STAR)[1] capability. The FD EV exploits this STAR capability along with digital self-interference cancellation [11]–[13] to attempt two concurrent attacks during each transmission phase against a certain UE in a massive MIMO system that lacks active attack detection capabilities or its detection algorithm fails. The concurrent attacks are: 1) Spoofing the training sequences of the attacked UE, and estimating the channel between the attacked UE and the EV during the channel estimation phase; 2) Eavesdropping the data intended for the attacked UE and jamming signal reception at the attacked UE during downlink transmission.

A. Related works

Having an insight into the capabilities of malicious active intruders is of paramount importance for improving secrecy design in wireless networks. This direction of research has attracted the research community, particularly, for adversary FD active attack (illegitimate attack) [14]–[16] and proactive FD active attack (legitimate attack) [17] . In [14], the FD EV estimates its channel to the legitimate transmitter and receiver during the channel training phase, then, uses this knowledge to simultaneously eavesdrop the transmitter data and jam the receiver using distinct antenna subsets. The authors in [15] assumed a jammer equipped with a massive antenna array. The jammer relies on reciprocity to estimate its channel to a legitimate user, then uses this estimate to direct jamming and destroy legitimate communication. More recently, in [17], the proactive spoofing-jamming attack, which is similar to the adversary attack but on suspicious users, has been considered. Complete self-interference cancellation and perfect channel knowledge are assumed by the cooperative FD spoofing and

Mahmoud Alageli is with the Faculty of Engineering, Garaboulli, Elmergib University, Al Khums, Libya (e-mail: mahmoud.alageli@elmergib.edu.ly).

Aissa Ikhlef is with the Department of Engineering, Durham University, Durham, DH1 3LE, U.K. (e-mail: aissa.ikhlef@durham.ac.uk).

Jonathon Chambers is with the School of Engineering, Newcastle University, Newcastle Upon Tyne, NE1 7RU, UK (e-mail: Jonathon.Chambers@ncl.ac.uk)..

[1]The STAR implementation relies on either a concentric-antenna configuration based on cross polarization antenna isolation [5]–[7] or a single-antenna configuration based on multiple-port circulator for isolating transmit and receive signals [8]–[11].

jamming hardware which are at separate locations. The spoofing nodes receive, manipulate, and forward the signal between two suspicious data users. Destructive relaying and jamming by spoofing and jamming nodes are cooperatively optimized.

B. Contributions

Given the existing works, and to the best of the authors' knowledge, the FD EV performing concurrent pilot spoofing and jamming-channel estimation in a massive MIMO system has not been considered in the literature. Employing this new type of concurrent attack using STAR antennas for both spoofing and jamming increases the attack severity and introduces a challenge in analyzing the correlated spoofing and jamming self-interferences. Next, we present our main contributions in relation to some relevant existing works

- From the system model perspective, all previous works have considered the active attack as being accomplished in orthogonal phases where the analysis deals with statistically independent signals. For example, in [14] the adversary EV takes advantage of legitimate downlink training to estimate the eavesdropping and jamming channels. During the downlink data phase (which is orthogonal to the training phase), the EV performs data eavesdropping and jamming towards the legitimate receiver. Both received data signal and jamming are statistically independent. However, in our work, due to the FD concurrent attacks, the eavesdropped signal, jamming and spoofing self-interferences are statistically dependent. The analyses take into account the non-zero mean and the dependencies between spoofing and jamming self-interference sub-channels. To the best of our knowledge, and at least for the FD active EV, such system model and analyses have not been undertaken in the literature.
- All existing works of active eavesdropping attack adopt simplified assumptions when dealing with self-interference at the full-duplex EV. For example, a perfect self-interfernce cancellation was assumed in [18]. The works in [14], [19] omitted the analysis of self-interference cancellation by assuming a given value of the self-interference impact. Also, the work in [20] has omitted both self-interference channel estimation and self-interference cancellation. The instantaneous impact of self-interference was obtained based on a given realization of self-interference channel. However, in our work, we perform rigorous analyses to calculate the statistics of residual self-interferences (spoofing and jamming self-interference) based on a practical model of the self-interference channel that follows the Rician distribution. The analyses take into account the non-zero mean and the dependencies between spoofing and jamming self-interference sub-channels.
- We propose a max-min algorithm at the base station (BS) to control the power allocated to the UEs. Also, we propose the optimizing of the active attack by the EV under three different cases: In the first case, the EV is aware of the optimized power allocation at the BS. In the second case, the EV is aware of equal power allocation at the BS. In the third case, the EV is unaware of the optimized power allocation at the BS and assumes equal power allocation at the BS.

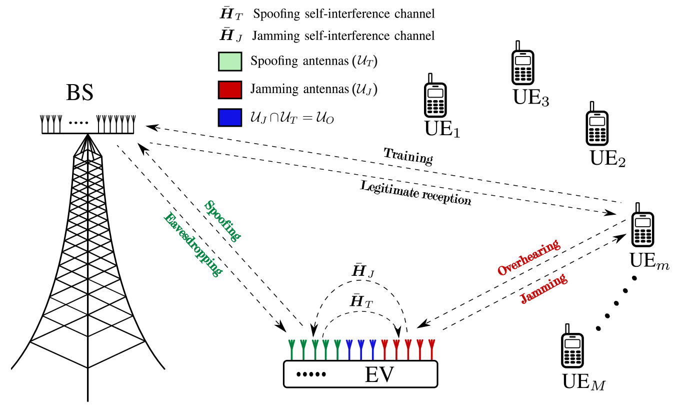

Fig. 1. An illustration of the proposed concurrent spoofing-jamming attack with a full-duplex EV.

Notation: Vectors and matrices are denoted by boldface lower case and upper case letters, respectively. $\boldsymbol{I}_N$ is the $N \times N$ identity matrix, and $\mathbf{1}_M$ is the $M \times 1$ vector with all entries equal to one. $\boldsymbol{I}_N$ and $\mathbf{1}_M$ might be written as $\boldsymbol{I}$ and $\mathbf{1}$, respectively, with dimensionality implied by context. diag$(\boldsymbol{S})$ is a column vector whose entries are the diagonal entries of matrix $\boldsymbol{S}$. Diag$(\boldsymbol{S})$ is a matrix whose entries are zero except the diagonal entries which are the same as the diagonal entries of matrix $\boldsymbol{S}$. The operators $(\cdot)^T$, $(\cdot)^H$, tr$(\cdot)$, $\log_2(\cdot)$, $(\cdot)^*$, $|\cdot|$, and $\|\cdot\|$ denote the transpose, conjugate transpose, trace of a matrix, logarithm to base 2, the complex conjugate, the absolute value, and the Euclidean norm, respectively. $\mathbb{R}^{m \times n}$ denotes the set of all real symmetric $m \times n$ matrices. $\mathbb{C}^{m \times n}$ denotes the set of all complex $m \times n$ matrices. $\boldsymbol{x} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$ denotes a circularly symmetric complex Gaussian random vector $\boldsymbol{x}$ with zero mean and covariance matrix $\boldsymbol{\Sigma}$. $\mathbb{E}[x]$ and var(x) denote the expected value and the variance of x, respectively. $\{a_n\}$ denotes a set of all vectors indexed by n. $|\mathcal{U}_J|$ denotes the cardinality of set $\mathcal{U}_J$. $\mathcal{U}_J \cap \mathcal{U}_T$ and $\mathcal{U}_J \cup \mathcal{U}_T$ denote intersection and union of the sets $\mathcal{U}_J$ and $\mathcal{U}_T$, respectively. $\boldsymbol{S}^{M^-}$ and $\boldsymbol{S}^{M^+}$ denote sub-matrices that span the first M columns and the last M columns of $\boldsymbol{S}$, respectively. $\boldsymbol{S}^{-M}$ and $\boldsymbol{S}^{+M}$ denote sub-matrices that span the first M rows and the last M rows of $\boldsymbol{S}$, respectively. $f(N) \xrightarrow{N \to \infty} a$ is equivalent to $\lim_{N \to \infty} f(N) = a$.

II. System Model

As illustrated in Fig. 1, we consider the downlink of a single-cell massive MIMO system consisting of a BS equipped with a large number of antennas N, M single-antenna UEs, $\{\text{UE}_i\}$, $i = 1, 2, ..., M$, and an FD active EV equipped with K antennas. The EV antennas have STAR capability and are used to illegitimately and actively eavesdrop and decode the information signal intended for a certain (attacked) UE, and to jam the attacked UE at the same time. During the uplink training phase, the FD EV performs two simultaneous tasks: 1) Transmits a copy of the training sequence of the

UE under attack, UE_m, $m \in \{1, 2, \ldots, M\}$, via a subset of its antennas, U_T, such that the BS estimates the uplink composite channel coefficients — which are equivalent to the downlink channel coefficients based on channel reciprocity in the time division duplexing mode — of both the UE_m channel and the channel(s) of the EV U_T antenna(s); 2) Utilizes the received training from the targeted UE to estimate the channel(s) between the EV U_J antenna(s) and UE_m. In the downlink transmission phase, the EV utilizes the estimated channel(s) of the U_J antenna(s) to perform jamming towards UE_m. Therefore, the FD EV attempts two concurrent attacks: active eavesdropping and jamming. We assume that the BS is unaware of the EV's active attack. This assumption can be justified by the lack or the non-accurate spoofing detection capabilities at the BS. Also, by this assumption, we restrict our focus to the attack from the knowledgeable EV's perspective, and make the pessimistic scenario that gives an insight into the ultimate severity of the proposed active attack, which is the main goal of this paper[2]. For simplicity, we assume that the EV attacks one UE noting that the corresponding analysis can be extended to the case of attacking multiple UEs. In contrast to some of the existing works that introduce pilot spoofing and jamming by two separate nodes [21], we assume a single FD EV performing concurrent pilot spoofing and jamming-channel estimation which introduces a new challenge in analyzing correlated spoofing and jamming self-interferences.

The locations of the BS, EV, and UEs are assumed to be within a highly scattered environment in which there is no dominant line-of-sight (LoS) propagation path between the transmitter and receiver pair. The EV and UEs are of slow mobility and the signal pulse duration is long enough to ensure the assumption of a flat slow fading signal whose envelope follows a Rayleigh distribution [22, Chapter 4], [23, Chapter 3]. Let $\mathbf{h}_i \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(\mathbf{0}, \beta_i \mathbf{I}_N) = \sqrt{\beta_i}\bar{\mathbf{h}}_i$ denote the uplink channel vector between UE_i and the BS, where β_i is a real positive scalar that represents the large-scale fading, and $\bar{\mathbf{h}}_i \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ represents the small-scale fading. $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{|U_T|}]$, $\mathbf{g}_k \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(\mathbf{0}, \beta_e \mathbf{I}_N) = \sqrt{\beta_e}\bar{\mathbf{g}}_k$, denotes the channel matrix between the U_T antenna(s) of the EV and the BS, where β_e is a real positive scalar that represents the large-scale fading, and $\bar{\mathbf{g}}_k \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ represents the small-scale fading. $\mathbf{f}_i \in \mathbb{C}^{|U_J|\times 1} \sim \mathcal{CN}(0, \beta_{e_i}\mathbf{I}_{|U_J|}) = \sqrt{\beta_{e_i}}\bar{\mathbf{f}}_i$ denotes the channel between the U_J antenna(s) of the EV and UE_i, where β_{e_i} is a real positive scalar that represents the large-scale fading, and $\bar{\mathbf{f}}_i \in \mathbb{C}^{|U_J|\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{|U_J|})$ represents the small-scale fading. $\mathbf{H} \in \mathbb{C}^{K\times K}$ denotes the symmetrical (uplink is equivalent to downlink) self-interference channel matrix of the EV, i.e., the kth column (which is equivalent to the kth row) of $\mathbf{H}$ that represents the channel response between the kth antenna and all K antennas of the EV. The distribution of $\mathbf{H}$ is discussed in Subsection II-B.

[2] The direction of research that focuses on understanding the abilities of the illegitimate attack adversaries has attracted wide attention from the research community [14]–[17]. The assumption of a knowledgeable EV and performing analyses from the perspective of the illegitimate attack side are crucial in quantifying the ultimate severity and are useful for designing potential countermeasures.

A. Uplink UE Channel Estimation

We assume that the user channels exhibit block fading, i.e., the channels remain constant over a coherence time-block and change independently from one block to another. Over each coherence time-block, the transmission occurs across two time slots: $T_U = \tau T_s$ for uplink training sequence transmission, where T_s is the duration of the transmitted symbol and $T_D = (Q - \tau) T_s$ for downlink data transmission, where Q is the total transmitted symbols per coherence time-block. During the uplink training phase, a pilot training sequence is sent from each UE with an average power P_I. The EV sends a copy of the training sequence of UE_m via a subset of its antennas, U_T, using part of its total energy, E, that results in an average power of $\frac{\phi E}{T_U}$, where $0 < \phi < 1$ and E is the total (per time-block) energy available at the EV. The remaining energy $(1 - \phi)E$ is used for transmitting jamming via another antenna subset, U_J, with an average power $\frac{(1-\phi)E}{T_D}$. Although training and jamming happen at orthogonal time slots, the EV still needs to decide how to share its limited energy between training and jamming. The jamming signal is beamformed towards the attacked UE. Based on the STAR property of the EV antennas, the same antenna(s) might belong to both sets such that $U_J \cap U_T = U_O$, i.e., U_O antennas are used simultaneously for both information decoding and jamming. The energies allocated for both transmissions of the EV constitute its total energy, i.e., $\phi E + (1 - \phi)E = E$. The training sequences of $\{\mathrm{UE}_i\}$, $\{\boldsymbol{\psi}_i \in \mathbb{C}^{\tau\times 1}\}$, are assumed to be orthogonal such that $\boldsymbol{\psi}_i^H \boldsymbol{\psi}_j = \delta_{i,j}\tau$, where $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise. The received training signal at the BS is

$$\mathbf{Y} \in \mathbb{C}^{N\times\tau} = \sqrt{P_I}\sum_i^M \mathbf{h}_i \boldsymbol{\psi}_i^T + \sum_{k\in U_T} \mathbf{g}_k \sqrt{\frac{E}{|U_T|T_U}} \boldsymbol{\psi}_m^T + \mathbf{N}, \tag{1}$$

where $\mathbf{N} \in \mathbb{C}^{N\times\tau}$ is the additive noise matrix with entries following the distribution $\mathcal{CN}(0, \sigma_b^2)$. Since the BS is unaware of the active-spoofing attack; i.e., the BS assumes $\mathbf{Y}$ in (1) without the second (spoofing) term; the minimum mean square error (MMSE) estimate of $\mathbf{h}_i$, $\hat{\mathbf{h}}_i = C_i \mathbf{Y}\boldsymbol{\psi}_i^*$, is given as

$$\hat{\mathbf{h}}_i = C_i\left(\sqrt{P_I}\mathbf{h}_i + \psi_{i,m}\sqrt{\frac{E}{|U_T|T_U}}\sum_{k\in U_T}\mathbf{g}_k + \mathbf{N}\boldsymbol{\psi}_i^*\right), \tag{2}$$

$$C_i = \frac{\psi_i\sqrt{P_I}}{\psi_i P_I + \sigma_b^2}, \tag{3}$$

where C_i is the MMSE estimation coefficient. UE_m is the attacked UE. The results in (2)-(3) follow from standard channel estimation theory [24], [25]. The estimations of the self-interference channel, $\mathbf{H}$, and the jamming channel $\mathbf{f}_m$, are left to be discussed in Subsections II-B and II-C, respectively.

B. Self-Interference Channel Estimation

For the STAR antennas of the EV, we assume the single-antenna configuration with 3-port circulator and a cascaded analog-digital interference-cancellation algorithm as in [11]–[13].

The analog cancellation is optimized for every single-antenna transceiver. Briefly, the analog cancellation is based on the subtraction of a linearly pre-processed version of the transmitted signal (measured at the transmit port of the circulator) from the input signal at the receiving port. The linear pre-processing of the subtractive signal is optimized to cancel the effect of transmit signal leakage from the circulator's transmit port to the receive port. For more details, please refer to the references [11]–[13]. The analog cancellation preserves the linearity of self-interference. After analog cancellation, the self-interference channel, $\boldsymbol{H}$, follows the widely adopted model of Rician distribution with large K-factor due to small LoS antenna separation [13], [26]. Therefore, $[\boldsymbol{H}]_{i,l}$ is modelled as

$$[\boldsymbol{H}]_{i,l} = \sqrt{\frac{\beta_{i,l}}{K_{i,l}+1}} \left(\sqrt{K_{i,l}}\, e^{-j\frac{2\pi d_{i,l}}{\lambda}} + h_{i,l} \right), \tag{4}$$

where $j = \sqrt{-1}$, $d_{i,l}$ is the separation between the ith and the lth antennas, λ is the carrier wavelength, $K_{i,l}$ is the K-factor, $\beta_{i,l}$ is the large-scale fading, and $h_{i,l} \sim \mathcal{CN}(0,1)$ is the non-line-of-sight (NLoS) small scale fading. For $l = i$, $K_{i,i} = \infty$, $d_{i,i} = 0$ and then $[\boldsymbol{H}]_{i,i} = \sqrt{\beta_{i,i}}$ represents the circulator isolation gain (after analog cancellation) of the ith antenna.

It has been shown that the K-factor is very large for antenna separation of less than 1 m [13], [26], [27]. For example, the experimental results in [13] show that the value of the K-factor for antenna separation between 10 cm and 40 cm is between 25 dB and 40 dB, i.e., the channel is largely dependent on the LoS component. This encourages the assumption that the coherence time of $\boldsymbol{H}$ is much larger than the coherence time of $\{\boldsymbol{g}_k\}$ and $\{\boldsymbol{f}_i\}$. This large coherence time allows the EV to estimate its self-interference channel in a stage prior to the uplink channel estimation and to exploit the calculated estimate in later digital self-interference cancellation.

The second stage of digital self-interference cancellation requires an estimate of the self-interference channel, $\boldsymbol{H}$. First, we introduce two different types of self-interference (and self-interference sub-channels) at the EV resulting from two transmissions:

1) The EV spoofing transmission (during uplink) that causes *spoofing self-interference* on the received training signal from UE$_m$. In this type, $\mathcal{U}_T$ antennas of the EV transmit the training (spoofing) signal that interferes with the UE$_m$'s training signal received at the antenna subset $\mathcal{U}_J$ of the EV. The spoofing self-interference channel $\boldsymbol{H}_T \in \mathbb{C}^{|\mathcal{U}_J| \times |\mathcal{U}_T|}$ (that could be a matrix, a vector, or a scalar) is optimally chosen, as will be justified later, to be the bottom-left $|\mathcal{U}_J| \times |\mathcal{U}_T|$ sub-matrix of $\boldsymbol{H}$.
2) The EV jamming transmission (during downlink) that causes *jamming self-interference* on the received information signal from the BS. In this type, $\mathcal{U}_J$ antennas of the EV transmit jamming that interferes with the received information signal at the antenna subset $\mathcal{U}_T$. The jamming self-interference channel $\boldsymbol{H}_J \in \mathbb{C}^{|\mathcal{U}_T| \times |\mathcal{U}_J|}$ (that could be a matrix, a vector, or a scalar) is optimally chosen, as will be justified later, to be the top-right $|\mathcal{U}_T| \times |\mathcal{U}_J|$ sub-matrix of $\boldsymbol{H}$. $\boldsymbol{H}$ can be written as

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{H}_J \\ \boldsymbol{A}_2 & \boldsymbol{A}_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1^T & \boldsymbol{A}_2^T \\ \boldsymbol{H}_T & \boldsymbol{A}_3^T \end{bmatrix}, \tag{5}$$

$$\boldsymbol{H}_T = \boldsymbol{H}_J^T, \tag{6}$$

where $\{\boldsymbol{A}_j\}$ are matrices ($\boldsymbol{A}_j$ might be a matrix, vector or scalar) that complete the entries of $\boldsymbol{H}$. Since the self-interference within each subset does not affect the performance of the EV, then the EV only opts to reduce the self-interference caused by one subset of antennas on the other subset of antennas by reducing the average LoS distances from the antennas in one subset to the antennas in the other subset. Intuitively, this is optimally achieved when the antennas within each subset are consecutive and the subsets are at the sides of the linear antenna array [3]. The EV performs one training pass across all antennas by using the training signal matrix $\boldsymbol{\Gamma} \in \mathbb{C}^{K \times \eta}$, $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^H = \eta \boldsymbol{I}_K$, and η is the length of training sequence transmitted by each antenna. The kth row in $\boldsymbol{\Gamma}$ represents the training sequence transmitted by the kth antenna. The received training signal at the EV, $\boldsymbol{Y}_{si} \in \mathbb{C}^{K \times \eta}$, and its projection on $\boldsymbol{\Gamma}$, $\boldsymbol{Y}_{si}\boldsymbol{\Gamma}^H$, are given as

$$\boldsymbol{Y}_{si} = \sqrt{P_{si}}\, \boldsymbol{H}\boldsymbol{\Gamma} + \boldsymbol{N}_e, \tag{7a}$$

$$\boldsymbol{Y}_{si}\boldsymbol{\Gamma}^H = \eta\sqrt{P_{si}}\, \boldsymbol{H} + \boldsymbol{N}_e\boldsymbol{\Gamma}^H, \tag{7b}$$

where P_{si} is the average training power. $\boldsymbol{N}_e \in \mathbb{C}^{K \times \eta}$ is the additive noise matrix with entries following the distribution $\mathcal{CN}(0, \sigma_e^2)$. Now, the transmitted spoofing signal by $\mathcal{U}_J$ antennas is received and projected at $\mathcal{U}_T$ antennas as $\boldsymbol{Y}_J \in \mathbb{C}^{|\mathcal{U}_T| \times |\mathcal{U}_J|}$; and the transmitted spoofing signal by $\mathcal{U}_T$ antennas is received and projected at $\mathcal{U}_J$ antennas as $\boldsymbol{Y}_T \in \mathbb{C}^{|\mathcal{U}_J| \times |\mathcal{U}_T|}$. $\boldsymbol{Y}_J$ and $\boldsymbol{Y}_T$ are obtained by matrix cropping as[4]

$$\boldsymbol{Y}_J = \boldsymbol{Y}_{si}^{|\mathcal{U}_T|\uparrow}\, {\boldsymbol{\Gamma}^{+|\mathcal{U}_J|\downarrow}}^H = \eta\sqrt{P_{si}}\, \boldsymbol{H}_J + \boldsymbol{N}_J, \tag{8a}$$

$$\boldsymbol{Y}_T = \boldsymbol{Y}_{si}^{+|\mathcal{U}_J|\downarrow}\, {\boldsymbol{\Gamma}^{|\mathcal{U}_T|\uparrow}}^H = \eta\sqrt{P_{si}}\, \boldsymbol{H}_T + \boldsymbol{N}_T, \tag{8b}$$

where $\boldsymbol{N}_J \in \mathbb{C}^{|\mathcal{U}_T| \times |\mathcal{U}_J|} = \boldsymbol{N}_e^{|\mathcal{U}_T|\uparrow}{\boldsymbol{\Gamma}^{+|\mathcal{U}_J|\downarrow}}^H$ and $\boldsymbol{N}_T \in \mathbb{C}^{|\mathcal{U}_J| \times |\mathcal{U}_T|} = \boldsymbol{N}_e^{+|\mathcal{U}_J|\downarrow}{\boldsymbol{\Gamma}^{|\mathcal{U}_T|\uparrow}}^H$ are the projected noise matrix/vector at $\mathcal{U}_T$ and $\mathcal{U}_J$ with zero mean and $\eta\sigma_e^2$ variance entries, respectively. Using (8a) and (8b), the MMSE estimate of $\boldsymbol{H}_J$, $\hat{\boldsymbol{H}}_J$; and the MMSE estimates of $\boldsymbol{H}_T$, $\hat{\boldsymbol{H}}_T$, are

$$\hat{\boldsymbol{H}}_J = \boldsymbol{C}_J\, \boldsymbol{Y}_J = \boldsymbol{C}_J \left(\eta\sqrt{P_{si}}\, \boldsymbol{H}_J + \boldsymbol{N}_J \right), \tag{9a}$$

$$\hat{\boldsymbol{H}}_T = \boldsymbol{C}_T\, \boldsymbol{Y}_T = \boldsymbol{C}_T \left(\eta\sqrt{P_{si}}\, \boldsymbol{H}_T + \boldsymbol{N}_T \right), \tag{9b}$$

where

$$\boldsymbol{C}_J \in \mathbb{C}^{|\mathcal{U}_T| \times |\mathcal{U}_T|} = \sqrt{P_{si}}\, \boldsymbol{R}_J \left(\eta P_{si} \boldsymbol{R}_J + |\mathcal{U}_J|\sigma_e^2 \boldsymbol{I} \right)^{-1}, \tag{10a}$$

[3]Please note that considering channel realization for selecting $\mathcal{U}_J$ and $\mathcal{U}_T$ antennas is not relevant since the EV is optimizing the ergodic eavesdropping performance, and also, only the statistical channel values are available at the uplink (attacking) phase.

[4]For self-interference cancellation, the EV only needs to estimate the whole self-interference channel, $\hat{\boldsymbol{H}}$, and use it to cancel the self-interference at all antennas. The details that explain the sub-channel estimates are provided to facilitate deriving the residual self-interference statistical values in (15) and (29) which are required by the EV to optimize its active attack.

$$\boldsymbol{C}_T \; 2\; \mathrm{C}^{jU_J j \quad jU_J j} = \sqrt{P_{si}}\boldsymbol{R}_T\left(\; P_{si}\boldsymbol{R}_T + jU_T j \; {}_e^2 \boldsymbol{I}\right)^{\;\;1}; \tag{10b}$$

$$\left[\boldsymbol{R}_T \; 2\; \mathrm{C}^{jU_J j \quad jU_J j} = \mathrm{E}\left[\boldsymbol{H}_T\boldsymbol{H}_T^H\right]\right]_{i,l} = \;\; {}_{i,l}\sum_{k=1}^{jU_T j} \;\; {}_{n(i),k} +$$

$$(1 \quad {}_{i,l})\sum_{k=1}^{jU_T j}\sqrt{\frac{{}_{n(i),k} \;\; {}_{n(l),k} K_{n(i),k}K_{n(l),k}}{\left(K_{n(i),k}+1\right)\left(K_{n(l),k}+1\right)}}e^{\;\; j\frac{2\pi\left(d_{n(i),k} \quad d_{n(l),k}\right)}{\lambda}}; \tag{10c}$$

$$\left[\boldsymbol{R}_J \; 2\; \mathrm{C}^{jU_T j \quad jU_T j} = \mathrm{E}\left[\boldsymbol{H}_J\boldsymbol{H}_J^H\right]\right]_{i,l} = \;\; {}_{i,l}\sum_{k=1}^{jU_J j} \;\; {}_{i,n(k)} +$$

$$(1 \quad {}_{i,l})\sum_{k=1}^{jU_J j}\sqrt{\frac{{}_{i,n(k)} \;\; {}_{l,n(k)} K_{i,n(k)}K_{l,n(k)}}{\left(K_{i,n(k)}+1\right)\left(K_{l,n(k)}+1\right)}}e^{\;\; j\frac{2\pi\left(d_{i,n(k)} \quad d_{l,n(k)}\right)}{\lambda}}; \tag{10d}$$

$$n(i) = K \quad jU_J j + i; \quad i = 1; \; 2; \; :::; \; jU_J j: \tag{10e}$$

To facilitate our later analyses of residual self-interference (in Subsections II-C and II-D), the channel estimation errors, $\hat{H}_J$ and $\hat{H}_T$, are expressed as

$$\hat{\boldsymbol{H}}_J = \hat{\boldsymbol{H}}_J \quad \boldsymbol{H}_J = \boldsymbol{C}_J\boldsymbol{H}_J + \boldsymbol{C}_J\boldsymbol{N}_J; \tag{11}$$

$$\hat{\boldsymbol{H}}_T = \hat{\boldsymbol{H}}_T \quad \boldsymbol{H}_T = \boldsymbol{C}_T\boldsymbol{H}_T + \boldsymbol{C}_T\boldsymbol{N}_T; \tag{12}$$

where $\boldsymbol{C}_J = \sqrt{P_{si}}\boldsymbol{C}_J \quad \boldsymbol{I}$ and $\boldsymbol{C}_T = \sqrt{P_{si}}\boldsymbol{C}_T \quad \boldsymbol{I}$.

Remark 1: Please note that there are no limitations on the size of U_T and U_J antennas. Therefore, the design might involve an overlap between U_T and U_J antennas, i.e., $0 \quad jU_O j \quad K$. This potential overlap introduces a challenge in the statistical analysis of the residual spoofing and jamming self-interferences as will be seen in Subsection III-B.

C. Jamming Channel Estimation

The EV can rely on its capability of cancelling the spoofing self-interference to facilitate estimating its channel towards the attacked UE, $\boldsymbol{f}_m$ (jamming channel). The received signal at U_J antennas of the EV during uplink training and after spoofing self-interference cancellation is

$$\boldsymbol{Y}_e \; 2\; \mathrm{C}^{jU_J j \quad \tau} = \sqrt{P_I}\sum_{i=1}^{M}\boldsymbol{f}_i \;\; {}_i^T + \sqrt{\frac{E}{jU_T jT_U}} \quad \hat{\boldsymbol{H}}_T\mathbf{1} \;\; {}_m^T + \boldsymbol{N}_{e_J}; \tag{13}$$

where $\boldsymbol{N}_{e_J} \;\; 2 \;\; \mathrm{C}^{jU_J j \quad \tau}$ is the additive noise matrix with entries following the distribution $\mathcal{CN}(0, \sigma_e^2)$. The term $\sqrt{\frac{\phi E}{jU_T j \; T_U}} \quad \hat{H}_T\mathbf{1} \;\; {}_m^T$ represents the residual spoofing self-interference at U_J antennas [28], [29]. The MMSE estimate of $\boldsymbol{f}_m$, $\hat{\boldsymbol{f}}_m$, is given as

$$\begin{aligned}\hat{\boldsymbol{f}}_m &= \boldsymbol{C}_{e_m}\boldsymbol{Y}_e \quad {}_m \\ &= \boldsymbol{C}_{e_m}\left(\; \sqrt{P_I} \; \boldsymbol{f}_m + \sqrt{\frac{E}{jU_T j \; T_U}} \quad \hat{\boldsymbol{H}}_T\mathbf{1}_{jU_T j} + \boldsymbol{N}_{e_J} \quad {}_m\right);\end{aligned} \tag{14a}$$

$$\boldsymbol{C}_{e_m} \; 2\; \mathrm{C}^{jU_J j \quad jU_J j} = \;\; {}_{e_m}\sqrt{P_I}\left(\;\; {}_{e_m}P_I\boldsymbol{I} + \quad \boldsymbol{R}_{T_{si}} + \; {}_e^2\boldsymbol{I}\right)^{\;\;1}; \tag{14b}$$

where $\boldsymbol{R}_{T_{si}}$ is the second-order non-central moment matrix of spoofing residual self-interference at U_J antennas, and it is given as

$$\boldsymbol{R}_{T_{si}} \; 2\; \mathrm{C}^{jU_J j \quad jU_J j} = \frac{E}{jU_T jT_U}\mathrm{E}[\quad \hat{\boldsymbol{H}}_T \; \mathbf{1}\mathbf{1}^T \quad \hat{\boldsymbol{H}}_T^H] = \frac{E}{jU_T jT_U}\Bigg(\boldsymbol{C}_T$$

$$\left(Q_{TL}\left(\mathbf{1}\mathbf{1}^T\right) + Q_{TN}\left(\mathbf{1}\mathbf{1}^T\right)\right)\boldsymbol{C}_T^H + jU_T j \quad {}_e^2\boldsymbol{C}_T\boldsymbol{C}_T^H\Bigg): \tag{15}$$

For presentation convenience, the expressions of Q_{TL} and Q_{TN} functions, and the details of obtaining the result in (15) are omitted here, and provided in the Appendix. The second-order non-central moment matrix of $\hat{\boldsymbol{f}}_m$ is

$$\boldsymbol{R}_{f_m} = \mathrm{E}[\hat{\boldsymbol{f}}_m\hat{\boldsymbol{f}}_m^H] = \tau\beta_{e_m}^2 P_I\left(\tau\beta_{e_m}P_I\boldsymbol{I} + \tau\boldsymbol{R}_{T_{si}} + \sigma_e^2\boldsymbol{I}\right)^{\;\;1}. \tag{16}$$

D. Downlink Transmission and Jamming

The BS employs transmit MF beamforming to direct the information signal vector $\sum_{i=1}^{M}\sqrt{p_i}\;\boldsymbol{w}_i\;x_i$, towards the UEs, where p_i is the power allocated to UE$_i$, x_i is the information symbol intended for UE$_i$, with $x_i \quad \mathcal{CN}(0,1)$, and $\boldsymbol{w}_i \; 2 \; \mathrm{C}^{N \quad 1} = \hat{\boldsymbol{h}}_i / k\hat{\boldsymbol{h}}_i k$. The EV uses what remains from its power and its knowledge of $\hat{\boldsymbol{f}}_m$ to jam UE$_m$ by the beamformed signal

$$\sqrt{\frac{(1 \quad)E}{T_D}}\;\boldsymbol{w}_m\;z; \quad \boldsymbol{w}_m \; 2\; \mathrm{C}^{jU_J j \quad 1} = \frac{\hat{\boldsymbol{f}}_m}{\mathrm{tr}^{\frac{1}{2}}(\boldsymbol{R}_{f_m})}; \tag{17}$$

where $z \qquad \mathcal{CN}(0,1)$ is a random jamming symbol. $\mathrm{E}[k\boldsymbol{w}_m z k^2] = 1$. By using the design parameters ϕ, $jU_J j$, $jU_T j$ and $jU_O j$, the EV can optimize its dual-active approach of uplink spoofing and downlink beamformed jamming to achieve a complex goal of increasing its eavesdropping information rate while decreasing the information rate of UE$_m$.

Given that UE$_m$ is the UE being attacked, then the received signals at UE$_i$, y_i; and at U_T antennas of the EV, $\boldsymbol{y}_{e_m}$, are

$$y_i = \sum_{j=1}^{M}\sqrt{p_j}\;\boldsymbol{h}_i^T\;\boldsymbol{w}_j\;x_j \;+\; \sqrt{\frac{(1 \quad \phi)E}{T_D}}\;\boldsymbol{f}_i^T\;\boldsymbol{w}_m\;z + n_i, \tag{18}$$

$$\begin{aligned}&\boldsymbol{y}_{e_m} \; 2\; \mathrm{C}^{jU_T j \quad 1} = \\ &\sum_{j=1}^{M}\sqrt{p_j}\;\boldsymbol{G}^T\;\boldsymbol{w}_j\;x_j \;+\; \sqrt{\frac{(1 \quad \phi)E}{T_D}} \quad \hat{H}_J\;\boldsymbol{w}_m\;z + \boldsymbol{n},\end{aligned} \tag{19}$$

where $n_i \quad \mathcal{CN}(0,\sigma_u^2)$ and $\boldsymbol{n} \quad \mathcal{CN}(\mathbf{0},\sigma_e^2\boldsymbol{I})$ are the noise at UE$_i$ and U_T antennas of the EV, respectively. The term $\sqrt{\frac{(1 \quad \phi)E}{T_D}} \quad \hat{H}_J\;\boldsymbol{w}_m\;z$ represents the residual jamming self-interference at U_T antennas of the EV [28], [29].

Remark 2: The assumption that the spoofing and jamming antennas might be mutually overlapping is made in order to consider the most general case for the proposed FD EV, and therefore, to ensure quantifying the ultimate severity of the active concurrent attack. This assumption implies the existence of a common part of entries between the spoofing self-interference channel, H_T, and jamming self-interference channel, H_J. Consequently, the analysis needs to take into account the statistical dependencies involved in $\boldsymbol{R}_{T_{si}}$; the variance of the sum of residual jamming self-interference at

U_T antennas, c_m; and the variance of jamming at the attacked UE $\mathrm{var}(\boldsymbol{f}_m^T \boldsymbol{w}_m z)$.

III. Achievable Rates

In this section, we derive the achievable rates of the attacked UE and the EV.

A. UE Achievable Rate

In massive MIMO systems, downlink channel estimation is practically challenging. Users rely on the statistical properties of their precoded channels to decode information. The lower bound on the ergodic achievable rate of UE$_i$, $\underline{R}_i$, is obtained by using Theorem 1 in [30] (given in (22) and (23)). The received signal y_i in (18) can be recast as

$$y_i = \sqrt{p_i}\, \mathsf{E}[a_{i,i}]\, x_i + \sqrt{p_i}\, (a_{i,i} - \mathsf{E}[a_{i,i}])\, x_i + Z_i, \tag{20}$$

where

$$Z_i = \sum_{j \neq i}^{M} \sqrt{p_j} a_{i,j} x_j + \sqrt{\frac{(1-\phi)E}{T_D}}\, \boldsymbol{f}_i^T \boldsymbol{w}_m z + n_i, \qquad a_{i,j} = \boldsymbol{h}_i^T \boldsymbol{w}_j. \tag{21}$$

$\sqrt{p_i}\, \mathsf{E}[a_{i,i}]$ is a positive real scalar which can be perfectly learned at UE$_i$. The additive terms in $\sqrt{p_i}\,(a_{i,i} - \mathsf{E}[a_{i,i}])\, x_i + Z_i$ are zero mean, mutually independent, statistically independent of x_i, and unknown at UE$_i$. For $N \to \infty$, according to the central limit theorem, the distribution of $\sqrt{p_i}\,(a_{i,i} - \mathsf{E}[a_{i,i}])\, x_i + Z_i$ converges to a zero mean complex Gaussian distribution. Since the BS is unaware of the active attack of the EV, consequently, it is also unaware of the jamming upon the attacked UE, i.e., the term $\sqrt{\frac{(1-\phi)E}{T_D}}\, \boldsymbol{f}_i^T \boldsymbol{w}_m$ in (21). Therefore, from the BS point of view, a lower bound on the ergodic rate of UE$_i$ is [1], [30]

$$\underline{R}_i = \log_2 (1 + \underline{\mathrm{SINR}}_i), \tag{22}$$

where

$$\underline{\mathrm{SINR}}_i = \frac{p_i\, |\mathsf{E}[a_{i,i}]|^2}{\sum_{j=1}^{M} p_j\, \mathrm{var}(a_{i,j}) + \sigma_u^2}. \tag{23}$$

By following a similar analysis to that in our previous work [3], the terms in (23) converge to

$$\mathsf{E}[a_{i,i}] \xrightarrow{N \to \infty} {}_i \sqrt{N P_I} \left(P_I \, {}_i + {}_{i,m} \frac{E}{T_U} {}_e + {}^2_u \right)^{-\frac{1}{2}}; \tag{24a}$$

$$\mathrm{var}(a_{i,i}) \xrightarrow{N \to \infty} \frac{{}_i \left({}_{i,m} \frac{\phi E}{T_U} {}_e + {}^2_u \right)}{P_I \, {}_i + {}_{i,m} \frac{\phi E}{T_U} {}_e + {}^2_u}; \tag{24b}$$

$$\mathrm{var}(a_{i,j \neq i}) \xrightarrow{N \to \infty} {}_i. \tag{24c}$$

B. EV Achievable Rate

We make the pessimistic assumption of a knowledgeable EV which perfectly knows its own channel, $\boldsymbol{G}$; and both the beamforming vector and the channel of the attacked UE, $\boldsymbol{h}_m$ and $\boldsymbol{w}_m$[5]. In the following, given that the EV is attacking UE$_m$, and from the EV point of view, we provide two different bounds: 1) A lower bound on the ergodic rate of the EV, $\underline{R}_{e_m}$; 2) An upper bound on the ergodic rate of UE$_m$, $\overline{R}_m$.

Using maximal ratio combining (MRC), the ergodic information rate of the EV is

$$R_{e_m} = \mathsf{E}[\log_2 (1 + \mathrm{SINR}_{e_m})], \tag{25}$$

where

$$\mathrm{SINR}_{e_m} = \frac{p_m\, |b_m|^2}{\sum_{j \neq m}^{M} p_j\, \mathrm{var}(b_j) + c_m + {}^2_e}, \tag{26}$$

$$b_j = \boldsymbol{1}^T \boldsymbol{G}^T \boldsymbol{w}_j, \tag{27}$$

and c_m represents the variance of the sum of residual jamming (intended for UE$_m$) self-interference at U_T antennas of the EV, and it is given as

$$\begin{aligned} c_m &= \frac{(1-\phi)E}{T_D}\, \mathrm{var}(\boldsymbol{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m z) \\ &= \frac{(1-\phi)E}{T_D \mathrm{tr}(\boldsymbol{R}_{f_m})} \left(|(c_{m_{11}} + c_{m_{12}})|^2 + \sum_{i=1}^{4} c_{m_{2i}} \right), \end{aligned} \tag{28}$$

where

$$c_{m_{11}} = \sqrt{\frac{E}{|U_T| T_U}}\, \boldsymbol{1}^T \boldsymbol{C}_J \left(Q_{JL} \left(\boldsymbol{C}_{e_m} \boldsymbol{C}_T \right) + Q_{JN} \left(\boldsymbol{C}_{e_m} \boldsymbol{C}_T \right) \right) \boldsymbol{1}; \tag{29a}$$

$$c_{m_{12}} = 1(U_O) \sqrt{\frac{E\, {}^4_e}{|U_T| T_U}}\, \boldsymbol{1}^T_{|U_T|} \boldsymbol{C}_J^{|U_O|^+} \boldsymbol{C}_T^{|U_O|\;H} \boldsymbol{C}_{em}^{|U_O|^H} \boldsymbol{1}_{|U_O|}; \tag{29b}$$

$$\begin{aligned} c_{m_{21}} = \frac{{}^2 E}{|U_T| T_U} \boldsymbol{1}^T \boldsymbol{C}_J \Big(& Q_{JL} \left(\boldsymbol{C}_{e_m} \boldsymbol{C}_T \right) \boldsymbol{1}\boldsymbol{1}^T Q_{JL} \left(\boldsymbol{C}_{e_m}^T \boldsymbol{C}_T^T \right) \\ & + Q_{JN} \left(\boldsymbol{C}_{e_m} \boldsymbol{C}_T \right) Q_{JN} \left(\boldsymbol{C}_{e_m}^T \boldsymbol{C}_T^T \right) \Big) \boldsymbol{C}_J^H \boldsymbol{1}; \end{aligned} \tag{29c}$$

$$c_{m_{22}} = \boldsymbol{1}^T \boldsymbol{C}_J \left(Q_{JL} (\boldsymbol{D}) + Q_{JL} (\boldsymbol{D}) \right) \boldsymbol{C}_J^H \boldsymbol{1}; \tag{29d}$$

$$\boldsymbol{D} = \boldsymbol{R}_{f_m} \frac{{}^2 E}{|U_T| T_U} \boldsymbol{C}_{e_m} \boldsymbol{C}_T \left(Q_{JN} \left(\boldsymbol{1}\boldsymbol{1}^T \right) + Q_{JN} \left(\boldsymbol{1}\boldsymbol{1}^T \right) \right) \boldsymbol{C}_{e_m}^T \boldsymbol{C}_T^T;$$

$$\begin{aligned} c_{m_{23}} = \frac{{}^2\,{}^2 E\, {}^4_e}{|U_T| T_U} \Big(& |U_T| \mathrm{tr} \left(\boldsymbol{C}_{e_m} \boldsymbol{C}_T \boldsymbol{C}_T^T \boldsymbol{C}_{e_m}^T \right) \boldsymbol{1}^T \boldsymbol{C}_J \boldsymbol{C}_J^H \boldsymbol{1} + 1(U_O) \boldsymbol{1}^T \boldsymbol{C}_J^{|U_O|^+} \\ & \boldsymbol{C}_T^{|U_O|\;H} \boldsymbol{C}_{em}^{|U_O|^H} \boldsymbol{1}\boldsymbol{1}^T \boldsymbol{C}_{em}^{|U_O|} \boldsymbol{C}_T^{|U_O|} \boldsymbol{C}_J^{|U_O|^{+H}} \boldsymbol{1} \Big); \end{aligned} \tag{29e}$$

$$c_{m_{24}} = {}^2_e \mathrm{tr} \left(\boldsymbol{R}_{f_m} \frac{{}^2 E\, {}^2_e}{T_U} \boldsymbol{C}_{e_m} \boldsymbol{C}_T \boldsymbol{C}_T^T \boldsymbol{C}_{e_m}^T \right) \boldsymbol{1}^T \boldsymbol{C}_J \boldsymbol{C}_J^H \boldsymbol{1}; \tag{29f}$$

[5]This assumption of a knowledgeable EV constitutes the worst-case scenario. We adopted this assumption along with concurrent FD spoofing-estimation capabilities of the EV to ultimately quantify how much worse the EV can degrade the legitimate communication. Answering this question is the main aim of this paper.

and $1(\mathcal{U}_O)$ represents an indicator function of the set $\mathcal{U}_O$; $1(\mathcal{U}_O)=1$ if $\mathcal{U}_O\neq\emptyset$ and $1(\mathcal{U}_O)=0$ if $\mathcal{U}_O=\emptyset$. For presentation convenience, the expressions of Q_{JL} and Q_{JN} functions, and the details of obtaining the results in (28)-(29f) are omitted here, and provided in the Appendix.

As $N\to\infty$, the useful signal power $|b_m|^2$ in (26) can be expressed in a sum of two values as

$$|b_m|^2=\left|\mathbf{1}^T\boldsymbol{G}\boldsymbol{w}_m\right|^2=\frac{C_m^2}{\|\hat{\boldsymbol{h}}_m\|^2}\left|\sum_{i\in\mathcal{U}_T}\boldsymbol{g}_i^T\left(\sqrt{\rho P_I}\boldsymbol{h}_m+\sqrt{\frac{E}{|\mathcal{U}_T|T_U}}\sum_{k\in\mathcal{U}_T}\boldsymbol{g}_k+\boldsymbol{N}\boldsymbol{\epsilon}_i\right)\right|^2$$

$$\xrightarrow{N\to\infty}\frac{E|\mathcal{U}_T|N\sigma_e^2}{T_U\left(\rho P_I\beta_m+\frac{\phi E\beta_e\tau}{T_U}+\sigma_e^2\right)}+\frac{1}{N\left(\rho P_I\beta_m+\frac{\phi E\beta_e\tau}{T_U}+\sigma_e^2\right)}\left|\sum_{i\in\mathcal{U}_T}\boldsymbol{g}_i^T\left(\sqrt{\rho P_I}\boldsymbol{h}_m+\sqrt{\frac{E}{|\mathcal{U}_T|T_U}}\sum_{k\neq i}\boldsymbol{g}_k+\boldsymbol{N}\boldsymbol{\epsilon}_i\right)\right|^2. \quad (30)$$

The asymptotic convergence in (30) follows from applying Corollary 1 in [31]. Based on this result, the useful signal power $|b_m|^2$ in (26) asymptotically converges to a sum of deterministic value $|b_m^{(1)}|^2$ and a random value $|b_m^{(2)}|^2$ given as

$$|b_m^{(1)}|^2=\frac{E|\mathcal{U}_T|N\sigma_e^2}{T_U\left(\rho P_I\beta_m+\frac{\phi E\tau\beta_e}{T_U}+\sigma_e^2\right)}, \quad (31a)$$

$$|b_m^{(2)}|^2=\frac{\left|\sum_{i\in\mathcal{U}_T}\boldsymbol{g}_i^T\left(\sqrt{\rho P_I\tau^2}\boldsymbol{h}_m+\sqrt{\frac{\phi E\tau^2}{|\mathcal{U}_T|T_U}}\sum_{k\neq i}\boldsymbol{g}_k+\boldsymbol{N}\boldsymbol{\epsilon}_i\right)\right|^2}{N\left(\rho P_I\beta_m+\frac{\phi E\beta_e\tau}{T_U}+\sigma_e^2\right)}. \quad (31b)$$

Since all the parameters in (31a) and (31b) are of a finite order of magnitude except the number $N\to\infty$, then, by utilizing the result in our previous work [32, Theorem 1], we have $\left(|b_m^{(1)}|^2/|b_m^{(2)}|^2\right)\to O(N)$. This implies that $|b_m^{(1)}|^2$ is greater than $|b_m^{(2)}|^2$ by $O(N)$ order of magnitude, and therefore, we have the following tight lower bound on SINR_{e_m}

$$\mathrm{SINR}_{e_m}\geq\underline{\mathrm{SINR}}_{e_m}=\frac{p_i\,|b_m^{(1)}|^2}{\sum_{j\neq m}^{M}p_j\mathrm{var}\left(b_j\right)+c_m+\sigma_e^2}. \quad (32)$$

The tightness of (32) is validated by evaluating $\frac{|b_m^{(1)}|^2}{|b_m^{(2)}|^2}$ numerically for different realizations of $\boldsymbol{G}$ and $\boldsymbol{h}_m$ in Table I. The values are for $N=1000$, $\rho=0.5$ and the system parameter values given in Section V. Based on (32), and since the value of $\underline{\mathrm{SINR}}_{e_m}$ is deterministic, the lower bound on the ergodic rate of the EV, $\underline{R}_{e_m}$ is

$$\underline{R}_{e_m}=\log_2\left(1+\underline{\mathrm{SINR}}_{e_m}\right)\leq R_{e_m}. \quad (33)$$

Based on the perfect knowledge of the EV, and by making use of Theorem 1 in [3], the EV is capable of calculating an upper bound on the ergodic rate of the attacked UE (the second bound), $\overline{R}_m$, as

TABLE I. RELATIVE VALUES OF $|b_m^{(1)}|^2$ AND $|b_m^{(2)}|^2$.

Realization	1st	2nd	3rd	4th
$\lvert b_m^{(1)}\rvert^2$	1.2×10^{-4}	1.2×10^{-4}	1.2×10^{-4}	1.2×10^{-4}
$\lvert b_m^{(2)}\rvert^2$	0.4×10^{-7}	0.1×10^{-6}	0.3×10^{-7}	0.2×10^{-7}
$\frac{\lvert b_m^{(1)}\rvert^2}{\lvert b_m^{(2)}\rvert^2}$	2.7×10^{3}	0.9×10^{3}	3.5×10^{3}	4.7×10^{3}

$$\overline{R}_m\xrightarrow{N\to\infty}\log_2\left(1+\overline{\mathrm{SINR}}_m\right), \quad (34)$$

where

$$\overline{\mathrm{SINR}}_m=\frac{p_i\,\mathbb{E}\left[|a_{m,m}|^2\right]}{\sum_{j\neq m}^{M}p_j\,\mathrm{var}\left(a_{m,j}\right)+\frac{(1-\phi)E}{T_D}\mathrm{var}\left(\boldsymbol{f}_m^T\boldsymbol{w}_m\right)+\sigma_u^2}. \quad (35)$$

The asymptotically converged values of $\mathrm{var}(b_{j\neq m})$ and $\mathbb{E}[|a_{m,m}|^2]$ in (32) and (35) (given in (36a) and (36b)), can be obtained by following a similar analysis to that in our previous work [3, Appendix C], while the value of $\mathrm{var}(\boldsymbol{f}_m^T\boldsymbol{w}_m z)$ (given in (36c)) can be obtained by following a similar analysis to that in Section C of the Appendix.

$$\mathrm{var}(b_{j\neq m})\xrightarrow{N\to\infty}|\mathcal{U}_T|\beta_e, \quad (36a)$$

$$\mathbb{E}\left[|a_{m,m}|^2\right]\xrightarrow{N\to\infty}\frac{\beta_m\left(N\rho P_I\beta_m+\frac{\phi E}{T_U}\beta_e+\sigma_u^2\right)}{\rho P_I\beta_m+\frac{\phi E}{T_U}\beta_e+\sigma_u^2}, \quad (36b)$$

$$\mathrm{var}(\boldsymbol{f}_m^T\boldsymbol{w}_m)=\frac{1}{\mathrm{tr}\left(\boldsymbol{R}_{f_m}\right)}\Bigg[\rho^2P_I\sigma_{e_m}^2\left(|\mathcal{U}_J|\mathrm{tr}\left(\boldsymbol{C}_{e_m}^H\boldsymbol{C}_{e_m}\right)+\mathrm{diag}^H\left(\boldsymbol{C}_{e_m}\right)\mathrm{diag}\left(\boldsymbol{C}_{e_m}\right)\right)$$
$$+\frac{\tau^2E\beta_{e_m}}{|\mathcal{U}_T|T_U}\mathrm{tr}\left(\boldsymbol{C}_{e_m}\left(\boldsymbol{C}_T\left(Q_{TL}\left(\mathbf{1}\mathbf{1}^T\right)+Q_{TN}\left(\mathbf{1}\mathbf{1}^T\right)\right)\boldsymbol{C}_T^T+|\mathcal{U}_T|\sigma_e^2\boldsymbol{C}_T\boldsymbol{C}_T^T\right)\boldsymbol{C}_{e_m}^T\right)+\beta_{e_m}\sigma_e^2\mathrm{tr}\left(\boldsymbol{C}_{e_m}\boldsymbol{C}_{e_m}^T\right)$$
$$-\rho^2P_I\sigma_{e_m}^2\mathrm{tr}^2\left(C_{e_m}\right)\Bigg]. \quad (36c)$$

IV. ACTIVE ATTACK SCENARIOS

The system objectives include the maximization of the minimum downlink UE rate, $\min_i\underline{R}_i$, by the BS which is unaware of the active attack of the EV; and a complex goal of increasing the eavesdropping information rate while decreasing the information rate of the attacked UE by the EV. The EV's objective is simply the minimization of rate difference

$$R_{D_m}=\overline{R}_m-\underline{R}_{e_m}. \quad (37)$$

Remark 3: Since we are considering the severity of the active attack from the EV's side, we use the rate difference, R_{D_m}, instead of the well-known secrecy rate metric, $\max(\overline{R}_m-\underline{R}_{e_m},0)$. That is because the zero and negative values of secrecy rate are equivalent from the BS's point of view, since the attacked UE loses total security in both cases. However, in our case, the active EV aims at maximizing its information rate over the attacked UE. Therefore, the negative values of R_{D_m} are better than zero values from the EV's point of view.

In this section, we will examine how badly the EV can do in degrading the information security under three different cases: 1) The BS uses an optimized power allocation for downlink, and the EV knows the optimized power values; 2) The BS uses equal power allocation for downlink, and the EV is aware of this equal power allocation; 3) The BS uses an optimized power allocation for downlink, and the EV is unaware of that and assumes equal power allocation at the BS [6].

Before discussing these three cases, let us derive the convex formulations for the max-min rate and eavesdropping optimization problems at the BS and the EV, respectively.

The downlink power allocation is formulated as a max-min optimization problem which aims at maximizing the minimum UE rate. Since the logarithmic function is monotonically increasing in its argument, the max-min optimization problem is formulated based on the epigraph of $\underline{\mathrm{SINR}}_i$ as

$$\begin{aligned} \underset{\{p_i\},\gamma}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & \underline{\mathrm{SINR}}_i \geq \gamma, && \text{(38a)} \\ & \sum_{i=1}^{M} p_i \leq P_t, \; p_i \geq 0, \; \forall i, && \text{(38b)} \end{aligned}$$

where P_t is the total power budget at the BS. Problem (38) is non-convex due to the upper bounded quadratic constraint (38a). However, we can recast (38) into a feasibility problem with a fixed value of γ (as given in (39)); and then we use it to search for the global optimal objective value, γ, using the bisection algorithm [33, Subsection 4.2.5].

$$\begin{aligned} \underset{\{p_i\}}{\text{maximize}} \quad & 0 \\ \text{subject to} \quad & p_i \, |\mathbb{E}[a_{i,i}]|^2 \geq \gamma \left(\sum_{j=1}^{M} p_j \, \mathrm{var}(a_{i,j}) + \sigma_u^2 \right), && \text{(39a)} \\ & \sum_{i=1}^{M} p_i \leq P_t, \; p_i \geq 0, \; \forall i. && \text{(39b)} \end{aligned}$$

Given that $\gamma^\star$ is the optimal value of the bisection algorithm search, the optimal power allocation is $\{p_i^\star\}$ that corresponds to $\gamma^\star$.

By looking at the definition of $\underline{\mathrm{SINR}}_i$ in (23), it is obvious that the optimized power allocation $\{p_i^\star\}$ is independent from the allocated spoofing and jamming antenna subsets U_T and U_J at the EV.

The EV is required to obtain the optimal balance between the following: 1) The power allocated for active spoofing, $\frac{\phi E}{T_U}$, and the power allocated for jamming, $\frac{(1-\phi)E}{T_D}$, under some total power constraint; 2) The number of antennas allocated for U_T and U_J subsets. Here, it is worth mentioning that the EV only needs to select the numbers of training and jamming antennas, but not their indices. This is because the EV is optimizing the ergodic performance; furthermore, the self-interference depends only on the cardinalities of U_T and U_J. The optimal balance should lead to the best rate difference and it is obtained by solving the following problem

$$\begin{aligned} \underset{\phi,\, |U_T|,\, |U_J|}{\text{minimize}} \quad & R_{D_m} \\ \text{subject to} \quad & 0 \leq \phi \leq 1, && \text{(40a)} \\ & |U_T \cup U_J| = K. && \text{(40b)} \end{aligned}$$

Constraint (40b) indicates that the EV uses all its antennas. Since R_{D_m} is convex in ϕ, and the selection ranges of ϕ, $|U_T|$ and $|U_J|$ defined by (40a) and (40b) are small, therefore (40) can be solved by a nested one-dimensional search (such as the bisection algorithm) over the range $0 \leq \phi \leq 1$ and $1 \leq |U_T|, |U_J| \leq K$ for the optimal (minimum) R_{D_m} value.

Our primary assumption of knowledgeable EV indicates that the EV perfectly knows its own channels and both the beamforming vector and the channel of the attacked UE. However, the EV awareness of the power allocation strategies at the BS is crucial and can have a significant impact on the severity of the dual-active attack. Therefore, we consider all possible cases (three cases) for EV awareness about the power allocation strategies at the BS. The severity of the dual-active attack is examined under the suggested cases as follows:

1) The first case (*case 1*) takes place when the BS performs optimal power allocation and the EV is aware of that optimized power. Thus, problem (40) is solved with $R_{D_m}(\{p_i\} = \{p_i^\star\})$. Given $\phi^\star$, $|U_T|^\star$ and $|U_J|^\star$ is the optimal solution of (40), the severity of the dual-active attack is examined by $R_{D_m}(\phi^\star, |U_T|^\star, |U_J|^\star, \{p_i^\star\})$.
2) The second case (*case 2*) takes place when the BS uses equal power allocation and the EV is aware of the equal power allocation. Thus, problem (40) is solved with $R_{D_m}(\{p_i\} = \{p_i = \frac{P_t}{M}\})$. The severity of the dual-active attack is examined by $R_{D_m}(\phi^\star, |U_T|^\star, |U_J|^\star, \{p_i\} = \{p_i = \frac{P_t}{M}\})$.
3) The third case (*case 3*) takes place when the BS performs optimal power allocation and the EV is unaware of that and assumes equal power allocation. Thus, problem (40) is solved with $R_{D_m}(\{p_i\} = \{p_i = \frac{P_t}{M}\})$. The severity of the dual-active attack is examined by $R_{D_m}(\phi^\star, |U_T|^\star, |U_J|^\star, \{p_i^\star\})$.

Remark 4: Although the work in this paper considers the existence of one EV, the provided analysis can be applied to the case of multiple EVs by considering these multiple EVs as one equivalent EV of LK antennas, where L is the number of EVs and K is the number of antennas for each EV. This equivalent EV can achieve the optimal active attack based on the following facts:

- The multiple EVs have statistically independent channels towards the attacked UE and the BS, then the individual jamming signals (received at the attacked UE) are also independent, and therefore, the variance of the total jamming signal is the sum of the individual jamming signal variances.
- Managing the inter-EV interference is not relevant since the EVs jam and intercept information over orthogonal time periods.
- Power balancing between EVs is not relevant due to the first fact and since each EV has its own power budget.

From the adversary attacker perspective, the advantages of using a single FD EV over multiple separate EVs are the ease of implementation and the saving of the resources required for sharing control information.

[6]Since the BS and the EV ergodic objectives, R and R_{D_m}, are functions of the statistical channel state informations (CSIs), the optimization solutions at the BS and the EV remain constant over multiple coherence time-blocks. Also, the EV knowledge of the optimized power allocation values is possible based on the EV knowledge of the statistical CSIs and power allocation method used at the BS.

V. Evaluations

This section evaluates the severity of the concurrent attack of the FD EV for the aforementioned three cases. In our evaluations, we assume that the UEs and the EV are randomly located within the BS coverage area, and the EV has no means to optimally localize itself with respect to the attacked UE. The large scale fading parameters are calculated based on the distance-based path loss model as $_i = d_i^{\ \gamma_u}\ 10^{\frac{\nu_u}{10}}$, $_e = d^{\ \gamma_e}\ 10^{\frac{\nu_e}{10}}$ and $_{e_m} = d_m^{\ \gamma_{eu}}\ 10^{\frac{\nu_{eu}}{10}}$, where d_i; d and d_m represent the distances from UE$_i$ and the EV to the BS; and between the EV and UE$_m$, respectively. $f\ _u;\ _e;\ _{eu}g$ are random shadowing fading coefficients (in dB) that follow a zero-mean Gaussian distribution with standard deviations $_U;\ _E$ and $_{EU}$ (also in dB), respectively. $f\ _u;\ _e;\ _{eu}g$ are the path loss exponents. Since the relative height of the BS antennas to the UE and the EV antenna(s) is larger than the relative height of the EV antennas to UE antenna, therefore, unless otherwise stated, we assume the following relative relations $_{eu} > \ _u;\ _e$ and $_{eu} > \ _u;\ _e$. The EV antennas are evenly and linearly spaced by distances $fd_{i,j} = ji\quad jj\ g$. Each antenna (dipole antenna) has a gain of 2.15 dB above an isotropic antenna. We assume the Friis free space equation to calculate the large-scale fading between different antenna channels as $_{i,j\ne i} = \frac{1.64}{(4\pi(i\quad j))^2}$. The self-antenna loss $_{i,i}$ is related to the analog self-interference cancellation and it is selected as 60 dB [11]. The span of the antenna array is in the range of tenths of centimeters; therefore, the values of K-factors are selected (randomly) between 40 dB and 60 dB with the condition $K_{i,j} > K_{i,m}$ for $ji\quad jj < ji\quad mj$ [13].

We assume a coherence bandwidth of B_C = 200 KHz and a coherence time of T_C = 1 ms, thus, the length of coherence time-block that ensures flat-and-slow fading channel is $Q = T_C B_C$ = 200 symbols and $T_s = 1{=}B_C = 5\quad s$ [22], [34]. $\ = M$ is the optimal value for massive MIMO networks without active eavesdropping as investigated in [35], however, we investigate our actively attacked network for $\quad M$.

Unless otherwise stated, the rest of the system parameters are selected as N = 128, M = 4, K = 10, $\ = K$ = 10, $f\ _b^2;\ _e^2;\ _u^2 g = \ 70$ dB, $f\ _U;\ _E g$ = 2, $_{EU}$ = 2:5, $_u = \ _e$ = 3, $_{eu}$ = 3:5, P_t = 1 W, $P_I = P_{si}$ = 1 W. The total energy at the EV is selected such that the average spoofing-jamming power $\frac{E}{Q\ T_s}$ = 1 W, and therefore E = 1 mJ. Next, we discuss the numerical simulations that examine the severity of the dual-active attack by the EV under the aforementioned cases.

We assume that the UEs and the EV are located randomly in a cell of 100 m diameter with the BS at the centre. The EV attacks the weakest UE that has the smallest channel gain such that the average received spoofing power relative to training power is at its maximum.

Fig. 2 shows how the information rates for both UE$_m$ and the EV change versus the energy splitting ratio for case 1. The corresponding rate difference is described by the square-marked red plot given in Fig. 4. For each value of $\ $, the optimal values $jU_T^\star j$ and $jU_J^\star j$ are obtained. It can be seen that as the EV allocates more power for active spoofing (i.e., less power for jamming), the rate difference decreases.

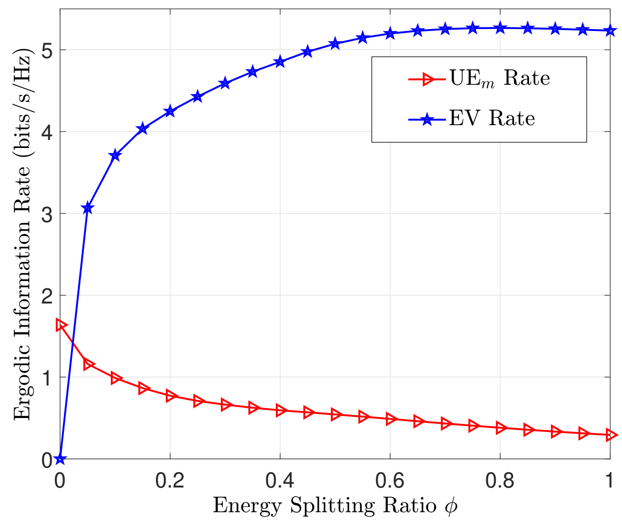

Fig. 2. UE$_m$ and the EV rates versus ϕ for case 1. γ_{eu} = 3.5 > γ_u = γ_e = 3.

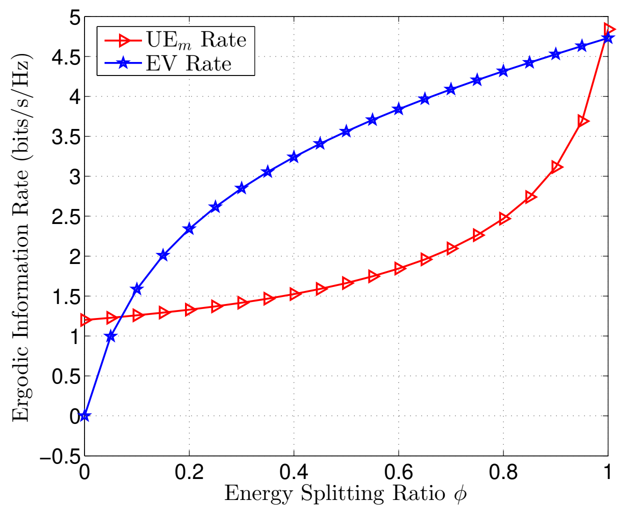

Fig. 3. UE$_m$ and the EV rates versus ϕ for case 1. γ_e = 3 > γ_u = 2.5 > γ_{eu} = 2.

The channel between the EV and the UE is weak; hence it is better for the EV to allocate all the power to spoofing. This inverse proportional relationship between R_{D_m} and $\ $ is further justified as follows. Increasing the spoofing power (i.e., increasing $\ $) results in:

1) Increasing the received signal power at the EV.
2) Decreasing the received signal power at UE$_m$.
3) Motivating the BS to allocate more power to UE$_m$ (please see (2) and its effect on (39)), which in turn increases the change mentioned in the first two points.
4) Decreasing the received jamming power at UE$_m$.

The consequences of increasing $\ $ are a consistent decrease in $\overline{R}_m$ (justified by 2), 3) and 4) and a consistent increase in $\underline{R}_{e_m}$ (justified by 1) and 3). This justification is valid by

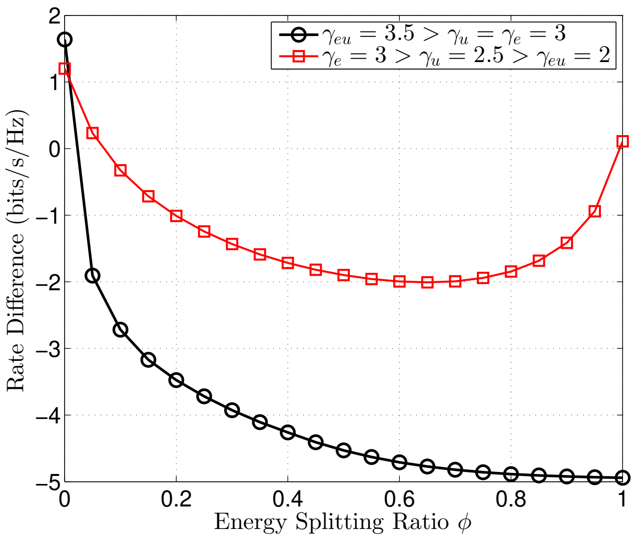

Fig. 4. Rate difference versus ϕ for case 1.

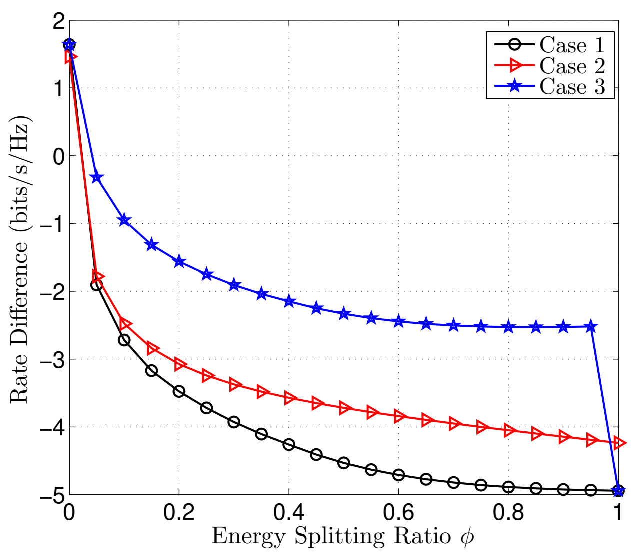

Fig. 5. Rate difference versus ϕ for all cases.

our assumption $\gamma_{eu} > \gamma_u, \gamma_e$, which gives $\underline{R}_{e_m}$ the advantage over $\overline{R}_m$, $\forall\phi$ (i.e., a consistent decrease in R_{D_m}). However, the assumptions $\gamma_{eu} > \gamma_u, \gamma_e$ do not necessarily reflect all channel environments in reality as will be seen next.

Fig. 3 shows the information rates for both UE$_m$ and the EV versus the energy splitting ratio for case 1. However, Fig. 3 considers a special case when UE$_m$ enjoys better channel conditions than the EV; and the EV has a strong jamming link towards UE$_m$ as $\gamma_e = 3 > \gamma_u = 2.5 > \gamma_{eu} = 2$. The corresponding rate difference is described by the circle-marked black plot given in Fig. 4. This special selection of channel condition gives a comparable advantage for both $\underline{R}_{e_m}$ and $\overline{R}_m$. It can be seen that the optimal energy splitting ratio is $\phi^\star = 0.65$. Using the same justification (the four points above) used

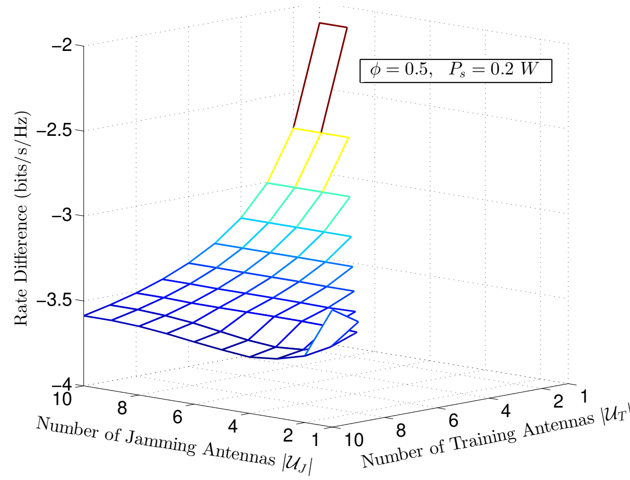

Fig. 6. Rate difference versus $(|\mathcal{U}_T|, |\mathcal{U}_J|)$ for case 1.

for discussing Fig. 2 we notice;

1) For $\phi < \phi^\star = 0.65$, increasing ϕ gives $\underline{R}_{e_m}$ the advantage over $\overline{R}_m$, and thus the rate of change inequality $\frac{\partial \underline{R}_{e_m}}{\partial \phi} > \frac{\partial \overline{R}_m}{\partial \phi}$ takes place for $\phi < \phi^\star = 0.65$ (i.e., a consistent decrease in R_{D_m}).
2) Conversely, for $\phi > \phi^\star = 0.65$, increasing ϕ gives $\overline{R}_m$ the advantage over $\underline{R}_{e_m}$, and thus $\frac{\partial \underline{R}_{e_m}}{\partial \phi} < \frac{\partial \overline{R}_m}{\partial \phi}$ takes place for $\phi > \phi^\star = 0.65$ (i.e., a consistent increase in R_{D_m}).

Fig. 5 shows the rate difference versus the energy splitting ratio for the three cases. It can be seen that case 1 is the most severe followed by cases 2 then 3. This is expected since in case 1 the EV knows $\{p_i^\star\}$ and can motivate the BS to allocate more power to UE$_m$, hence its designed attack is optimal. However, in case 2, the EV knows the equal power allocation at the BS and can not motivate the BS to allocate more power to UE$_m$, hence its designed attack is less optimal than case 1. In the less severe case 3, the EV does not know the equal power allocation at the BS and can not motivate the BS to allocate more power to UE$_m$.

For $\phi = 1$, case 1 and 3 achieve the same rate difference as the EV will use all the power for spoofing in both cases and hence $|\mathcal{U}_T| = K$ and $|\mathcal{U}_J| = 0$ is the optimal solution for both cases and the dual-active attack is examined equivalently for both cases as $R_{D_m}(\phi^\star = 1, |\mathcal{U}_T|^\star = K, |\mathcal{U}_J|^\star = 0, \{p_i^\star\})$.

The impact of the numbers of spoofing and jamming antennas $|\mathcal{U}_T|$ and $|\mathcal{U}_J|$ is shown in Fig. 6 for $\phi = 0.5$ and $P_{si} = 0.2\ W$. Since the pairs of $(|\mathcal{U}_T|, |\mathcal{U}_J|)$ used to generate the 3-D plot satisfy $|\mathcal{U}_T \cup \mathcal{U}_J| = K$, then $\{(|\mathcal{U}_T|, |\mathcal{U}_J|)\} \subset \{1, 2, \ldots, |\mathcal{U}_T|\} \times \{1, 2, \ldots, |\mathcal{U}_J|\}$ and that is noticeable from the missing part in the 3-D plot in Fig. 6. We can see that the optimal numbers are $|\mathcal{U}_T^\star| = 10$, and $|\mathcal{U}_J^\star| = 5$. Generally, a larger number of training antennas $|\mathcal{U}_T|$ performs better. This might be due to the performance increment obtained by the $|\mathcal{U}_T|$ antenna diversity gain (see (25)-(26)) at the EV, which always exceeds the decrement imposed by spoofing self-interference at the jamming antennas (during estimation phase). On the other side, the number of jamming antennas needs to find a

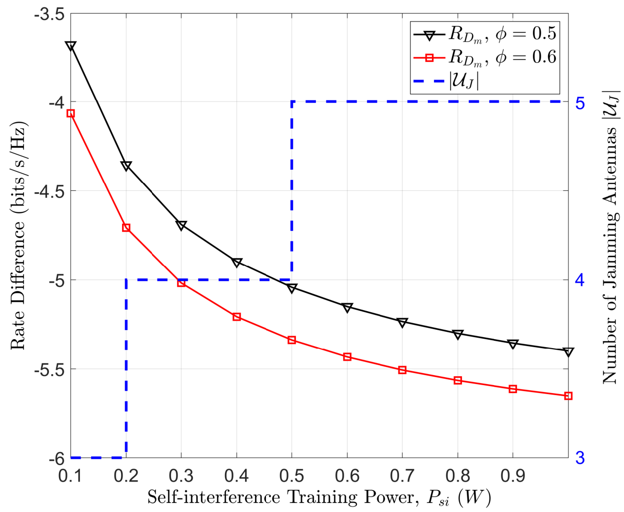

Fig. 7. Rate difference versus self-interference training power.

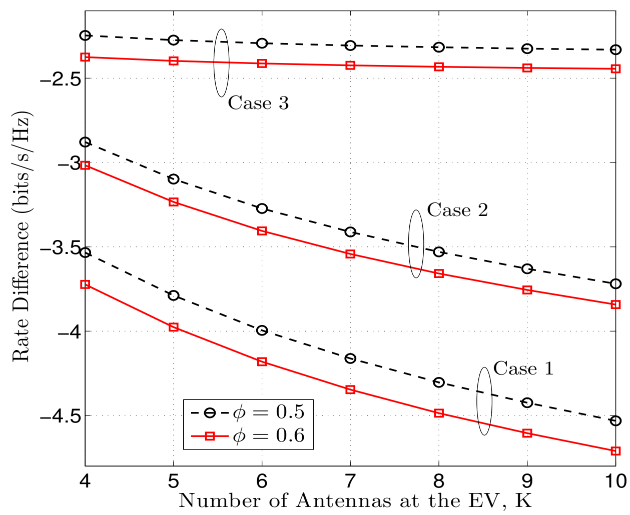

Fig. 8. Rate Difference versus number of antennas at the EV.

balance between the jamming power at UE_m and the jamming self-interference power at U_T. This balance is affected by the self-interference cancellation (i.e., the estimation accuracy $\boldsymbol{H}$) as will be seen next.

In Fig. 7, we demonstrate the impact of self-interference cancellation (via changing the value of self-interference training power P_{si}) on both the optimal rate difference, R_{D_m}, and the associated number of jamming antennas $|U_J^\star|$, for $|U_T| = K$ and for two values of the energy splitting ratio $= 0.5$ and $= 0.6$. The figure has two vertical axes; the left-axis is for R_{D_m} and the right-axis is for the staircase function that represents the number of jamming antennas $|U_J^\star|$. As expected, R_{D_m} decreases as P_{si} increases (self-interference decreases). The same trend was observed for both $= 0.5$ and $= 0.6$. Also, at high self-interference values (low P_{si} values), the optimal number of U_J antennas is small. As the self-interference decreases (P_{si} increases), more antennas are used for jamming.

The rate difference versus the number antennas at the EV, K, and length of training sequences, , are depicted in Figs. 8 and 9, respectively. We can see that R_{D_m} decreases with K as the larger values of K give the EV more freedom to balance its attack parameters, particularly $|U_T|$ which determines the receive diversity gain of the EV. Fig 9, shows that R_{D_m} increases with . For both results, case 1 is the most severe followed by case 2 then case 3.

In Fig. 10, we demonstrate the impact of the self-interference channel training length, , on the self-interference channel estimation errors. For that purpose we use the mean square errors (MSEs) given in (41) and (42); E_J and E_T; and their approximations, $\hat{E}_J$ and $\hat{E}_T$, that are valid over the range of larger values of based on the fact $\lim_{\eta \to K} \boldsymbol{C}_J = \lim_{\eta \to K} \boldsymbol{C}_J^H \frac{I_{|U_T|}}{\eta \sqrt{P_{si}}}$. The figure depicts the plots of E_J, $\hat{E}_J$, and the MSE $\mathrm{E}[|| \; \hat{\boldsymbol{H}}_J||^2]$ obtained by averaging over 10000 realizations for $|U_J| = |U_T| = \frac{K}{2} = 5$. At the shortest length, $= K$, we can calculate the order of the average MSE per channel

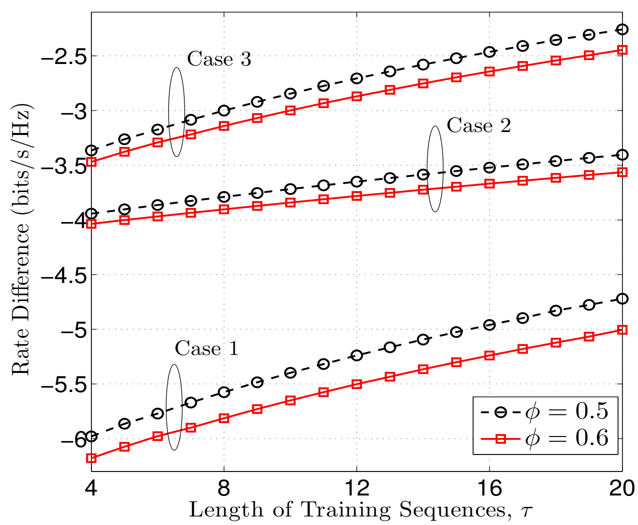

Fig. 9. Rate Difference versus length of training sequences.

coefficient as $\frac{0.8 \quad 10^{\;6}}{|U_J| \; |U_T|} \quad 10^{\;8}$. This figure is very much smaller than the average self-interference channel coefficient gain $\mathrm{E}[\frac{1.64}{(4\pi(i \quad j))^2}] \quad 10^{\;4}$, $i; \; j \; 2 \; f1; \; 2; \; :::; K g$. Accordingly, the selection of the shortest length, $= K$, will result in a very accurate estimate, and hence, the impact of increasing on the performance will not be noticeable.

$$E_J(\;) = \mathrm{E}\left[\left\| \quad \hat{\boldsymbol{H}}_J \right\|^2\right] = \mathrm{tr}\left(\boldsymbol{C}_J \boldsymbol{R}_J \boldsymbol{C}_J^H + \quad {}^2_e |U_J| \boldsymbol{C}_J \boldsymbol{C}_J^H\right); \tag{41}$$

$$E_T(\;) = \mathrm{E}\left[\left\| \quad \hat{\boldsymbol{H}}_T \right\|^2\right] = \mathrm{tr}\left(\boldsymbol{C}_T \boldsymbol{R}_T \boldsymbol{C}_T^H + \quad {}^2_e |U_T| \boldsymbol{C}_T \boldsymbol{C}_T^H\right); \tag{42}$$

$$\hat{E}_J(\;) = \hat{E}_J(\;) = \frac{{}^2_e |U_J||U_T|}{P_{si}}: \tag{43}$$

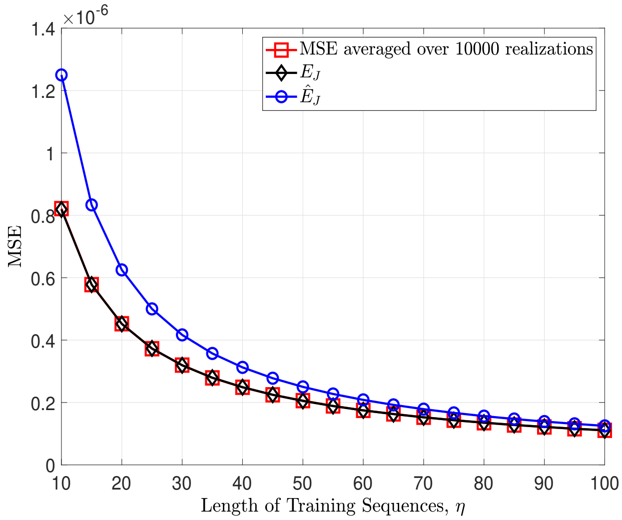

Fig. 10. MSE versus length of training sequences, η.

VI. Conclusions

In this paper, the severity of a concurrent spoofing-jamming attack of an FD multi-antenna EV in a massive MIMO system has been considered. The general expression for the ergodic rate difference for any possible overlap between spoofing and jamming antenna subsets has been derived. The residual spoofing and jamming self-interferences and their statistical dependencies have been considered in the derived expression. The most important insights for the considered attack are as follows: 1) Based on the derived expressions for the statistics of residual self-interferences (correlated spoofing and jamming self-interferences), we showed that the proposed FD EV can explore those statistics to optimize its attack by adjusting its spoofing and jamming signal powers; 2) Without the knowledge of the statistical residual self-interference, the optimal selection of spoofing and jamming antennas sets is not achievable; 3) Numerical results revealed a significant security threat to the legitimate communication even with a small number of antennas and a power budget equal to that of the attacked UE; 4) A marginal increase in the rate difference was obtained by increasing the length of training sequences. These insights are useful for designing potential countermeasures for this new type of active attack. Considering the same attack in cell-free massive MIMO systems is a potential future work.

Appendix

A. Expressions of Q_{TL} and Q_{TN} Functions

Let us write $\boldsymbol{H}_T$ and $\boldsymbol{H}_J$ as a sum of the deterministic LoS and the random NLoS parts as $\boldsymbol{H}_T = \boldsymbol{H}_{TL} + \boldsymbol{H}_{TN}$ and $\boldsymbol{H}_J = \boldsymbol{H}_{JL} + \boldsymbol{H}_{JN}$, respectively, where

$$\mathrm{E}[\boldsymbol{H}_{TL}]_{i,l} = \bar{\boldsymbol{H}}_{TL} = \sqrt{\frac{{}_{n(i),l} K_{n(i),l}}{K_{n(i),l} + 1}}\, e^{\,j\frac{2\pi d_{n(i),l}}{\lambda}}; \tag{44}$$

$$\mathrm{E}[\boldsymbol{H}_{TN}]_{i,l} = \bar{\boldsymbol{H}}_{TN} = \sqrt{\frac{{}_{n(i),l}}{K_{n(i),l} + 1}}; \tag{45}$$

$$\mathrm{E}[\boldsymbol{H}_{JL}]_{i,l} = \bar{\boldsymbol{H}}_{JL} = \sqrt{\frac{{}_{i,n(l)} K_{i,n(l)}}{K_{i,n(l)} + 1}}\, e^{\,j\frac{2\pi d_{i,n(l)}}{\lambda}}; \tag{46}$$

$$\mathrm{E}[\boldsymbol{H}_{JN}]_{i,l} = \bar{\boldsymbol{H}}_{JN} = \sqrt{\frac{{}_{i,n(l)}}{K_{i,n(l)} + 1}}; \tag{47}$$

$$n(i) = K \quad |U_J| + i; \quad i = 1; 2; \ldots; |U_J|. \tag{48}$$

For an arbitrary matrix $\boldsymbol{A}$, it is simple to conclude that $\mathrm{E}[\boldsymbol{H}_T \boldsymbol{A} \boldsymbol{H}_T^H] = \bar{\boldsymbol{H}}_{TL} \boldsymbol{A} \bar{\boldsymbol{H}}_{TL}^H + \mathrm{Diag}(\bar{\boldsymbol{H}}_{TN} \mathrm{Diag}(\boldsymbol{A}) \bar{\boldsymbol{H}}_{TN}^H)$. For notational convenience, the terms in the last expression are represented by the functions $Q_{TL}(\boldsymbol{A}) = \bar{\boldsymbol{H}}_{TL} \boldsymbol{A} \bar{\boldsymbol{H}}_{TL}^H$ and $Q_{TN}(\boldsymbol{A}) = \mathrm{Diag}\left(\bar{\boldsymbol{H}}_{TN} \mathrm{Diag}(\boldsymbol{A}) \bar{\boldsymbol{H}}_{TN}^H\right)$. Similarly, $Q_{JL}(\boldsymbol{A}) = \bar{\boldsymbol{H}}_{JL} \boldsymbol{A} \bar{\boldsymbol{H}}_{JL}^H$ and $Q_{JN}(\boldsymbol{A}) = \mathrm{Diag}\left(\bar{\boldsymbol{H}}_{JN} \mathrm{Diag}(\boldsymbol{A}) \bar{\boldsymbol{H}}_{JN}^H\right)$.

B. Derivation of the $\boldsymbol{R}_{T_{si}}$ Matrix in (15)

By using (12), we can write

$$\mathrm{E}\left[\hat{\boldsymbol{H}}_T \mathbf{1}\mathbf{1}^T \hat{\boldsymbol{H}}_T^H\right] = \mathrm{E}\left[\boldsymbol{C}_T \boldsymbol{H}_T \mathbf{1}\mathbf{1}^T \boldsymbol{H}_T^H \boldsymbol{C}_T^H\right] + \mathrm{E}[\boldsymbol{C}_T \boldsymbol{N}_T \mathbf{1}\mathbf{1}^T \boldsymbol{N}_T^H \boldsymbol{C}_T^H] = \boldsymbol{C}_T \left(Q_{TL}\left(\mathbf{1}\mathbf{1}^T\right) + Q_{TN}\left(\mathbf{1}\mathbf{1}^T\right)\right) \boldsymbol{C}_T^H + |U_T| \quad {}_e^2 \boldsymbol{C}_T \boldsymbol{C}_T^H. \tag{49}$$

The first term on the right-hand side of the second equality follows from using the functions Q_{TL} and Q_{TN}. The second term on the right-hand side of the second equality follows from the statistical independence of the $\boldsymbol{N}_T$'s rows and vectors. By substituting the value of (49) in $\boldsymbol{R}_{T_{si}} = \frac{\phi E}{|U_T| T_U} \mathrm{E}[\hat{\boldsymbol{H}}_T \mathbf{1}\mathbf{1}^T \hat{\boldsymbol{H}}_T^H]$, we get the result in (15).

C. Derivation of the Variance c_m in (29)

Starting from $c_m = \frac{(1-\phi)E}{T_D} \mathrm{var}(\mathbf{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m z)$, we apply the definition of statistical variance to calculate

$$\mathrm{var}\left(\mathbf{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m z\right) = \mathrm{E}\left[\left|\mathbf{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m z - \mathrm{E}\left[\mathbf{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m z\right]\right|^2\right]$$
$$= \mathrm{E}\left[\mathbf{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m \boldsymbol{w}_m^H \hat{\boldsymbol{H}}_J^H \mathbf{1}\right] - \left|\mathrm{E}\left[\mathbf{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m\right]\right|^2. \tag{50}$$

$$\mathrm{E}\left[\mathbf{1}^T \hat{\boldsymbol{H}}_J \boldsymbol{w}_m\right] = \sqrt{\frac{E}{|U_T| T_U \mathrm{tr}(\boldsymbol{R}_{f_m})}} \mathbf{1}^T \boldsymbol{C}_J \mathrm{E}\left[\boldsymbol{H}_J \boldsymbol{C}_{e_m} \boldsymbol{C}_T \boldsymbol{H}_J^H\right] \boldsymbol{C}_J^H \mathbf{1}$$
$$+ \sqrt{\frac{E}{|U_T| T_U \mathrm{tr}(\boldsymbol{R}_{f_m})}} \mathbf{1}^T \boldsymbol{C}_J \mathrm{E}[\boldsymbol{N}_J \boldsymbol{C}_{e_m} \boldsymbol{C}_T \boldsymbol{N}_T] \mathbf{1}. \tag{51}$$

The last equation, (51), is obtained by making use of (12), (14a), (17), and excluding the terms which are made-up of multiplication of independent zero mean random values. The first term can be represented as

$$\sqrt{\frac{E}{|U_T||T_U|\,\mathrm{tr}\,(\boldsymbol{R}_{f_m})}}\mathbf{1}^T\boldsymbol{C}_J\left(Q_{JL}\left(\boldsymbol{C}_{e_m}\boldsymbol{C}_T\right)+Q_{JN}\left(\boldsymbol{C}_{e_m}\boldsymbol{C}_T\right)\right)\mathbf{1}$$
$$=\mathrm{tr}^{-\frac{1}{2}}\left(\boldsymbol{R}_{f_m}\right)c_{m_{11}}. \tag{52}$$

The expectation of the second term in (51), $\mathrm{E}[\boldsymbol{N}_J\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{N}_T]$, exists only if there is an overlap between the training antennas U_T and the jamming antennas U_J, i.e, $1 \quad |U_O| \quad K$. In this case, we should note that the last $|U_O|$ rows of $\boldsymbol{N}_J$, $\boldsymbol{N}_J^{+|U_O|}$, are equal to the first $|U_O|$ rows of $\boldsymbol{N}_T$, $\boldsymbol{N}_T^{+|U_O|}$. Based on the previous and the property of the unitary matrix multiplication, $\boldsymbol{\Gamma}^{|U_J|^+}{}^H\boldsymbol{\Gamma}^{|U_T|} = K\boldsymbol{I}_{|U_O|}$, $\mathrm{E}[\boldsymbol{N}_J\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{N}_T]$ will result in an $|U_T|$-by-$|U_T|$ matrix whose entries are zeros except the top-right $|U_O|$-by-$|U_O|$ sub-matrix which consists of non-zero entries. By utilizing this sparse structure, the value of the second term in (51) is simplified as

$$1\,(U_O)\quad\sqrt{\frac{E\quad{}^4_e}{|U_T||T_U|}}\;\mathbf{1}^T_{|U_T|}\boldsymbol{C}_J^{|U_O|^+}\;\boldsymbol{C}_T^{|U_O|}\quad{}^H\boldsymbol{C}_{e_m}^{|U_O|}{}^H\mathbf{1}_{|U_O|}$$
$$=\mathrm{tr}^{-\frac{1}{2}}\left(\boldsymbol{R}_{f_m}\right)c_{m_{12}}. \tag{53}$$

By following comparable steps as in the analysis of (51), the expectation $\mathrm{E}[\mathbf{1}^T\quad\hat{\boldsymbol{H}}_J\boldsymbol{w}_m\boldsymbol{w}_m^H\quad\hat{\boldsymbol{H}}_J^H\mathbf{1}]$ in (50) can be analysed as a sum of the following four expectations

$$\frac{{}^2\;E}{|U_T||T_U|\mathrm{tr}(\boldsymbol{R}_{f_m})}\mathbf{1}^T\boldsymbol{C}_J\mathrm{E}\left[\boldsymbol{H}_J\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{H}_J^H\mathbf{1}\mathbf{1}^H\boldsymbol{H}_J\boldsymbol{C}_T^T\boldsymbol{C}_{e_m}^T\boldsymbol{H}_J^H\right]$$
$$\boldsymbol{C}_J^H\mathbf{1}=\frac{c_{m_{21}}}{\sqrt{\boldsymbol{R}_{f_m}}}; \tag{54}$$

$$\mathbf{1}^T\boldsymbol{C}_J\mathrm{E}\left[\boldsymbol{H}_J\left(\boldsymbol{w}_m\quad\sqrt{\frac{{}^2\;E}{|U_T||T_U|\mathrm{tr}(\boldsymbol{R}_{f_m})}}\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{H}_J^H\mathbf{1}\right)\left(\boldsymbol{w}_m\right.\right.$$
$$\left.\left.\sqrt{\frac{{}^2\;E}{|U_T||T_U|\mathrm{tr}(\boldsymbol{R}_{f_m})}}\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{H}_J^H\mathbf{1}\right)^H\boldsymbol{H}_J^H\right]\boldsymbol{C}_J^H\mathbf{1}=\frac{c_{m_{22}}}{\sqrt{\boldsymbol{R}_{f_m}}} \tag{55}$$

$$\frac{{}^2\;E}{|U_T||T_U|\mathrm{tr}\,(\boldsymbol{R}_{f_m})}\mathbf{1}^T\boldsymbol{C}_J\mathrm{E}\left[\boldsymbol{N}_J\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{N}_T\mathbf{1}\mathbf{1}^T\boldsymbol{N}_T^T\boldsymbol{C}_T^T\boldsymbol{C}_{e_m}^T\boldsymbol{N}^H\right]$$
$$\boldsymbol{C}_J^H\mathbf{1}=\frac{c_{m_{23}}}{\sqrt{\boldsymbol{R}_{f_m}}} \tag{56}$$

$$\mathbf{1}^T\boldsymbol{C}_J\mathrm{E}\left[\boldsymbol{N}_J\left(\boldsymbol{w}_m\quad\sqrt{\frac{{}^2\;E}{|U_T||T_U|\mathrm{tr}(\boldsymbol{R}_{f_m})}}\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{N}\quad\mathbf{1}\right)\left(\boldsymbol{w}_m\right.\right.$$
$$\left.\left.\sqrt{\frac{{}^2\;E}{|U_T||T_U|\mathrm{tr}(\boldsymbol{R}_{f_m})}}\boldsymbol{C}_{e_m}\boldsymbol{C}_T\boldsymbol{N}\quad\mathbf{1}\right)^H\boldsymbol{N}_J^H\right]\boldsymbol{C}_J^H\mathbf{1}=\frac{c_{m_{24}}}{\sqrt{\boldsymbol{R}_{f_m}}}; \tag{57}$$

where $\{c_{m_{21}}, c_{m_{22}}, \ldots, c_{m_{24}}\}$ are as defined in (29a)-(29f).

REFERENCES

[1] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Jan. 2013.

[2] H. Akhlaghpasand, E. Björnson, and S. M. Razavizadeh, "Jamming-robust uplink transmission for spatially correlated massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3495–3504, Mar. 2020.

[3] M. Alageli, A. Ikhlef, and J. Chambers, "SWIPT massive MIMO systems with active eavesdropping," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 233–247, Sep. 2018.

[4] 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.

[5] M. Ranjbar Nikkhah, J. Wu, H. Luyen, and N. Behdad, "A Concurrently Dual-Polarized, Simultaneous Transmit and Receive (STAR) Antenna," *IEEE Trans. Antennas Propag.*, vol. 68, no. 8, pp. 5935–5944, Aug. 2020.

[6] M. V. Kuznetcov, S. K. Podilchak, A. J. McDermott, and M. Sellathurai, "Dual-Polarized High-Isolation Antenna Design and Beam Steering Array Enabling Full-Duplex Communications for Operation Over a Wide Frequency Range," *IEEE Open Journal of Antennas Propag.*, vol. 2, pp. 521–532, 2021.

[7] Q.-C. Ye, Y.-M. Zhang, J.-L. Li, G. F. Pedersen, and S. Zhang, "High-Isolation Dual-Polarized Leaky-Wave Antenna With Fixed Beam for Full-Duplex Millimeter-Wave Applications," *IEEE Trans. Antennas Propag.*, vol. 69, no. 11, pp. 7202–7212, Nov. 2021.

[8] M. N. A. Tarek, M. R. Guerra, A. Nunez, M. N. Uddin, and E. A. Alwan, "Improving Isolation in Monostatic Simultaneous Transmit and Receive Systems Using a Quasi-Symmetrical Self-Interference Cancellation Architecture," *IEEE Journal of Microwaves*, pp. 1–10, Dec. 2022.

[9] E. I. Ackerman, C. H. Cox, H. V. Roussell, and P. S. Devgan, "Broadband simultaneous transmit and receive from a single antenna using improved photonic architecture," in *Proc. Int. Microw. Symp. (IMS). IEEE*, Jun. 2019, pp. 778–781.

[10] K. Kolodziej, J. Doane, and B. Perry, "Single antenna in-band full-duplex isolation-improvement techniques," in *Proc. Int. Symp. on Antennas and Propag. (APSURSI). IEEE*, Jun. 2016, pp. 1661–1662.

[11] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. ACM SIGCOMM Conf. on SIGCOMM*, Aug. 2013, pp. 375–386.

[12] A. Hamza, A. Nagulu, A. F. Davidson, J. Tao, C. Hill, H. AlShammary, H. Krishnaswamy, and J. Buckwalter, "A Code-Domain, In-Band, Full-Duplex Wireless Communication Link With Greater Than 100-dB Rejection," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 1, pp. 955–968, Nov. 2021.

[13] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Nov. 2012.

[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] M. Karlsson and E. G. Larsson, "Massive MIMO as a cyber-weapon," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput. IEEE*, Nov. 2014, pp. 661–665.

[16] X. Tang, P. Ren, Y. Wang, and Z. Han, "Combating full-duplex active eavesdropper: A hierarchical game perspective," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1379–1395, Dec. 2016.

[17] J. Moon, H. Lee, C. Song, S. Kang, and I. Lee, "Relay-assisted proactive eavesdropping with cooperative jamming and spoofing," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6958–6971, Aug. 2018.

[18] J. Choi, "Full-Duplexing Jamming Attack for Active Eavesdropping," in *2016 6th International Conference on IT Convergence and Security (ICITCS)*, 2016, pp. 1–5.

[19] L. Li, A. P. Petropulu, and Z. Chen, "MIMO Secret Communications Against an Active Eavesdropper," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 10, pp. 2387–2401, Oct. 2017.

[20] Z. Shen and K. Xu, "Location-Aided Secure Transmission for Uplink Massive MIMO System Against Full-Duplex Jammer," in *Proc. IEEE Conf. on Computer Commun. (INFOCOM). IEEE*, Jul. 2020, pp. 526–531.

[21] C. Si, H. Sun, M. Sheng, X. Wang, and J. Li, "Physical layer security with hostile jammers and eavesdroppers: Security transmission capacity," *Proc. 27th Annu. Int. Symp. Pres. Mobile Radio Commun. (PTMRC)*, pp. 1–6, Sep. 2016.

[22] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. PTR Prentice Hall New Jersey, 1996, vol. 2.

[23] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.

[24] A. K. Jagannatham, "NOC: Estimation for wireless communications: MIMO/OFDM cellular and sensor networks," 2016, [Online] Available: http://nptel.ac.in/courses/117104118/ [Accessed: 2 Apr. 2022].

[25] S. M. Kay, *Fundamentals of statistical signal processing, volume I: Estimation theory*. Englewood Cliffs, PTR Prentice-Hall, 1993.

[26] X. Wu, Y. Shen, and Y. Tang, "Propagation characteristics of the full-duplex self-interference channel for the indoor environment at 2.6 GHz," in *Proc. Antennas and Propag. Soc. Int. Symp. (APSURSI). IEEE*, 2014, pp. 1183–1184.

[27] S. N. Venkatasubramanian, L. Laughlin, K. Haneda, and M. A. Beach, "Wideband self-interference channel modelling for an on-frequency repeater," in *Proc. 10th Eur. Conf. on Antennas and Propag. (EuCAP). IEEE*, Apr. 2016, pp. 1–5.
[28] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Aug. 2012.
[29] A. Mukherjee and A. L. Swindlehurst, "A full-duplex active eavesdropper in MIMO wiretap channels: Construction and countermeasures," in *Proc. Conf. Signals, Syst. Comput. (ASILOMAR). IEEE*, Nov. 2011, pp. 265–269.
[30] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
[31] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2059–2078, Sep. 2000.
[32] M. Alageli, A. Ikhlef, F. Alsifiany, M. A. Abdullah, G. Chen, and J. Chambers, "Optimal downlink transmission for cell-free SWIPT massive MIMO systems with active eavesdropping," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1983–1998, Nov. 2019.
[33] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
[34] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Magazine*, vol. 54, no. 2, pp. 114–123, Feb. 2016.
[35] H. Q. Ngo, M. Matthaiou, and E. G. Larsson, "Massive MIMO with optimal power and training duration allocation," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 605–608, Sep. 2014.

Mahmoud Alageli was born in Houn, Libya. He received the B.Sc. degree (Hons.) in electrical and electronic engineering from the Engineering Academy Tajoura, Tripoli, Libya, in 1999, the M.Eng. degree in communication and computer from the National University of Malaysia, Malaysia, in 2006, and the Ph.D. degree in communications and signal processing from Newcastle University, Newcastle upon Tyne, U.K., in 2019. From 2007 to 2011, he was an Assistant Lecturer with the Engineering Academy Tajoura. Since 2012, he has been a Lecturer with the Faculty of Engineering, Garaboulli, Elmergib University, Libya. His current research interests include energy harvesting communications, physical layer security, and massive MIMO. He was a recipient of the Bright Bestowal 2006 organized by the Center of Graduate Studies at the National University of Malaysia.

Aissa Ikhlef (Senior Member, IEEE) received the Ph.D. degree in Signal Processing and Telecommunications from the University of Rennes 1, France, in 2008. From 2008 to 2010 he was a Postdoctoral Fellow with the Communication and Remote Sensing Laboratory, Catholic University of Louvain, Belgium. From 2010 to 2013 he was with the data communications group, University of British Columbia, Vancouver, Canada, as a Postdoctoral Fellow. From 2013 to 2014 he was with Toshiba Research Europe Limited, UK, as a Senior Research Engineer. From 2014 to 2016 he was with the School of Electrical and Electronic Engineering, Newcastle University, UK, as a Lecturer (Assistant Professor). Since 2016 he has been with the Department of Engineering, Durham University, Durham, UK, where he is currently an Associate Professor. His current research interests include machine learning, physical layer security, cloud radio access networks, and massive MIMO.

Jonathon Chambers (S'83–M'90–SM'98–F'11) received the Ph.D. and D.Sc. degrees in signal processing from the Imperial College of Science, Technology and Medicine (Imperial College London), London, U.K., in 1990 and 2014, respectively. In 2015, he joined the School of Electrical and Electronic Engineering, Newcastle University, where he was Professor of signal and information processing and led the ComS2IP group and is now a Visiting Professor. He has advised approaching 100 researchers through to Ph.D. graduation and published more than 600 conference papers and journal articles, many of which are in IEEE journals. Dr. Chambers is a Fellow of the Royal Academy of Engineering, U.K., and the Institution of Electrical Engineers. He was Technical Program Cochair for the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, Prague, Czech Republic. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING for three terms over the periods 1997-1999, 2004- 2007, and as a Senior Area Editor between 2011-2014.