Maximum Likelihood Detection in Single-Input Double-Output Non-Gaussian Barrage-Jammed Systems

Khalid A. Almahorg and Ramy H. Gohary

***Abstract*—We derive the likelihood functions and the maximum likelihood (ML) detectors for four classes of single-input double-output (SIDO) communication systems, i.e., systems with one transmit and two receive antennas. For all classes, the received signals are contaminated by a Gaussian noise component and a non-Gaussian component induced by the Gaussian transmissions of a proactive continuous single-antenna jammer over an unknown complex 2×1 Gaussian vector channel. The considered classes correspond to whether full channel distribution information (CDI), or partial CDI about the transmitter channel and the jammer channel is available at the receiver. Unlike their scalar counterparts, the vector channels considered herein interweave the components of the received signal, rendering the derivation of the likelihood function a daunting task for more than two receive antennas. Furthermore, the interweaving of the received signal components in the vector channel case prevents the optimal ML detector for unit-norm constellations from reducing to the corresponding Gaussian approximation-based detector. This is in sharp contrast with the scalar case, wherein the two detectors are equivalent for unit-norm constellations. Confirming our analytical findings, experimental results show that the difference between the two detectors can be significant, especially when the transmitter-receiver and jammer-receiver channels have substantial line-of-sight components. Although the computational cost of performing optimal ML detection in the presence of non-Gaussian jamming is higher in the case of two receive antennas, the performance advantage over the single antenna case justifies it.**

I. Introduction

The services offered by wireless communications have proliferated our everyday life, from entertainment to banking and defence-centric services. Disrupting these services can cause distress, evoke chaos and potentially incur catastrophic losses. The open nature of the wireless channel makes it vulnerable to jamming which compromises the reliability and effectiveness of these services and, in extreme cases, can result in denial of service altogether [1], [2]. Although traditionally, jamming was considered in the context of warfare, its impact on civil applications cannot be ignored. For instance, cellular systems [3], positioning systems [4], sensing and monitoring networks [5], Vehicular Ad hoc Networks (VANETs) [6], cognitive radio networks [7], and Internet-of-Things (IoT) systems [8] to name a few can all be disrupted by jamming activities. Unlike their traditional stationary counterparts, modern jammers can be mobile or airborne [9], [10], giving rise to intricate wireless channels and received signal models.

The authors are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada.

Launching jamming attacks is readily feasible and inexpensive. For instance, a protocol-specific jammer can use off-the-shelf software defined radio to exploit vulnerabilities of system protocols [11], e.g., the medium access control protocol (MAC) [12] and channel estimation procedures [13]. Alternatively, a protocol-neutral jammer can use a noise generator to disrupt communications by degrading the signal-to-noise ratio (SNR) at the receiver. Such a jammer can simultaneously impact the performance of multiple wireless systems [5].

Jammers can be either reactive or proactive depending on the timing of their attacks [12]. Whereas reactive jammers launch their attacks whenever a communication activity is detected, proactive ones launch their attacks blindly, irrespective of the communication activity. The transmissions of proactive jammers can be either continuous or random. Random jamming is more power efficient, but continuous jamming is more impactful [2].

Jammers can also be classified according to the bandwidth they target. For instance, a jammer can focus its transmissions on a single or a particular set of discrete tones or on a contiguous frequency band. A particular class of jammers is the so-called barrage jammers which are characterized by broadband transmissions [14]. Such jammers typically emit white Gaussian noise that occupies the entire bandwidth of the target system. In the absence of prior information about target systems, proactive continuous barrage jamming is known to be the most deleterious [9]. When prior channel information is available, proactive jamming can be mitigated with the help of a friendly eavesdropper [15] or by adjusting the transmission power or rate, or both, of legitimate users using adaptive water-filling-like protocols [16]–[18]. However, these techniques are not readily applicable when channel information is not available, which is the case considered herein. This is due to the fact that when channel information is not available, the signal induced by the jammer at the receiver is not Gaussian, thereby violating a fundamental assumption of those techniques.

The signal induced at the receiver by the barrage jammer is assumed to be an additive Gaussian noise in the majority of literature, see e.g., [9], [15], [16], [19]–[21]. However, this assumption is valid if the jammer transmits Gaussian-distributed signals and the channel state information (CSI) of the jammer-receiver channel is known. The fact that the jammer does not send pilots or prescribed waveforms makes acquiring this CSI a difficult task. The problem becomes more pronounced when either the jammer and/or the receiver are

mobile or airborne [22], rendering the Gaussianity assumption on the jamming signal observed at the receiver questionable. For a mobile jammer and/or receiver, the statistical distribution of the jamming signal observed at the receiver is that of the product of a Gaussian random variable representing the emitted jamming signal and a random vector representing the multiple-antenna jammer-receiver channel. In the particular case in which the receiver has a single antenna and the jammer-receiver channel is Rayleigh or Rician fading, the distribution of the jamming signal at the receiver is that of the product of two Gaussian random variables, which is not Gaussian [23], [24]. Having expressions for the conditional probability density functions (pdf's) of the received signal in the presence of barrage jamming is necessary for optimal maximum likelihood (ML) detection. Such expressions depend on the channel statistical information available at the receiver.

Obtaining conditional pdf expressions for the received signal involves deriving the distributions of various products of Gaussian random variables. Such distributions have been derived for real and complex Gaussian random variables in, e.g., [25]–[29]. For instance, the pdf of the inner product of two real Gaussian random vectors was derived in [25]. The case of the inner product of two complex Gaussian vectors was considered in [26]. Therein, an expression of the characteristic function was first derived, and subsequently, the corresponding pdf was obtained by evaluating its inverse Fourier transform numerically. The joint pdf of the magnitude and phase of the product of two independent complex Gaussian scalars was derived in [27], whereas that of the random scalar resulting from adding a complex Gaussian scalar to the product of two complex Gaussian scalars was derived in [30]. The vector counterpart of [30], i.e., the pdf of the random vector resulting from adding a Gaussian vector to the product of a Gaussian scalar and a Gaussian vector was derived in [28], [31]. A more general version in which the pdf of the vector resulting form adding a Gaussian vector to the product of a Gaussian matrix and a Gaussian vector was derived in [32] and [33]. The vector considered in [31] resulted due to considering a barrage-jammed single-input single-output system. In that case, all channels were assumed to be block fading with independent Gaussian scalar gains, rendering the components of the received signal vector separable. This separability has been key in deriving the conditional pdf of the received signal with arbitrary symbol dimensions. Unfortunately, this separability does not hold for the case in which the receiver has multiple antennas. This significantly complicates the derivation of the conditional pdf of the received signal vector and prevents the derivation in [31] from generalizing to the case of receivers with more than two antennas.

In this paper, we consider a scenario in which a single-antenna transmitter sends multi-dimensional complex symbols to a double-antenna receiver in the presence of a single-antenna barrage jammer. The transmitter-receiver and the jammer-receiver channels are assumed to be block fading independent Gaussian matrices with coherence time that is larger than the time required for transmitting the multi-dimensional symbol. The jammer's signal is represented by a vector of independent complex Gaussian random variables. Given the transmitted symbols, the received signal can be represented by the random vector resulting from adding a Gaussian vector to the product of a block diagonal Gaussian matrix and a Gaussian vector. We derive the conditional pdf of the received signal vector for the cases in which either full or partial channel distribution information (CDI) about the transmitter-receiver and jammer-receiver channels is available at the receiver. The block fading assumption along with the multiplication inherent in this model interweaves the components of the received signal, thereby complicating the derivation of the required conditional pdf and rendering it intractable for receivers with more than two antennas. To the best of our knowledge, the scenario considered herein was not considered in the literature and, in a sense, complements the results reported in [27], [30], and [31].

The paper is organized as follows. In Section II we present the system model and the problem statement for four different cases. Expressions for the likelihood functions corresponding to each of these cases are presented in Section III. This section also presents expressions for the likelihood function obtained by treating the jammer's component of the received signal as if it were Gaussian distributed. In Section IV we present simulation results and Section V concludes the paper. For ease of flow, most derivations are relegated to the appendices.

Standard notations will be used throughout the paper. Boldface uppercase and lowercase letters will be used to denote random vectors and random scalars, respectively; regular face letters will be used to denote deterministic quantities. The Euclidean norm of vector V is denoted by $\|V\|$. The determinant of a matrix M, its inverse, if exists, and its Hermitian transpose will be denoted by $|M|$, M^{-1}, and $M^{\dagger}$ respectively. The identity matrix of dimension N will be denoted by I_N. The vectorized version of matrix $M \in \mathbb{C}^{m\times n}$ resulting from stacking the columns of M on top of each other is given by $\text{vec}(M) \in \mathbb{C}^{mn\times 1}$. For a complex variable z, the complex conjugate, the magnitude, the real part, the imaginary part and the phase will be denoted by z^*, $|z|$, $\Re\{z\}$, $\Im\{z\}$ and $\measuredangle z$, respectively. The modified Bessel functions of the first kind of order zero will be denoted by $I_0(\cdot)$. The Kronecker product will be denoted by $\otimes$.

II. System Model and Problem Statement

We consider a system composed of a single-antenna transmitter communicating with a double-antenna receiver in the presence of a single-antenna barrage jammer. The transmitter-receiver channel, $\boldsymbol{H} \in \mathbb{C}^{2\times 1}$, and the jammer-receiver channel, $\boldsymbol{G} \in \mathbb{C}^{2\times 1}$, are assumed to be block fading vector channels, i.e., they remain constant during the transmission of any data symbol and take independent realizations for subsequent symbols. Unlike their scalar counterparts, the block fading vector channels considered herein interweave the components of the received signal, rendering the derivation of the likelihood function intractable for more than two receive antennas.

Both $\boldsymbol{H}$ and $\boldsymbol{G}$ are modeled as independent proper complex Gaussian random vectors with mean vectors $\boldsymbol{\mu}_H$ and $\boldsymbol{\mu}_G$, and covariance matrices $\sigma_h^2 I_2$ and $\sigma_g^2 I_2$, respectively, where

σ_h and σ_g are known scalars, i.e., $\boldsymbol{H} \sim \mathcal{CN}(\boldsymbol{\mu}_H, \sigma_h^2 I_2)$ and $\boldsymbol{G} \sim \mathcal{CN}(\boldsymbol{\mu}_G, \sigma_g^2 I_2)$. The transmitter-receiver and the jammer-receiver distances are assumed to be much larger than the receiver's inter-antenna spacing, this assumption allows us to express the mean vectors of $\boldsymbol{H}$ and $\boldsymbol{G}$ as $\boldsymbol{\mu}_H = |\boldsymbol{\mu}_h|[e^{-\jmath\boldsymbol{\theta}_{h_1}}\ e^{-\jmath\boldsymbol{\theta}_{h_2}}]^\dagger$ and $\boldsymbol{\mu}_G = |\boldsymbol{\mu}_g|[e^{-\jmath\boldsymbol{\theta}_{g_1}}\ e^{-\jmath\boldsymbol{\theta}_{g_2}}]^\dagger$, where $|\boldsymbol{\mu}_h|$ are $|\boldsymbol{\mu}_g|$ are known scalars, whereas $\boldsymbol{\theta}_{h_1}, \boldsymbol{\theta}_{h_2}, \boldsymbol{\theta}_{g_1}$, and $\boldsymbol{\theta}_{g_2}$ are $i.i.d$ random scalars uniformly distributed in $[0, 2\pi)$. The transmitter, the jammer, and the receiver can be mobile or airborne, which is the case in vehicle-to-vehicle (V2V) and unmanned aerial vehicle (UAVs) communications. Hence, both channels, $\boldsymbol{H}$ and $\boldsymbol{G}$, are assumed to be subject to independent fading processes. Both channels can be Rayleigh or Rician fading depending on whether their respective means, $\boldsymbol{\mu}_H$ and $\boldsymbol{\mu}_G$, are zero or non-zero vectors [34]. For the case in which the transmitter-receiver channel, $\boldsymbol{H}$, is Rayleigh fading, i.e., $|\boldsymbol{\mu}_h| = 0$, we assume that the receiver has full CDI about $\boldsymbol{H}$. On the other hand, for the case in which $\boldsymbol{H}$ is Rician fading, e.g., the case when either the transmitter and/or the receiver is a UAV [35], we assume that the receiver has a partial CDI about $\boldsymbol{H}$ represented by the relative strength of its line-of-sight (LOS) components given by its k-factor, $k_h = \dfrac{|\boldsymbol{\mu}_h|}{\sigma_h}$. Analogously, $\boldsymbol{G}$ is either Rayleigh fading and the receiver knows σ_g, or $\boldsymbol{G}$ is Rician fading and the receiver knows σ_g and $k_g = \dfrac{|\boldsymbol{\mu}_g|}{\sigma_g}$.

The transmitted symbol, the jamming signal, and the additive noise at the receiver are denoted by $\boldsymbol{X} \in \mathbb{C}^T$, $\boldsymbol{V} \in \mathbb{C}^T$, and $\bar{\boldsymbol{Z}} \in \mathbb{C}^{2\times T}$, respectively. The jamming vector, $\boldsymbol{V}$, and the vectorized additive white noise, $\boldsymbol{Z} = \text{vec}(\bar{\boldsymbol{Z}})$, are assumed to be independent circularly symmetric complex Gaussian random vectors, i.e., $\boldsymbol{V} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_V^2 I_T)$ and $\boldsymbol{Z} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_Z^2 I_{2T})$, where σ_V and σ_Z are known scalars. The received signal is denoted by $\bar{\boldsymbol{Y}} \in \mathbb{C}^{2\times T}$ and is given by:

$$\bar{\boldsymbol{Y}} = \boldsymbol{H}(\boldsymbol{X}^*)^\dagger + \boldsymbol{G}(\boldsymbol{V}^*)^\dagger + \bar{\boldsymbol{Z}}. \tag{1}$$

It can be shown that the vectorized version of (1) is given by:

$$\boldsymbol{Y} = (I_T \otimes \boldsymbol{H})\boldsymbol{X} + (I_T \otimes \boldsymbol{G})\boldsymbol{V} + \boldsymbol{Z}, \tag{2}$$

where $\boldsymbol{Y} \in \mathbb{C}^{2T}$.

This paper aims at developing the ML detector for each of the four cases outlined hereinabove. For all these cases, we show that the ML detector outperforms the detector based on the Gaussian approximation of the signal induced at the receiver by the jammer transmission. This contrasts with the case in which the receiver has a single antenna. In that case, the exact detector and the Gaussian approximation detector coincide when the constellation symbols have unit norm.

The four cases considered herein are the following.

1) *Statistical distribution of both $\boldsymbol{H}$ and $\boldsymbol{G}$ partially known:* The ML detector decides in favor of the constellation point $\hat{X}$ given by

$$\hat{X} = \arg\max_{X\in\mathcal{C}} p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X; |\boldsymbol{\mu}_h|, \sigma_h^2, |\boldsymbol{\mu}_g|, \sigma_g^2\right). \tag{3}$$

In this case, the phases of the mean vectors of $\boldsymbol{H}$ and $\boldsymbol{G}$, i.e., $\boldsymbol{\theta}_{\mu_{h_1}}$, $\boldsymbol{\theta}_{\mu_{h_2}}$, $\boldsymbol{\theta}_{\mu_{g_1}}$, and $\boldsymbol{\theta}_{\mu_{g_2}}$ are random and unknown. This situation arises when both the transmitter and the jammer channels have LOS components. The exact and Gaussian approximation-based likelihood functions for this case are given in (7) and (9), respectively.

2) *Statistical distribution of $\boldsymbol{H}$ partially known and statistical distribution of $\boldsymbol{G}$ known:*
The ML detector decides in favor of the constellation point $\hat{X}$ given by

$$\hat{X} = \arg\max_{X\in\mathcal{C}} p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X; |\boldsymbol{\mu}_h|, \sigma_h^2, \sigma_g^2\right). \tag{4}$$

In this case the phases of the mean vector of $\boldsymbol{H}$, i.e., $\boldsymbol{\theta}_{\mu_{h_1}}$ and $\boldsymbol{\theta}_{\mu_{h_2}}$, are random and unknown while the mean vector of $\boldsymbol{G}$ is zero. This situation arises when the transmitter channel has an LOS component but the jammer channel does not. The exact and Gaussian approximation-based likelihood functions for this case are given in (10) and (9), respectively.

3) *Statistical distribution of $\boldsymbol{H}$ known and statistical distribution of $\boldsymbol{G}$ partially known:* The ML detector decides in favor of the constellation point $\hat{X}$ given by

$$\hat{X} = \arg\max_{X\in\mathcal{C}} p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X; |\boldsymbol{\mu}_g|, \sigma_g^2, \sigma_h^2\right). \tag{5}$$

Complementary to the previous case, the phases of the mean vector of $\boldsymbol{G}$, i.e., $\boldsymbol{\theta}_{\mu_{g_1}}$ and $\boldsymbol{\theta}_{\mu_{g_2}}$, are random and unknown while the mean vector of $\boldsymbol{H}$ is zero. In this situation the jammer channel has an LOS component but the transmitter channel does not. The exact and Gaussian approximation-based likelihood functions for this case are given in (12) and (14), respectively.

4) *Statistical distribution of both $\boldsymbol{H}$ and $\boldsymbol{G}$ known:* The ML detector decides in favor of the constellation point $\hat{X}$ given by

$$\hat{X} = \arg\max_{X\in\mathcal{C}} p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X; \sigma_h^2, \sigma_g^2\right). \tag{6}$$

In this case both $\boldsymbol{H}$ and $\boldsymbol{G}$ have zero mean vectors, i.e., neither the transmitter channel nor the jammer channel has an LOS component. The exact and Gaussian approximation-based likelihood function for this case are given in (15) and (14), respectively.

In all of the aforementioned cases, the transmitted symbol that maximizes the corresponding likelihood function is obtained using exhaustive search over constellation symbols. In the next section, we will provide exact and approximate ML detection rules for each of these cases.

III. Likelihood function expressions for the different CSI cases

In this section, we present the exact likelihood function expressions for the cases outlined in the previous section along with the likelihood expressions based on the Gaussian approximation of the non-Gaussian signal resulting from the jammer's transmissions.

A. Statistical distribution of both $\boldsymbol{H}$ and $\boldsymbol{G}$ partially known:

In this case the transmitter channel and the jammer channel have LOS components. This might be the case when the jammer and the transmitter are airborne, or when the receiver is airborne or highly mounted.

1) The exact likelihood function: The exact likelihood function for this case is $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X;|\boldsymbol{\mu}_h|,\sigma_h^2,|\boldsymbol{\mu}_g|,\sigma_g^2\right)$, (cf. (3)) and is given in the following theorem:

Theorem 1. *Let* $\boldsymbol{Y} = (I_T \otimes \boldsymbol{H})\boldsymbol{X} + (I_T \otimes \boldsymbol{G})\boldsymbol{V} + \boldsymbol{Z}$, *where* $\boldsymbol{X} \in \mathbb{C}^T$, $\boldsymbol{V} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_V^2 I_T)$, $\boldsymbol{Z} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_Z^2 I_{2T})$, $\boldsymbol{H} \sim \mathcal{CN}\left(|\boldsymbol{\mu}_h|[e^{-\jmath\boldsymbol{\theta}_{\mu_{h_1}}}\ e^{-\jmath\boldsymbol{\theta}_{\mu_{h_2}}}]^\dagger, \sigma_h^2 I_2\right)$ *and* $\boldsymbol{G} \sim \mathcal{CN}\left(|\boldsymbol{\mu}_g|[e^{-\jmath\boldsymbol{\theta}_{\mu_{g_1}}}\ e^{-\jmath\boldsymbol{\theta}_{\mu_{g_2}}}]^\dagger, \sigma_g^2 I_2\right)$. *The scalars* $|\boldsymbol{\mu}_h|$, $|\boldsymbol{\mu}_g|$, σ_h, σ_g, σ_Z, *and* σ_V *are known, and the phases* $\boldsymbol{\theta}_{\mu_{h_1}}$, $\boldsymbol{\theta}_{\mu_{h_2}}$, $\boldsymbol{\theta}_{\mu_{g_1}}$, *and* $\boldsymbol{\theta}_{\mu_{g_2}}$ *are i.i.d. random scalars uniformly distributed over* $[0, 2\pi)$. *Let* $k_h = \dfrac{|\boldsymbol{\mu}_h|}{\sigma_h}$ *and* $k_g = \dfrac{|\boldsymbol{\mu}_g|}{\sigma_g}$. *The likelihood function of* $\boldsymbol{Y}$ *conditioned on* $\boldsymbol{X} = X$ *is given by:*

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{\exp(-2k_g^2)}{4\pi^{2T+2}\sigma_g^4\sigma_v^4|\Sigma|^2}\times$$
$$\int_0^{2\pi}\int_0^{2\pi}\exp\Big(-(Y_o-k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)^\dagger\Sigma^{-1}(Y_o-k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)\Big)$$
$$\times\exp\Big(-(Y_e-k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)^\dagger\ \Sigma^{-1}(Y_e-k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)\Big)$$
$$\times\Psi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}})\,d\theta_{\mu_{h_1}}d\theta_{\mu_{h_2}}, \quad (7)$$

where $\Sigma = \sigma_h^2 XX^\dagger + \sigma_z^2 I_T$, *and*

$$\Psi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}}) = \int_0^1\int_0^1 \frac{t\ (\ \bar{t}\ )^{T-3}}{|\Lambda t+\bar{t}\ I_T|}\exp\left(\frac{-t}{\sigma_g^2\sigma_v^2\bar{t}}\right)$$
$$\times\exp\Big(\bar{u}(Y_o-k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)^\dagger\Upsilon(\Sigma,t)(Y_o-k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)\Big)$$
$$\times\exp\Big(u(Y_e-k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)^\dagger\Upsilon(\Sigma,t)(Y_e-k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)\Big)$$
$$\times I_0\Big(2\sqrt{u\bar{u}}t|(Y_o-k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)^\dagger\Upsilon(\Sigma,t)(Y_e-k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)|\Big)$$
$$\times I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{\bar{u}\ t}{\bar{t}}}\right)I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{u\ t}{\bar{t}}}\right)du\ dt, \quad (8)$$

where Y_o *contains the odd components of* Y, *i.e.,* $Y_o = [Y_1^*,\ldots,Y_{2i-1}^*,\ldots,Y_{2T-1}^*]^\dagger$ *and* Y_e *contains the even components of* Y, *i.e.,* $Y_e = [Y_2^*,\ldots,Y_{2i}^*,\ldots,Y_{2T}^*]^\dagger$. *The parameters* $\bar{t} = 1-t$ *and* $\bar{u} = 1-u$, *and the diagonal matrix* Λ *contains the eigenvalues of* Σ^{-1}, *and the matrix* $\Upsilon(\Sigma,t) = \Sigma^{-1}\left(I_T + \frac{\bar{t}}{t}\Sigma\right)^{-1}$.

Proof. See Appendix A. □

We note that $\boldsymbol{Y}_o$ and $\boldsymbol{Y}_e$ are the outputs of the first and second receive antennas, respectively. We also note the symmetry of the expressions in (7) and (8) in $\boldsymbol{Y}_o$ and $\boldsymbol{Y}_e$, which reflects the symmetry of the channels observed by the two antennas. The computation of (8) is compounded by the fact that a term of the integrand in (8), namely, $I_0\Big(2\sqrt{u\bar{u}}\ t|(Y_o-k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)^\dagger\Upsilon(\Sigma,t)(Y_e-k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)|\Big)$, contains the inner product of the output of the first and second receive antennas. This portends the difficulties that are likely to be encountered if the proposed technique were to be used to obtain the likelihood function for the case with more than two receive antennas.

2) Gaussian Approximation:

Due to mathematical tractability, the signal induced by the jammer at the receiver has been commonly assumed to be Gaussian distributed, see, e.g., [9], [15], [16], [19]–[21]. In this paper, we investigate this assumption for all considered cases.

Approximating the signal resulting form the jammer transmission at the i-th receive antenna, $i \in \{1,2\}$, by a Gaussian vector $\boldsymbol{J}_i = \boldsymbol{g}_i\boldsymbol{V}$, where $\boldsymbol{J}_i \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_g^2\sigma_v^2 I_T\right)$ yields the likelihood function in the following lemma.

Lemma 1. *Let* $\boldsymbol{Y}$, $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{Z}$, $\boldsymbol{H}$, $\boldsymbol{G}$, k_h *and* k_g *be as defined in Theorem 1, and let* $\boldsymbol{J} = \boldsymbol{g}_i\boldsymbol{V}$ *where* $i \in \{1,2\}$. *Assume that* $\boldsymbol{J} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_g^2\sigma_V^2 I_T)$, *then the likelihood function of* $\boldsymbol{Y}$ *conditioned on* $\boldsymbol{X} = X$ *is given by:*

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{1}{\pi^{2T}|\Sigma_N|^2}$$
$$\times\exp\left(-Y_o^\dagger\Sigma_N^{-1}Y_o - Y_e^\dagger\Sigma_N^{-1}Y_e - 2k_h^2\sigma_h^2X^\dagger\Sigma_N^{-1}X\right)$$
$$\times I_0\left(2k_h\sigma_h|Y_o^\dagger\ \Sigma_N^{-1}X|\right)I_0\left(2k_h\sigma_h|Y_e^\dagger\ \Sigma_N^{-1}X|\right), \quad (9)$$

where $\Sigma_N = \sigma_h^2XX^\dagger + (\sigma_g^2\sigma_v^2 + \sigma_Z^2)I_T$, *and* Y_o *and* Y_e *are defined in Theorem 1.*

Proof. See Appendix B. □

It can be noticed from (9) that in this case, and in the other cases as will be seen later, the likelihood function based on the Gaussian approximation does not incorporate the k-factor of the jammer's channel, k_g. Not incorporating this information renders the performance of the detector based on this likelihood function inferior to the one based on the exact one, cf. Section IV below. It is worth mentioning that the non-zero mean of the transmitter's channel, μ_h, appears in the Gaussian approximation expression in the form of $|\mu_h| = k_h\sigma_h$. In contrast, the non-zero mean of the jammer's channel, μ_g, does not appear in the likelihood expression based on the Gaussian approximation. This is due to the fact that the signal resulting form the jammer's transmission, $\boldsymbol{J} = \boldsymbol{gV}$, is approximated by a Gaussian signal with mean $\mu_j = \mu_g\mu_v$. The zero mean of the jamming signal, μ_v, renders μ_j zero. Hence, μ_g is suppressed regardless of its value.

B. Statistical distribution of $\boldsymbol{H}$ *partially known and statistical distribution of* $\boldsymbol{G}$ *known:*

The conditional statistical distribution of the received signal for this case is required for optimal detection when the jammer-receiver channel does not have an LOS component, whereas the transmitter-receiver channel has an LOS component with a known k-factor. Such a scenario may arise when both the jammer and the receiver are ground-based and the transmitter is airborne or highly mounted.

1) The exact likelihood function: The exact likelihood function for this case is $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X;|\boldsymbol{\mu}_h|,\sigma_h^2,\sigma_g^2\right)$, (cf. (4)) and is given in the following corollary.

Corollary 1. *Let* $\boldsymbol{Y}$, $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{Z}$, $\boldsymbol{H}$, Σ, $\boldsymbol{\theta}_{\mu_{h_1}}$, $\boldsymbol{\theta}_{\mu_{h_2}}$ *and* k_h *be as defined in Theorem 1, and let* $\boldsymbol{G} \sim \mathcal{CN}\left(\boldsymbol{0}, \sigma_g^2 I_2\right)$ *where* σ_g

is a known scalar. The likelihood function of $\boldsymbol{Y}$ conditioned on $\boldsymbol{X} = X$ is given by:

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{1}{4\pi^{2T+2}\sigma_g^4\sigma_v^4|\Sigma|^2}\times \int_0^{2\pi}\int_0^{2\pi} \exp\Big(-(Y_o - k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)^\dagger \Sigma^{-1}(Y_o - k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)\Big) \times \exp\Big(-(Y_e - k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)^\dagger\ \Sigma^{-1}(Y_e - k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)\Big) \times \Phi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}})\ d\theta_{\mu_{h_1}}d\theta_{\mu_{h_2}}, \quad (10)$$

where

$$\Phi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}}) = \int_0^1\int_0^1 \frac{t\ (\ \bar{t}\ )^{T-3}}{|\Lambda t + \bar{t}\ I_T|}\exp\left(\frac{-t}{\sigma_g^2\sigma_v^2\bar{t}}\right) \times \exp\Big(\bar{u}(Y_o - k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)^\dagger \Upsilon(\Sigma,t)(Y_o - k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)\Big) \times \exp\Big(u(Y_e - k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)^\dagger \Upsilon(\Sigma,t)(Y_e - k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)\Big) I_0\Big(2\sqrt{u\bar{u}t}|(Y_o - k_h\sigma_h e^{\jmath\theta_{\mu_{h_1}}}X)^\dagger \Upsilon(\Sigma,t)(Y_e - k_h\sigma_h e^{\jmath\theta_{\mu_{h_2}}}X)|\Big) \times du\ dt, \quad (11)$$

where Λ, $\bar{t}$, Y_o, Y_e, $\Upsilon(\Sigma,t)$, and $\bar{u}$ are defined in Theorem 1.

Proof. Setting k_g to zero in (7) and (8) and noticing that $I_0(0) = 1$ yields (10) and (11), respectively. □

It can be noticed that the role of the function $\Phi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}})$ in (10) is similar to the role of the function $\Psi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}})$ in (7). In particular, both functions render the outputs of both antennas inseparable, which compounds the evaluation of the respective likelihood functions. However, in comparison with $\Psi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}})$ in (8), the computation of $\Phi_{X,Y,k_h}(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}})$ in (11) is less demanding because it does not contain the Bessel functions involving k_g.

2) Gaussian Approximation: From Lemma 1, it can be seen that the likelihood function based on the Gaussian approximation does not incorporate the k-factor of the jammer's channel, k_g. Hence, it can be readily verified that approximating the received signal components corresponding to the jammer's transmission by Gaussian vectors as in Section III-A2 yields the same likelihood function given in (9).

C. Statistical distribution of $\boldsymbol{H}$ known and statistical distribution of $\boldsymbol{G}$ partially known:

Complementary to the previous case, in this case the jammer-receiver channel has an LOS component with a known k-factor, whereas the transmitter-receiver channel has no LOS component. Such a case arises when the jammer is airborne or highly mounted whereas both the transmitter and the receiver are ground-based.

1) The exact likelihood function: The exact likelihood function for this case is $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X;|\boldsymbol{\mu}_g|,\sigma_g^2,\sigma_h^2\right)$, (cf. (5)) and is given in the following corollary.

Corollary 2. *Let $\boldsymbol{Y}$, $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{Z}$, $\boldsymbol{G}$, Σ and k_g be as defined in Theorem 1, and let $\boldsymbol{H} \sim \mathcal{CN}\left(\boldsymbol{0},\sigma_h^2 I_2\right)$ where σ_h is a known scalar. The likelihood function of $\boldsymbol{Y}$ conditioned on $\boldsymbol{X} = X$ is given by:*

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{\exp(-2k_g^2)}{\pi^{2T}\sigma_g^4\sigma_v^4|\Sigma|^2} \times \exp\left(-\left(Y_o^\dagger\ \Sigma^{-1}\ Y_o\right) - \left(Y_e^\dagger\ \Sigma^{-1}Y_e\right)\right)\Omega_{X,Y,k_h}, \quad (12)$$

where

$$\Omega_{X,Y,k_h} = \int_0^1\int_0^1 \frac{t\ (\ \bar{t}\ )^{T-3}}{|\Lambda t + \bar{t}\ I_T|}\exp\left(-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right) \times \exp\left(\bar{u}Y_o^\dagger\Upsilon(\Sigma,t)Y_o + uY_e^\dagger\Upsilon(\Sigma,t)Y_e\right) I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{\bar{u}\ t}{\bar{t}}}\right) \times I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{u\ t}{\bar{t}}}\right) I_0\left(2\sqrt{u\bar{u}}t|Y_o^\dagger\ \Upsilon(\Sigma,t)Y_e|\right) dudt, \quad (13)$$

where Λ, $\bar{t}$, Y_o, Y_e, $\Upsilon(\Sigma,t)$, and $\bar{u}$ are defined in Theorem 1.

Proof. Setting k_h to zero in (7) and (8) and noticing that $(2\pi)^{-2}\int_0^{2\pi}\int_0^{2\pi} d\theta_{\mu_{h_1}}d\theta_{\mu_{h_2}} = 1$, and $I_0(0) = 1$ yields (12) and (13), respectively. □

Comparing (12) with (7) and (10), we notice that (12) does not involve integration over $\theta_{\mu_{h_1}}$ and $\theta_{\mu_{h_2}}$. This is due to the fact that $\boldsymbol{H}$ is zero mean, which renders the computation of the expression in (12) much less demanding than its counterparts in (7) and (10).

2) Gaussian Approximation: Approximating the signal resulting form the jammer transmission at the i-th receive antenna, $i \in \{1,2\}$, by a Gaussian vector $\boldsymbol{J}_i = \boldsymbol{g}_i\boldsymbol{V}$, where $\boldsymbol{J}_i \sim \mathcal{CN}\left(\boldsymbol{0},\sigma_g^2\sigma_v^2 I_T\right)$ yields the likelihood function in the following lemma.

Lemma 2. *Let $\boldsymbol{Y}$, $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{Z}$, $\boldsymbol{G}$, Σ and k_g be as defined in Theorem 1, and let $\boldsymbol{H} \sim \mathcal{CN}\left(\boldsymbol{0},\sigma_h^2 I_2\right)$ where σ_h is a known scalar. Let $\boldsymbol{J} = \boldsymbol{g}_i\boldsymbol{V}$ where $i \in \{1,2\}$. Assume that $\boldsymbol{J} \sim \mathcal{CN}(\boldsymbol{0},\sigma_g^2\sigma_V^2 I_T)$, then the likelihood function of $\boldsymbol{Y}$ conditioned on $\boldsymbol{X} = X$ is given by:*

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{1}{\pi^{2T}|\Sigma_N|^2}\exp\left(-\left(Y_o^\dagger\Sigma_N^{-1}Y_o + Y_e^\dagger\Sigma_N^{-1}Y_e\right)\right), \quad (14)$$

where $\Sigma_N = \sigma_h^2 XX^\dagger + (\sigma_g^2\sigma_v^2 + \sigma_Z^2)I_T$, and Y_o and Y_e are defined in Theorem 1.

Proof. Setting k_h to zero in (9) and noticing that $I_0(0) = 1$ yields (14). □

As in Lemma 1, it can be seen form (14) that the likelihood function based on the Gaussian approximation does not incorporate the k-factor of the jammer's channel, k_g. Moreover, in comparison with (9), we notice that the computation of likelihood function in (14) is less demanding because it does not contain the Bessel functions involving k_h.

D. Statistical distribution of both $\boldsymbol{H}$ and $\boldsymbol{G}$ known:

In this case, neither the transmitter-receiver channel nor the jammer-receiver channel has an LOS component. This might be the case when all parties are ground-based.

1) The exact likelihood function: The exact likelihood function for this case is $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = p_{\boldsymbol{Y}|\boldsymbol{X}}\left(Y\middle|X;\sigma_h^2,\sigma_g^2\right)$, (cf. (6)) and is given in the following corollary:

Corollary 3. *Let $\boldsymbol{Y}$, $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{Z}$ and Σ be as defined in Theorem 1, and let $\boldsymbol{H} \sim \mathcal{CN}\left(\boldsymbol{0},\sigma_h^2 I_2\right)$ and $\boldsymbol{G} \sim \mathcal{CN}\left(\boldsymbol{0},\sigma_g^2 I_2\right)$, where σ_h and σ_g are known scalars. Then, the likelihood function of $\boldsymbol{Y}$ conditioned on $\boldsymbol{X} = X$ is given by:*

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{1}{\pi^{2T}\sigma_g^4\sigma_v^4|\Sigma|^2} \times \exp\left(-\left(Y_o^\dagger\ \Sigma^{-1}\ Y_o\right) - \left(Y_e^\dagger\ \Sigma^{-1} Y_e\right)\right)\Gamma_{X,Y,k_h}, \quad (15)$$

where

$$\Gamma_{X,Y,k_h} = \int_0^1\int_0^1 \frac{t\ (\ \bar{t}\ )^{T-3}}{|\Lambda t + \bar{t}\ I_T|} \times \exp\left(\bar{u}\ Y_o^\dagger\ \Upsilon(\Sigma,t)\ Y_o + u\ Y_e^\dagger\ \Upsilon(\Sigma,t)\ Y_e - \frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right) \times I_0\left(2\sqrt{u\bar{u}t}\ |Y_o^\dagger\ \Upsilon(\Sigma,t)\ Y_e|\right) du\ dt, \quad (16)$$

where Λ, $\bar{t}$, Y_o, Y_e, $\Upsilon(\Sigma,t)$ and $\bar{u}$ are defined in Theorem 1.

Proof. Setting k_h and k_g to zeros in (7) and (8) and noticing that $(2\pi)^{-2}\int_0^{2\pi}\int_0^{2\pi} d\theta_{\mu_{h_1}} d\theta_{\mu_{h_2}} = 1$, and $I_0(0) = 1$ yields (15) and (16), respectively. □

The fact that (15) does not involve the double integration over $\theta_{\mu_{h_1}}$ and $\theta_{\mu_{h_2}}$, and that Γ_{X,Y,k_h} in (16) does not include the two Bessel functions involving k_g that were present in (13) makes the exact likelihood function of this case the least computationally demanding among all considered cases.

2) Validation: In validating the correctness of the exact likelihood expression in (15) by comparing it to the histogram produced by a Monte Carlo simulation, we faced the problem that the analytical and the Monte Carlo results are multivariate. For instance, a constellation with $T = 2$ will produce a likelihood function that has eight independent real variables. Plotting such a likelihood function or a histogram is not an easy task. To overcome this obstacle, we set T to be 2, and resorted to deriving the conditional marginal distribution of $f_{\boldsymbol{r_1}|\boldsymbol{X}}(r_1|X)$, where r_1 is the magnitude of the first output of the first antenna, i.e. $r_1 = |\boldsymbol{Y}_1|$. The expression of $f_{\boldsymbol{r_1}|\boldsymbol{X}}(r_1|X)$ is given in the following corollary[1].

Corollary 4. *Let $\boldsymbol{Y}$ be the random vector whose likelihood function, $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$, is given in Corollary 3. Let $T = 2$. Then, the marginal distribution of $\boldsymbol{r}_1 = |\boldsymbol{Y}_1|$ conditioned on $\boldsymbol{X}$ is given by:*

$$f_{\boldsymbol{r_1}|\boldsymbol{X}}(r_1|X) = \frac{2r_1}{\sigma_g^4\sigma_v^4|\Sigma|^2}\sum_{k=0}^{\infty}\sum_{\ell=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{n=0}^{k-2\ell}\frac{(n+\ell)!(k-\ell-n)!^2\ r_1^{2n+2\ell}}{k!\ \ell!^2\ n!\ (k-2\ell-n)!} \times \int_0^1\int_0^1 \frac{(u\bar{u}t^2)^k t(\bar{t})^{-1}}{\prod_{i=1}^2(\lambda_i t+\bar{t})}\exp\left(\frac{-t}{\sigma_g^2\sigma_v^2\bar{t}} - \frac{\lambda_1^2 tu + \lambda_1\bar{t}}{\lambda_1 t+\bar{t}}r_1^2\right) \times \left(\frac{\lambda_1 t+\bar{t}}{\lambda_1^2 t\bar{u}+\lambda_1\bar{t}}\right)^{n+\ell+1}\left(\frac{\lambda_2 t+\bar{t}}{\lambda_2^2 tu+\lambda_2\bar{t}}\right)^{k-\ell-n+1}\left(\frac{\lambda_2^2}{\lambda_2 t+\bar{t}}\right)^{2k}$$

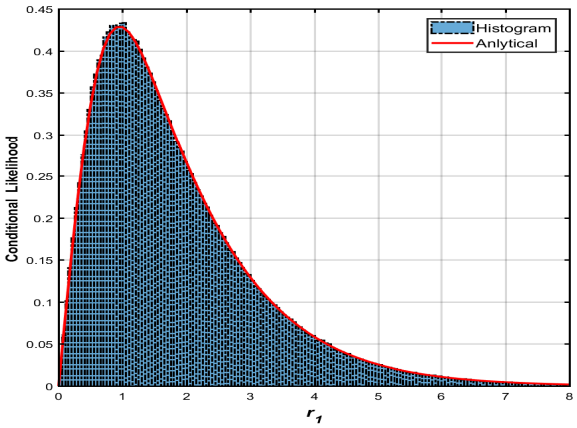

Fig. 1. The Monte Carlo histogram versus the analytical marginal pdf for $T=6$, $k_g = k_h = 0$.

$$\times \left(\frac{\lambda_2 t+\bar{t}}{\lambda_2^2 t\bar{u}+\lambda_2\bar{t}}\right)^{k-\ell-n+1}\left(\frac{\lambda_1^2(\lambda_2 t+\bar{t})}{\lambda_2^2(\lambda_1 t+\bar{t})}\right)^{2(n+\ell)} du dt, \quad (17)$$

where λ_i, $i \in \{1,2\}$, and $\bar{t}$ are defined in Theorem 1.

Proof. See Appendix C. □

The plots of $f_{\boldsymbol{r_1}|\boldsymbol{X}}(r_1|X)$ and the histogram produced by a Monte-Carlo simulation are given in Figure 1. It can be seen form this figure that the derived conditional marginal distribution closely match the histogram.

3) Gaussian Approximation: Due to the fact that the likelihood function based on the Gaussian approximation does not incorporate the jammer's channel k-factor, k_g, cf. Lemma 1 and Lemma 2, approximating the signal resulting form the jammer's transmission at the receiver by Gaussian vectors as in Sections III-A2 and III-C2 yields a likelihood expression identical to the one in (14).

4) The Optimal covariance matrix for Gaussian Approximation: The Gaussian approximation likelihood expression for the cases in which $k_h = 0$, i.e. cases in Sections III-C and III-D, can be optimized by using the covariance matrix that minimizes the Kullback-Leibler distance measure [36] between the Gaussian approximation-based likelihood function in (14) and the exact likelihood functions in (12) and (15). This covariance matrix is given in the following lemma[2].

Lemma 3. *Let $\boldsymbol{Y}$, $\boldsymbol{Y_o}$, $\boldsymbol{Y_e}$, and $\boldsymbol{X}$ be as defined in Theorem 1, and let $\boldsymbol{H} \sim \mathcal{CN}\left(\boldsymbol{0},\sigma_h^2 I_2\right)$ where σ_h is a known positive scalar. The covariance matrix that minimizes the Kullback-Leibler distance measure between the Gaussian approximation-based likelihood function in (14) and the exact likelihood functions in (12) and (15) is given by*

$$\Xi^\star(X) = \frac{\Xi_{Y_o|X} + \Xi_{Y_e|X}}{2}, \quad (18)$$

where $\Xi_{Y_o|X} = \mathrm{E}\{\boldsymbol{Y}_o\boldsymbol{Y}_o^\dagger|\boldsymbol{X}\} = \int_{Y_o} f_{\boldsymbol{Y}_o|\boldsymbol{X}}(Y_o|X)\left(Y_o Y_o^\dagger\right)dY_o$, and $\Xi_{Y_e|X} = \mathrm{E}\{\boldsymbol{Y}_e\boldsymbol{Y}_e^\dagger|\boldsymbol{X}\} =$

[1]The validation has been kindly suggested by an anonymous Reviewer.

[2]This approach has been kindly suggested by an anonymous Reviewer.

$\int_{Y_e} f_{\boldsymbol{Y}_e|\boldsymbol{X}}(Y_e|X)\big(Y_e Y_e^{\dagger}\big)dY_e$, $f_{\boldsymbol{Y}_o|\boldsymbol{X}}$ and $f_{\boldsymbol{Y}_e|\boldsymbol{X}}(Y_e|X)$ are the exact likelihood functions of $\boldsymbol{Y}_o$ and $\boldsymbol{Y}_e$ respectively (See [31]).

Proof. See Appendix D. □

The result of this lemma is rather intuitive. In particular, to approximate a given distribution by a Gaussian one, the optimal covariance matrix that can be used for the Gaussian distribution is the covariance matrix of the realizations of the original distribution.

IV. Simulation Results

In this section, we compare the performance of the ML detector based on the exact likelihood functions provided hereinabove with that of the approximate ML detector based on the Gaussian assumption of the signals received due to the jammer's transmissions. Moreover, we examine the performance gain achieved by using double instead of a single receive antenna in addition to the impact of the symbol length, T, on performance.

In all simulations, we used constellations of unit norm symbols constructed with the method proposed in [37]. In particular, to generate a complex constellation $\mathcal{C}$ with 2^n T-dimensional symbols, we use the following procedure.

1) Generate a pool of I random matrices of size $T \times 2^n$ from the standard complex Gaussian distribution. For each matrix, compute the QR-decomposition, where Q is the unitary component of the matrix and R is its lower triangular component. Let Q_i be the Q-component of the i-th matrix, $i \in \{1, \ldots, I\}$. The set $\{Q_i\}$ is known to be asymptotically isotropically distributed on the group of unitary matrices [37].
2) Out of all the generated matrices, select the matrix whose Q component has the largest angular separation between its columns, that is, select the i^*'th matrix, where $i^* = \arg\min_i \max_{m,n,m\neq n} |(q_m^i)\dagger q_n^i|$, q_r^i is the r-th column of Q_i, $r \in \{1, \ldots, 2^n\}$.
3) Choose the columns of Q_{i^*} to be the points of the signalling constellation.

To facilitate the presentation of results, we fix the total power of the noise plus jamming signal to 10 dB. To do so, in all simulations, we set the noise variance to $\sigma_Z^2 = 0.5$ and the jamming signal variance to $\sigma_V^2 = 4.5$. To maintain a constant average channel power gain for the transmitter and jammer channels under different values of k_h and k_g, the transmitter channel variance, σ_h^2, and the jammer channel variance, σ_g^2, are set to $\sigma_h^2 = \dfrac{1}{k_h^2+1}$ and $\sigma_g^2 = \dfrac{1}{k_g^2+1}$, respectively. In all forthcoming examples, except Example 4 and Example 5, the symbol length is set to $T = 4$, and the data rate is set to 0.75 bits per channel use (bpcu).

Example 1: In this example, we set the k-factor of the jammer's channel, k_g, to zero for various values of the k-factor of the transmitter's channel, k_h. This scenario arises when the transmitter is highly mounted or airborne while the jammer and the receiver are ground-based, cf. Section III-A. Using these settings, we obtained the symbol error rates (SERs) reported in Figure 2 for both the exact and approximate ML detectors. The superiority of the exact ML detector over the approximate one is apparent in Figure 2. Furthermore, it can be seen that the performance of both detectors improves when k_h increases, i.e., when the LOS of the transmitter's channel becomes stronger. For instance, at a signal-to-jamming-plus-noise ratio (SJNR) of 10 dB, increasing the k-factor of the transmitter channel from $k_h = 1$ to $k_h = 3$ reduces the SER from around 9×10^{-3} to about 2×10^{-3} for the exact ML detector and form about 2×10^{-2} to 9×10^{-3} for the approximate one. However, the advantage of the exact ML detector is more pronounced at lower values of k_h. For instance, at an SER of 10^{-3}, the exact ML detector exhibits an advantage of 3 dB when $k_h = 1$ and 2 dB when $k_h = 3$.

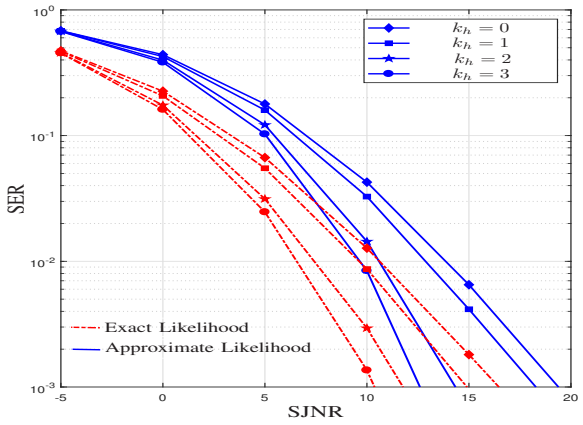

Fig. 2. $T = 4$, $k_g = 0$ and various values of k_h.

This figure illustrates the sharp contrast between the behaviour of the exact and approximate ML detectors for the cases of single and double receive antennas. In particular, as shown in [31], the performance of these detectors coincides when the signalling constellation has a unit norm and the receiver has one antenna. However, the exact ML detector offers a significant performance advantage over its approximate counterpart when the receiver has two antennas, albeit with a higher computational cost. □

Example 2: In this example, we consider a case complementary to that considered in Example 1. In particular, in this example, the k-factor of the transmitter's channel was set to $k_h = 0$, and various values were used for the k-factor of the jammer's channel, k_g. This scenario arises when the jammer is highly mounted or airborne while the transmitter and the receiver are ground-based, cf. Section III-C.

The SER curves obtained in this example are reported in Figure 3 for both the exact and approximate ML detectors. It can be seen from Figure 3 that, as in the cases considered in Example 1, the exact ML detector significantly outperforms the approximate one. For instance, at an SER of 1×10^{-2}, the exact ML detector has an advantage of 3 dB at $k_g = 0$ and 3.25 dB at $k_g = 8$. Furthermore, it can be noticed that, although the detection performance of both detectors is not highly sensitive to the changes in k_g, the detection performance of the exact ML detector improves with the

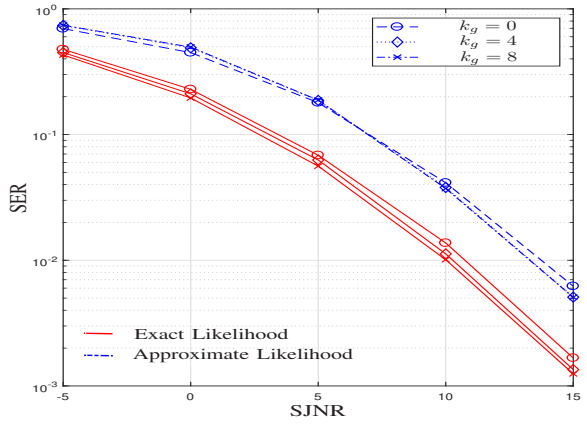

Fig. 3. $T=4$, $k_h=0$ and various values of k_g.

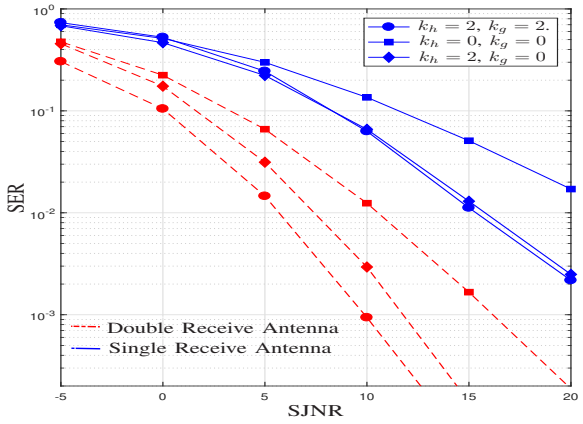

Fig. 4. Detection performance of double receive antenna vs. single receive antenna for $T=4$ and various values of k_h and k_g.

strength of the jammer's LOS component, i.e., increasing k_g decreases the SER. For instance, at an SJNR of 10 dB, the SER drops form 1.4×10^{-2} at $k_g=0$ to 1×10^{-2} at $k_g=8$.

This example suggests that a receiver with an intense LOS component to the jammer is more resilient to jamming. In a complementary fashion, it is more effective for the jammer to be occluded from, than for it to have an LOS component to, the receiver.

Example 3: In this example, we compare the performance achieved by a double-antenna receiver with that achieved by a single-antenna one for various values of k_h and k_g. The SER curves obtained in this example are presented in Figure 4. This figure illustrates the robustness of double-antenna receivers to jamming. For instance, at an SER of 3.3×10^{-3}, the exact ML detector yielded a gain of 11 dB when $k_h=k_g=2$, a gain of 6 dB when $k_h=k_g=0$, and a gain of 9.5 dB when $k_h=2$ and $k_g=0$.

This example suggests that equipping the receiver with more antennas will likely offer significant jamming resilience. Unfortunately, however, as previously pointed out in Section III, deriving the optimal ML detector for cases with more than two receive antennas seems daunting. Furthermore, performing optimal detection with a large number of receive antennas is likely to be computationally expensive unless closed-form expressions are obtained for the respective likelihood functions.

Example 4: In this example, we investigate the effect of the symbol length, T, on the performance of the exact and the Gaussian approximation-based ML detectors. Two values of T were considered, namely, $T=2$ and $T=4$. For each value of T, performance is evaluated at three transmission rates, namely, 0.5, 1, and 2 bpcu. The k-factors of both channels are chosen to be zero, i.e., $k_h=k_g=0$. The results obtained in this example are presented in Figure 5. It can be noticed from this figure that, as commonly noted in the literature, e.g. [38], [39], spreading information across more time slots tends to yield better performance at higher SJNRs. However, this phenomenon does not necessarily carry over to the low SJNR regime. For instance, at an SER of 1×10^{-2}, using $T=4$ has an SJNR advantage of 1.5 dB when the data rate

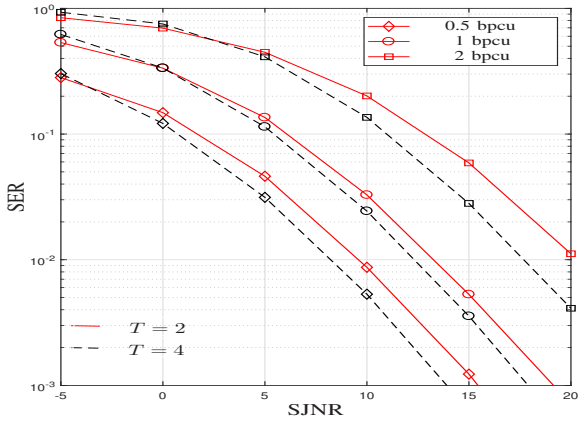

Fig. 5. Performance at different rates and values of T for $k_h=k_g=0$.

is 0.5 bpcu and 2.5 dB when the data rate is 2 bpcu over the case of $T=2$.

Example 5: In this example, we investigate the effect of the optimal covariance matrix in Lemma 3 on the performance of the detector based on Gaussian approximation when $T=6$ and the constellation contains 8 symbols. For ease of computation, we used the sample covariance matrices of Y_o and Y_e to approximate the optimal covariance matrix corresponding to each constellation symbol, X. The results of this investigation are presented in Figure 6. It can be seen form this figure that using the optimized Gaussian approximation detector yields valuable gain over its commonly-used counterpart. For instance, at an SER of 10^{-3}, this gain is about 1 dB. The advantage of the detector based on the exact likelihood function over the commonly-used Gaussian approximation detector at this SER is about 8.5 dB.

V. Conclusion

We considered a scenario in which a single-antenna transmitter communicates with a double-antenna receiver in the presence of a continuous proactive single-antenna jammer. The

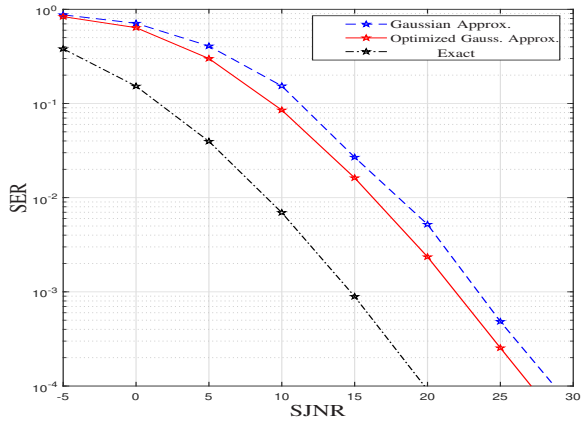

Fig. 6. Detection performance of the Gaussian Approximate, Optimized Gaussian Approximate, and the Exact detector for $T = 6$, $k_h = 0$ and $k_g = 0$.

jammer's transmitted signal, the transmitter-receiver, and the jammer-receiver channels are assumed to be Gaussian, thereby inducing a non-Gaussian jamming signal at the receiver. For this scenario, the likelihood functions of the received signals are derived for the cases in which the receiver has full or partial CDI about the transmitter's channel and the jammer's channel. These likelihood functions are used to develop the optimal ML detector. Furthermore, the likelihood functions based on the Gaussian approximation of the signal component resulting from the jammer's transmission are derived, and the corresponding detector is developed. The detection performance of both detectors is tested through simulations. These simulations revealed the superiority of the ML detector based on the exact likelihood functions over the one based on Gaussian approximations for the case of unit-norm constellations. This is in sharp contrast with the single antenna case wherein the two detectors are equivalent for this type of constellations [31]. Simulation results showed the significant power advantage of the exact ML detector, especially when the transmitter has a strong LOS component. Furthermore, these results showed that a double-antenna receiver with strong LOS components to the transmitter and the jammer is more resilient to proactive barrage jamming than its single-antenna counterpart. Extending these results to receivers with more than two antennas seems daunting and computationally expensive, but it will likely yield valuable performance gains.

Appendix A
Proof of Theorem 1

An outline of this proof is as follows.

- Conditioned on $\boldsymbol{X}$, $\boldsymbol{G}$, $\boldsymbol{V}$ and $\boldsymbol{\theta}_{\mu_h}$, invoke statistical independence to express the pdf of $\boldsymbol{Y}$ as the product of the pdfs of $\boldsymbol{Y}_o$ and $\boldsymbol{Y}_e$.
- To facilitate the separation of the components of $\boldsymbol{V}$, use the eigendecomposition to diagonalize the covariance matrix, Σ^{-1}.
- Average out $\boldsymbol{V}$ to obtain the pdf of $\boldsymbol{Y}$ given $\boldsymbol{X}$, $\boldsymbol{G}$, and $\boldsymbol{\theta}_{\mu_h}$.
- Average out $\boldsymbol{G}$ to obtain the pdf of $\boldsymbol{Y}$ given $\boldsymbol{X}$ and $\boldsymbol{\theta}_{\mu_h}$.
- Average out $\boldsymbol{\theta}_{\mu_h}$ to obtain the pdf of $\boldsymbol{Y}$ given $\boldsymbol{X}$.

Equation (2) can be written as:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{2T-1} \\ Y_{2T} \end{bmatrix} = \begin{bmatrix} H & 0 & \dots & 0 & 0 \\ 0 & H & \dots & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & \dots & H & 0 \\ 0 & 0 & \dots & 0 & H \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{T-1} \\ X_T \end{bmatrix} + \begin{bmatrix} G & 0 & \dots & 0 & 0 \\ 0 & G & \dots & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & \dots & G & 0 \\ 0 & 0 & \dots & 0 & G \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{T-1} \\ V_T \end{bmatrix} + \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{2T-1} \\ Z_{2T} \end{bmatrix}. \tag{19}$$

From this equation, it can be seen that, conditioned on $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{G} = \begin{bmatrix} \boldsymbol{g}_1^* & \boldsymbol{g}_2^* \end{bmatrix}^\dagger$, k_h and $\boldsymbol{\theta}_{\mu_h} = \begin{bmatrix} \boldsymbol{\theta}_{\mu_{h_1}} & \boldsymbol{\theta}_{\mu_{h_2}} \end{bmatrix}^\dagger$, the pdf of the received signal vector, $\boldsymbol{Y}$, is the product of two Gaussian pdfs:

$$f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{V},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,V,\theta_{\mu_h}) = f_{\boldsymbol{Y}_o|\boldsymbol{X},\boldsymbol{V},\boldsymbol{g_1},\boldsymbol{\theta}_{\mu_{h_1}}}(Y_o|X,V,g_1,\theta_{\mu_{h_1}}) \times f_{\boldsymbol{Y}_e|\boldsymbol{X},\boldsymbol{V},\boldsymbol{g_2},\boldsymbol{\theta}_{\mu_{h_2}}}(Y_e|X,V,g_2,\theta_{\mu_{h_2}}), \tag{20}$$

where $Y_o = [Y_1^*, \dots, Y_{2i-1}^*, \dots, Y_{2T-1}^*]^\dagger$, $Y_e = [Y_2^*, \dots, Y_{2i}^*, \dots, Y_{2T}^*]^\dagger$,

$$f_{\boldsymbol{Y}_o|\boldsymbol{X},\boldsymbol{V},\boldsymbol{g}_1,\boldsymbol{\theta}_{\mu_{h_1}}}(Y_o|X,V,g_1,\theta_{\mu_{h_1}}) = \frac{1}{\pi^T|\Sigma|} \times \exp\left(- \left[\tilde{Y}_o - g_1 V\right]^\dagger \Sigma^{-1} \left[\tilde{Y}_o - g_1 V\right] \right), \tag{21}$$

and

$$f_{\boldsymbol{Y}_e|\boldsymbol{X},\boldsymbol{V},\boldsymbol{g}_2,\boldsymbol{\theta}_{\mu_{h_2}}}(Y_e|X,V,g_2,\theta_{\mu_{h_2}}) = \frac{1}{\pi^T|\Sigma|} \times \exp\left(- \left[\tilde{Y}_e - g_2 V\right]^\dagger \Sigma^{-1} \left[\tilde{Y}_e - g_2 V\right] \right), \tag{22}$$

where $\tilde{Y}_o = Y_o - k_h \sigma_h e^{j\boldsymbol{\theta}_{\mu_{h_1}}} X$, $\tilde{Y}_e = Y_e - k_h \sigma_h e^{j\boldsymbol{\theta}_{\mu_{h_2}}} X$, and $\Sigma = \sigma_h^2 X X^\dagger + \sigma_Z^2 I_T$. Using the eigenvalue decomposition, we write $\Sigma^{-1} = Q \Lambda Q^\dagger$. Hence,

$$\begin{aligned} & f_{\boldsymbol{Y}_o|\boldsymbol{X},\boldsymbol{V},\boldsymbol{g}_1,\boldsymbol{\theta}_{\mu_{h_1}}}(Y_o|X,V,g_1,\theta_{\mu_{h_1}}) \\ &= \frac{1}{\pi^T|\Sigma|} \exp\left(- \left[\check{Y}_o - g_1 \hat{V}\right]^\dagger \Lambda \left[\check{Y}_o - g_1 \hat{V}\right] \right), \\ &= \frac{1}{\pi^T|\Sigma|} \exp\left(- \sum_{i=1}^{T} \lambda_i |\check{y}_{2i-1} - g_1 \hat{v}_i|^2 \right), \end{aligned} \tag{23}$$

and

$$\begin{aligned} & f_{\boldsymbol{Y}_e|\boldsymbol{X},\boldsymbol{V},\boldsymbol{g}_2,\boldsymbol{\theta}_{\mu_{h_2}}}(Y_e|X,V,g_2,\theta_{\mu_{h_2}}) \\ &= \frac{1}{\pi^T|\Sigma|} \exp\left(- \left[\check{Y}_e - g_2 \hat{V}\right]^\dagger \Lambda \left[\check{Y}_e - g_2 \hat{V}\right] \right), \\ &= \frac{1}{\pi^T|\Sigma|} \exp\left(- \sum_{i=1}^{T} \lambda_i |\check{y}_{2i} - g_2 \hat{v}_i|^2 \right), \end{aligned} \tag{24}$$

where $\hat{V}=Q^{\dagger}V$, $\check{Y}_o=Q^{\dagger}\tilde{Y}_o$ and $\check{Y}_e=Q^{\dagger}\tilde{Y}_e$. Using (23) and (24) in (20) yields:

$$\begin{aligned}
&f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{V},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,V,\theta_{\mu_h})\\
&=\frac{1}{\pi^{2T}|\Sigma|^2}\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}-g_1\hat{v}_i|^2+|\check{y}_{2i}-g_2\hat{v}_i|^2\right)\right),\\
&=\frac{1}{\pi^{2T}|\Sigma|^2}\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\\
&\times\exp\left(-\sum_{i=1}^{T}\lambda_i\left(\|g\|^2|\hat{v}_i|^2-2\Re\{(g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i})\,\hat{v}_i^*\}\right)\right).
\end{aligned}\tag{25}$$

where $\|g\|^2=|g_1|^2+|g_2|^2$. Note that $\Re\{(g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i})\,\hat{v}_i^*\}=|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}||\hat{v}_i|\cos\left(\theta_{\hat{v}_i^*}-\alpha_i\right)$, where $\alpha_i=\angle\left(g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}\right)$, and $\theta_{\hat{v}_i^*}=\angle\hat{v}_i^*=-\angle\hat{v}_i$.

To obtain an expression for $f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h})$, we average both sides of (25) over the random vector $\boldsymbol{V}$. Note that, since Q is a unitary matrix, i.e., $|Q|=1$; hence, $\hat{\boldsymbol{V}}$ has the same pdf as $\boldsymbol{V}$. Doing so yields:

$$\begin{aligned}
&f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h})=\\
&\qquad\int_{\boldsymbol{V}}f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{V},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,V,\theta_{\mu_h})f_{\boldsymbol{V}}(V)dV,
\end{aligned}\tag{26}$$

where $f_{\boldsymbol{V}}(V)$ is given by:

$$f_{\boldsymbol{V}}(V)=\frac{1}{\pi^T\sigma_v^{2T}}\exp\left(-\frac{\|V\|^2}{\sigma_v^2}\right).\tag{27}$$

Using (25) and (27) in (26) and taking into account that $\|V\|=\|\hat{V}\|$ and that $\cos\left(\theta_{\hat{v}_i^*}-\alpha_i\right)=\cos\left(\theta_{\hat{v}_i}+\alpha_i\right)$ yields:

$$\begin{aligned}
&f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h})=\\
&\frac{1}{\pi^{3T}\sigma_v^{2T}|\Sigma|^2}\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\times\\
&\int_{\hat{\boldsymbol{V}}}\exp\left(2\sum_{i=1}^{T}\lambda_i|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}||\hat{v}_i|\cos\left(\theta_{\hat{v}_i}+\alpha_i\right)\right)\\
&\qquad\times\exp\left(-\sum_{i=1}^{T}\left(\lambda_i\|g\|^2+\frac{1}{\sigma_v^2}\right)|\hat{v}_i|^2\right)d\hat{V}.
\end{aligned}\tag{28}$$

Using polar coordinates with $|\hat{v}_i|=r_i$, $\theta_i=\theta_{\hat{v}_i}$ and $d\hat{V}=\prod_{i=1}^{T}r_i dr_i d\theta_i$ yields:

$$\begin{aligned}
&f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h})=\\
&\frac{1}{\pi^{3T}\sigma_v^{2T}|\Sigma|^2}\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\times\\
&\prod_{i=1}^{T}\int_0^{\infty}r_i\exp\left(-\left(\lambda_i\|g\|^2+\frac{1}{\sigma_v^2}\right)r_i^2\right)\times\\
&\int_0^{2\pi}\exp\left(2\lambda_i|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}|r_i\cos\left(\theta_i+\alpha_i\right)\right)d\theta_i\\
&\qquad\times dr_i.
\end{aligned}\tag{29}$$

We note that the zeroth-order modified Bessel function of the first kind is given by

$$\begin{aligned}
I_0(z)&=\frac{1}{2\pi}\int_0^{2\pi}\exp\left(z\cos\left(\theta+\alpha\right)\right)d\theta,\\
&=\sum_{k=0}^{\infty}\frac{(z/2)^{2k}}{(k!)^2}.
\end{aligned}$$

Hence, the inner integration in (29) can be expressed as:

$$\begin{aligned}
&2\pi I_0\left(2\lambda_i|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}|r_i\right)=\\
&\qquad 2\pi\sum_{\ell=0}^{\infty}\frac{\left(\lambda_i^2|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}|^2r_i^2\right)^{\ell}}{(\ell!)^2}.
\end{aligned}\tag{30}$$

Using (30) in (29) yields:

$$\begin{aligned}
&f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h})=\frac{1}{\pi^{2T}\sigma_v^{2T}|\Sigma|^2}\times\\
&\qquad\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\times\\
&\qquad\prod_{i=1}^{T}\sum_{\ell=0}^{\infty}\frac{\left(\lambda_i^2|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}|^2\right)^{\ell}}{(\ell!)^2}\times\\
&\qquad\int_0^{\infty}r_i^{2\ell}\exp\left(-\left(\lambda_i\|g\|^2+\frac{1}{\sigma_v^2}\right)r_i^2\right)2r_i dr_i
\end{aligned}\tag{31}$$

$$\begin{aligned}
&=\frac{1}{\pi^{2T}|\Sigma|^2}\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\times\\
&\qquad\prod_{i=1}^{T}\left(\frac{1}{\lambda_i\sigma_v^2\|g\|^2+1}\right)\times\\
&\qquad\sum_{\ell=0}^{\infty}\left(\frac{\lambda_i^2\sigma_v^2|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}|^2}{\lambda_i\sigma_v^2\|g\|^2+1}\right)^{\ell}\frac{1}{(\ell!)}
\end{aligned}\tag{32}$$

$$\begin{aligned}
&=\frac{1}{\prod_{i=1}^{T}(\lambda_i\sigma_v^2\|g\|^2+1)}\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\\
&\qquad\times\frac{1}{\pi^{2T}|\Sigma|^2}\exp\left(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2|g_1^*\check{y}_{2i-1}+g_2^*\check{y}_{2i}|^2}{\lambda_i\sigma_v^2\|g\|^2+1}\right),
\end{aligned}\tag{33}$$

where, to write (32), we have used the fact that the integration in (31) is equal to $(\ell!)\left(\frac{\sigma_v^2}{\lambda_i\sigma_v^2\|g\|^2+1}\right)^{\ell+1}$, and to write (33), we have replaced the summation in (32) with the exponential function.

Let S_1 be the summation in the argument of the last exponential in (33). This summation is given by

$$\begin{aligned}
S_1=\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2}{\lambda_i\sigma_v^2\|g\|^2+1}\Big(&|\check{y}_{2i-1}|^2|g_1|^2+|\check{y}_{2i}|^2|g_2|^2\\
&+2|\check{y}_{2i-1}||\check{y}_{2i}||g_1||g_2|\cos\left(\theta_{g_1}-\theta_{g_2}-\theta_i\right)\Big),
\end{aligned}\tag{34}$$

where $\theta_{g_\ell}=\angle g_\ell$, $\ell\in\{1,2\}$, and $\theta_i=\angle\check{y}_{2i-1}-\angle\check{y}_{2i}$. The last term in (34) can be written as:

$$2|g_1||g_2|\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2|\check{y}_{2i-1}||\check{y}_{2i}|\cos\left(\theta_{g_1}-\theta_{g_2}-\theta_i\right)}{\lambda_i\sigma_v^2\|g\|^2+1}$$

$$= 2|g_1||g_2|\Big(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2|\check{y}_{2i-1}||\check{y}_{2i}|\cos(\theta_i)}{\lambda_i\sigma_v^2\|g\|^2+1}\Big)\cos(\theta_{g_1}-\theta_{g_2})$$
$$+ 2|g_1||g_2|\Big(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2|\check{y}_{2i-1}||\check{y}_{2i}|\sin(\theta_i)}{\lambda_i\sigma_v^2\|g\|^2+1}\Big)\sin(\theta_{g_1}-\theta_{g_2}),$$
$$= 2|g_1||g_2|\bar{C}_{\check{Y}}(\|g\|^2)\times$$
$$(\cos(\beta)\cos(\theta_{g_1}-\theta_{g_2})+\sin(\beta)\sin(\theta_{g_1}-\theta_{g_2})),$$
$$= 2|g_1||g_2|\bar{C}_{\check{Y}}(\|g\|^2)\cos(\theta_{g_1}-\theta_{g_2}-\beta), \quad (35)$$

where $\bar{C}_{\check{Y}}(\|g\|^2) = \sqrt{C_1^2+C_2^2}$, $\beta = \arctan\left(\frac{C_2}{C_1}\right)$, and $C_1 = \sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2|\check{y}_{2i-1}||\check{y}_{2i}|\cos(\theta_i)}{\lambda_i\sigma_v^2\|g\|^2+1}$, and $C_2 = \sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2|\check{y}_{2i-1}||\check{y}_{2i}|\sin(\theta_i)}{\lambda_i\sigma_v^2\|g\|^2+1}$. Using $\tilde{\Lambda}$ to denote the diagonal matrix whose ℓ-th element is given by $\tilde{\Lambda}_\ell = \frac{\lambda_\ell^2\sigma_v^2}{\lambda_\ell\sigma_v^2\|g\|^2+1}$, it can be verified that $C_1 = \Re\{\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e\}$ and $C_2 = \Im\{\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e\}$. Hence,

$$\bar{C}_{\check{Y}}^2(\|g\|^2) = \sigma_v^4\sum_{\substack{m=1\\n=1}}^{T}\frac{|\check{y}_{2m-1}||\check{y}_{2n-1}||\check{y}_{2m}||\check{y}_{2n}|\cos(\theta_m-\theta_n)}{\lambda_m^{-2}\lambda_n^{-2}(\lambda_m\sigma_v^2\|g\|^2+1)(\lambda_n\sigma_v^2\|g\|^2+1)},$$
$$= |\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e|^2. \quad (36)$$

Using (35) and (36) in (34) yields:

$$S_1 = \Big(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2\left(|\check{y}_{2i-1}|^2|g_1|^2+|\check{y}_{2i}|^2|g_2|^2\right)}{\lambda_i\sigma_v^2\|g\|^2+1} + 2|g_1||g_2||\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e|\cos(\theta_{g_1}-\theta_{g_2}-\beta)\Big). \quad (37)$$

Substituting from (37) in (33) yields

$$f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h}) = \frac{1}{\pi^{2T}|\Sigma|^2\prod_{i=1}^{T}(\lambda_i\sigma_v^2\|g\|^2+1)}$$
$$\times\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\times$$
$$\exp\left(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2\left(|\check{y}_{2i-1}|^2|g_1|^2+|\check{y}_{2i}|^2|g_2|^2\right)}{\lambda_i\sigma_v^2\|g\|^2+1}\right)$$
$$\times\exp\Big(2|g_1||g_2||\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e|\cos(\theta_{g_1}-\theta_{g_2}-\beta)\Big). \quad (38)$$

To obtain an expression for $f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h})$, we average $f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h})$ over the distributions of $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$. In particular, we write

$$f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h}) = \int_{g_1}\int_{g_2} f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{G},\boldsymbol{\theta}_{\mu_h}}(Y|X,G,\theta_{\mu_h}) \times f_{\boldsymbol{g}_1}(g_1)f_{\boldsymbol{g}_2}(g_2)dg_1dg_2, \quad (39)$$

where we have used the fact that $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ are independent and identically distributed.

We will derive $f_{\boldsymbol{g}_1}(g_1)$ and the derivation of $f_{\boldsymbol{g}_2}(g_2)$ follows *mutatis mutandis*.

$$f_{\boldsymbol{g}_1}(g_1) = \frac{1}{2\pi}\int_0^{2\pi} f_{\boldsymbol{g}_1|\bar{\theta}_{\mu_{g_1}}}(g_1|\theta_{\mu_{g_1}})d\theta_{\mu_{g_1}},$$
$$= \frac{1}{2\pi^2\sigma_g^2}\int_0^{2\pi}\exp\left(\frac{-|g_1-k_g\sigma_g e^{j\theta_{\mu_{g_1}}}|^2}{\sigma_g^2}\right)d\theta_{\mu_{g_1}},$$
$$= \frac{\exp\left(-k_g^2-\frac{|g_1|^2}{\sigma_g^2}\right)}{2\pi^2\sigma_g^2}\int_0^{2\pi}\exp\left(\frac{2k_g|g_1|\cos\left(\theta_{\mu_{g_1}}-\theta_{g_1}\right)}{\sigma_g}\right)d\theta_{\mu_{g_1}}$$
$$= \frac{1}{\pi\sigma_g^2}\exp\left(-k_g^2-\frac{|g_1|^2}{\sigma_g^2}\right)I_0\left(\frac{2k_g|g_1|}{\sigma_g}\right). \quad (40)$$

Analogously, we have

$$f_{\boldsymbol{g}_2}(g_2) = \frac{1}{\pi\sigma_g^2}\exp\left(-k_g^2-\frac{|g_2|^2}{\sigma_g^2}\right)I_0\left(\frac{2k_g|g_2|}{\sigma_g}\right). \quad (41)$$

Using (38), (40) and (41) in (39) yields:

$$f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h}) =$$
$$\frac{\exp(-2k_g^2)}{\pi^{2T+2}\sigma_g^4|\Sigma|^2}\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\times$$
$$\int_{g_1}\int_{g_2}\frac{I_o\left(\frac{2k_g|g_1|}{\sigma_g}\right)I_o\left(\frac{2k_g|g_2|}{\sigma_g}\right)}{\prod_{i=1}^{T}(\lambda_i\sigma_v^2\|g\|^2+1)}\exp\left(-\frac{|g_1|^2+|g_2|^2}{\sigma_g^2}\right)$$
$$\times\exp\left(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2\left(|\check{y}_{2i-1}|^2|g_1|^2+|\check{y}_{2i}|^2|g_2|^2\right)}{\lambda_i\sigma_v^2\|g\|^2+1}\right)\times$$
$$\exp\Big(2|g_1||g_2||\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e|\cos(\theta_{g_1}-\theta_{g_2}-\beta)\Big)\,dg_1dg_2, \quad (42)$$

Note that $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ are complex variables. Hence, the double integration in (42) is equivalent to quadruple real integrations.

Let χ_0 be the integration in (42). Using polar representation, we write $|g_\ell| = r_\ell$, and $dg_\ell = r_\ell dr_\ell d\theta_{g_\ell}$, $\ell\in\{1,2\}$, and hence

$$\chi_0 = \int_0^\infty\int_0^\infty\frac{I_o\left(\frac{2k_gr_1}{\sigma_g}\right)I_o\left(\frac{2k_gr_2}{\sigma_g}\right)}{\prod_{i=1}^{T}(\lambda_i\sigma_v^2(r_1^2+r_2^2)+1)}\times$$
$$\exp\left(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2\left(|\check{y}_{2i-1}|^2r_1^2+|\check{y}_{2i}|^2r_2^2\right)}{\lambda_i\sigma_v^2(r_1^2+r_2^2)+1}-\frac{r_1^2+r_2^2}{\sigma_g^2}\right)\times$$
$$\int_0^{2\pi}\int_0^{2\pi}\exp\Big(2r_1r_2|\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e|\cos(\theta_{g_1}-\theta_{g_2}-\beta)\Big)\,d\theta_{g_1}d\theta_{g_2}$$
$$\times r_1r_2dr_1dr_2. \quad (43)$$

Let χ_1 be the inner double integration in (43), it can be shown that

$$\chi_1 = 4\pi^2I_0\left(2r_1r_2|\check{Y}_o^\dagger\ \tilde{\Lambda}\ \check{Y}_e|\right),$$
$$= 4\pi^2I_0\left(2r_1r_2\ \bar{C}_{\check{Y}}\left(r_1^2+r_2^2\right)\right), \quad (44)$$

Using the following change of variable: $\rho^2 = r_1^2+r_2^2$, $\phi = \arctan\left(\frac{r_2}{r_1}\right)$, $r_1 = \rho\cos(\phi)$, $r_2 = \rho\sin(\phi)$, $dr_1dr_2 = \rho d\rho d\phi$ yields:

$$f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h}) = \frac{2\exp(-2k_g^2)}{\pi^{2T}\sigma_g^4|\Sigma|^2}$$
$$\times\exp\left(-\sum_{i=1}^{T}\lambda_i\left(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\right)\right)\int_0^\infty\int_0^{\frac{\pi}{2}}\frac{2\rho^3\sin(\phi)\cos(\phi)}{\prod_{i=1}^{T}(\lambda_i\sigma_v^2\rho^2+1)}$$

$$\times I_0\Big(\frac{2k_g\rho\cos(\phi)}{\sigma_g}\Big)I_0\Big(\frac{2k_g\rho\sin(\phi)}{\sigma_g}\Big)I_0\big(\rho^2\sin(2\phi)\ \ \bar{C}_{\check{Y}}\left(\rho^2\right)\big)$$
$$\times\exp\left(\sum_{i=1}^{T}\frac{\lambda_i^2\sigma_v^2\left(|\check{y}_{2i-1}|^2\cos^2(\phi)+|\check{y}_{2i}|^2\sin^2(\phi)\right)\rho^2}{\lambda_i\sigma_v^2\rho^2+1}-\frac{\rho^2}{\sigma_g^2}\right)$$
$$\times d\phi d\rho. \quad (45)$$

Using the change of variables: $u = \sin^2(\phi)$, $du = 2\sin(\phi)\cos(\phi)\,d\phi$, $t = \dfrac{\sigma_v^2\ \rho^2}{\sigma_v^2\ \rho^2+1}$ and $\dfrac{dt}{\sigma_v^2(1-t)^2} = 2\rho d\rho$ and defining $\bar{t} = 1-t$ and $\bar{u} = 1-u$ yields:

$$f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h}) = \frac{\exp\left(-2k_g^2\right)}{\pi^{2T}\sigma_g^4\sigma_v^4|\Sigma|^2}$$
$$\times\exp\left(-\sum_{i=1}^{T}\lambda_i\big(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\big)\right)\int_0^1\int_0^1\frac{t\ \bar{t}^{T-3}}{\prod_{i=1}^{T}(\lambda_i t+\bar{t})}$$
$$\times I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{t\bar{u}}{\bar{t}}}\right)I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{tu}{\bar{t}}}\right)I_0\left(2t\sqrt{u\bar{u}}\bar{C}_{\check{Y}}\left(\frac{t}{\sigma_v^2\bar{t}}\right)\right)$$
$$\times\exp\left(\sum_{i=1}^{T}\frac{\lambda_i^2\left(|\check{y}_{2i-1}|^2\bar{u}+|\check{y}_{2i}|^2u\right)t}{\lambda_i t+\bar{t}}-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right)dudt, \quad (46)$$

It can be shown that the argument of the last Bessel function can be written as:

$$2t\sqrt{u\bar{u}}\bar{C}_{\check{Y}}\left(\frac{t}{\sigma_v^2\bar{t}}\right) = |\check{Y}_o^{\dagger}\ \bar{\Lambda}(u,t)\ \check{Y}_e|.$$

where $\bar{\Lambda}(u,t)$ is a diagonal matrix whose ℓ-th element is given by $\bar{\Lambda}_\ell = \dfrac{2\ \lambda_\ell^2\sqrt{u\bar{u}}\ t}{\lambda_\ell t+\bar{t}}$.

To obtain an expression for $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$ we average $f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h})$ over the independent uniformly distributed phases $\boldsymbol{\theta}_{\mu_{h_1}}$ and $\boldsymbol{\theta}_{\mu_{h_2}}$. In particular, we write

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h})d\theta_{\mu_{h_1}}d\theta_{\mu_{h_2}}. \quad (47)$$

Using (46) in (47) yields:

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{\exp(-2k_g^2)}{4\pi^{2T+2}\sigma_g^4\sigma_v^4|\Sigma|^2}\times$$
$$\int_0^{2\pi}\int_0^{2\pi}\exp\left(-\sum_{i=1}^{T}\lambda_i\big(|\check{y}_{2i-1}|^2+|\check{y}_{2i}|^2\big)\right)\Psi(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}})d\theta_{\mu_{h_1}}d\theta_{\mu_{h_2}}, \quad (48)$$

where

$$\Psi(\theta_{\mu_{h_1}},\theta_{\mu_{h_2}}) = \int_0^1\int_0^1\frac{t\bar{t}^{T-3}}{\prod_{i=1}^{T}(\lambda_i t+\bar{t})}I_0\left(|\check{Y}_o^{\dagger}\bar{\Lambda}(u,t)\check{Y}_e|\right)$$
$$\times\exp\left(\sum_{i=1}^{T}\frac{\lambda_i^2\left(|\check{y}_{2i-1}|^2\bar{u}+|\check{y}_{2i}|^2u\right)t}{\lambda_i t+\bar{t}}-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right)$$
$$\times I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{\bar{u}t}{\bar{t}}}\right)I_0\left(\frac{2k_g}{\sigma_g\sigma_v}\sqrt{\frac{ut}{\bar{t}}}\right)dudt, \quad (49)$$

The desired pdf is given by the expressions in (48) and (49). These expressions can be presented more succinctly by introducing matrix notation. In particular, using the fact that $\check{Y}_o = Q^{\dagger}\tilde{Y}_o = Q^{\dagger}\left(Y_o - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_1}}}X\right)$ and $\check{Y}_e = Q^{\dagger}\tilde{Y}_e = Q^{\dagger}\left(Y_e - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_2}}}X\right)$, along with the fact that $\Sigma^{-1} = Q\Lambda Q^{\dagger}$, for (48) we can write

$$\sum_{i=1}^{T}\lambda_i|\check{y}_{2i-1}|^2 = (Y_o - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_1}}}X)\Sigma^{-1}(Y_o - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_1}}}X), \quad (50)$$

$$\sum_{i=1}^{T}\lambda_i|\check{y}_{2i}|^2 = (Y_e - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_2}}}X)^{\dagger}\Sigma^{-1}(Y_e - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_2}}}X), \quad (51)$$

and, for (49) we can write:

$$\sum_{i=1}^{T}\frac{\lambda_i^2|\check{y}_{2i-1}|^2\bar{u}t}{\lambda_i t+\bar{t}} = \bar{u}(Y_o - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_1}}}X)^{\dagger}\Upsilon(\Sigma,t)(Y_o - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_1}}}X), \quad (52)$$

where $\Upsilon(\Sigma,t) = \Sigma^{-1}(I_T + \frac{\bar{t}}{t}\Sigma)^{-1}$. Analogously, it can be shown that:

$$\sum_{i=1}^{T}\frac{\lambda_i^2|\check{y}_{2i}|^2ut}{\lambda_i t+\bar{t}} = u(Y_e - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_2}}}X)^{\dagger}\Upsilon(\Sigma,t)(Y_e - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_2}}}X), \quad (53)$$

$$|\check{Y}_o^{\dagger}\ \bar{\Lambda}(u,t)\ \check{Y}_e| = 4u\bar{u}t\times |(Y_o - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_1}}}X)^{\dagger}\Upsilon(\Sigma,t)(Y_e - k_h\sigma_h e^{\jmath\boldsymbol{\theta}_{\mu_{h_2}}}X)|, \quad (54)$$

and

$$\prod_{i=1}^{T}(\lambda_i t+\bar{t}) = |\Lambda t+\bar{t}I_T|. \quad (55)$$

Using equations (50), (51), (52), (53), (54), and (55) in (48) and (49) completes the proof.

Appendix B
Proof of Lemma 1

To prove this lemma, we assume in (19) that the signal resulting form the jammer transmission at the i-th receive antenna, $\boldsymbol{J}_i = \boldsymbol{g}_i\boldsymbol{V}$, $i \in \{1,2\}$ is Gaussian distributed, i.e., $\boldsymbol{J}_i \sim \mathcal{CN}\left(\boldsymbol{0},\sigma_g^2\sigma_v^2I_T\right)$. Under this assumption, it can be seen from (19) that, conditioned on the vector of transmitted symbols, $\boldsymbol{X}$, the k-factor of the transmitter channel, k_h, and the phases vector of the transmitter channel mean $\boldsymbol{\theta}_{\mu_h} = [\ \boldsymbol{\theta}_{\mu_{h_1}}^*\ \boldsymbol{\theta}_{\mu_{h_2}}^*\ ]^{\dagger}$, the pdf of the received signal vector, $\boldsymbol{Y}$, is given by:

$$f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h}) = f_{\boldsymbol{Y}_o|\boldsymbol{X},\boldsymbol{\theta}_{\mu_{h_1}}}(Y_o|X,\theta_{\mu_{h_1}})f_{\boldsymbol{Y}_e|\boldsymbol{X},\boldsymbol{\theta}_{\mu_{h_2}}}(Y_e|X,\theta_{\mu_{h_2}}), \quad (56)$$

where $\boldsymbol{Y}_o$ and $\boldsymbol{Y}_e$ are Gaussian distributed with the following pdfs:

$$f_{\boldsymbol{Y}_o|\boldsymbol{X},\boldsymbol{\theta}_{\mu_{h_1}}}(Y_o|X,\theta_{\mu_{h_1}}) = \frac{1}{\pi^T|\Sigma_N|}\times$$

$$\exp\left(-(Y_o - k_h\sigma_h e^{j\theta_{\mu_{h_1}}}X)^\dagger \Sigma_N^{-1}(Y_o - k_h\sigma_h e^{j\theta_{\mu_{h_1}}}X)\right), \tag{57}$$

and

$$f_{\boldsymbol{Y}_e|\boldsymbol{X},\boldsymbol{\theta}_{\mu_{h_2}}}(Y_e|X,\theta_{\mu_{h_2}}) = \frac{1}{\pi^T|\Sigma_N|}\times \exp\left(-(Y_e - k_h\sigma_h e^{j\theta_{\mu_{h_2}}}X)^\dagger \Sigma_N^{-1}(Y_e - k_h\sigma_h e^{j\theta_{\mu_{h_2}}}X)\right), \tag{58}$$

where $\Sigma_N = \sigma_h^2 XX^\dagger + (\sigma_g^2\sigma_v^2 + \sigma_Z^2)I_T$. Using the eigendecomposition $\Sigma_N^{-1} = Q_N\Lambda_N Q_N^\dagger$, yields:

$$\begin{aligned} &f_{\boldsymbol{Y}_o|\boldsymbol{X},\boldsymbol{\theta}_{\mu_{h_1}}}(Y_o|X,\theta_{\mu_{h_1}}) \\ &= \frac{1}{\pi^T|\Sigma_N|}\exp\left(-\sum_{i=1}^{T}\lambda_i|\hat{y}_{2i-1} - k_h\sigma_h\hat{x}_i e^{j\theta_{\mu_{h_1}}}|^2\right) \\ &= \frac{1}{\pi^T|\Sigma_N|}\exp\left(-\sum_{i=1}^{T}\lambda_i(|\hat{y}_{2i-1}|^2 + k_h^2\sigma_h^2|\hat{x}_i|^2)\right) \\ &\quad\times\exp\left(2k_h\sigma_h\sum_{i=1}^{T}\lambda_i|\hat{y}_{2i-1}||\hat{x}_i|\cos(\theta_{\mu_{h_1}} + \alpha_i)\right), \end{aligned} \tag{59}$$

where $\hat{y}_{2i-1}$, $\hat{y}_{2i}$ and $\hat{x}_i$ are the i-th elements of $\hat{Y}_o = Q_N^\dagger Y_o$, $\hat{Y}_e = Q_N^\dagger Y_e$, $\hat{X} = Q_N^\dagger X$, respectively, and $\alpha_i = \angle\hat{x}_i - \angle\hat{y}_{2i-1}$.

Denoting the summation in the argument of the last exponential in (59) by S_N, we write

$$\begin{aligned} S_N &= \left(\sum_{i=1}^{T}\lambda_i|\hat{y}_{2i-1}||\hat{x}_i|\cos(\alpha_i)\right)\cos(\theta_{\mu_{h_1}}) \\ &\quad + \left(\sum_{i=1}^{T}\lambda_i|\hat{y}_{2i-1}||\hat{x}_i|\sin(\alpha_i)\right)\sin(\theta_{\mu_{h_1}}) \\ &= \sqrt{W_1^2 + W_2^2}\left(\cos(\beta_o)\cos(\theta_{\mu_{h_1}}) + \sin(\beta_o)\sin(\theta_{\mu_{h_1}})\right) \\ &= \sqrt{W_1^2 + W_2^2}\cos(\theta_{\mu_{h_1}} - \beta_o), \end{aligned} \tag{60}$$

where $W_1 = \sum_{i=1}^{T}\lambda_i|\hat{y}_{2i-1}||\hat{x}_i|\cos(\alpha_i)$, $W_2 = \sum_{i=1}^{T}\lambda_i|\hat{y}_{2i-1}||\hat{x}_i|\sin(\alpha_i)$, and $\beta_o = \arctan\left(\frac{W_2}{W_1}\right)$. It can be shown that $W_1 = \Re\{\hat{Y}_o^\dagger\,\Lambda_N\,\hat{X}\}$ and $W_2 = \Im\{\hat{Y}_o^\dagger\,\Lambda_N\,\hat{X}\}$, which yields that

$$\begin{aligned} &W_1^2 + W_2^2 \\ &= \sum_{m,n=1}^{T}\lambda_m\lambda_n|\hat{y}_{2m-1}||\hat{y}_{2n-1}||\hat{x}_{2m}||\check{x}_{2n}|\cos(\alpha_m - \alpha_n) \\ &= |\hat{Y}_o^\dagger\,\Lambda_N\,\hat{X}|^2, \end{aligned} \tag{61}$$

where Λ_N is the diagonal matrix of eigenvalues of Σ_N^{-1}. Using (60) and (61) in (59) yields:

$$\begin{aligned} f_{\boldsymbol{Y}_o|\boldsymbol{X},\boldsymbol{\theta}_{\mu_{h_1}}}(Y_o|X,\theta_{\mu_{h_1}}) &= \frac{1}{\pi^T|\Sigma_N|} \\ &\times\exp\left(-\sum_{i=1}^{T}\lambda_i(|\hat{y}_{2i-1}|^2 + k_h^2\sigma_h^2|\hat{x}_i|^2)\right) \\ &\times\exp\left(2k_h\sigma_h|\hat{Y}_o^\dagger\Lambda_N\hat{X}|\cos(\theta_{\mu_{h_1}} - \beta_o)\right). \end{aligned} \tag{62}$$

Analogously, it can be shown that:

$$\begin{aligned} f_{\boldsymbol{Y}_e|\boldsymbol{X},\boldsymbol{\theta}_{\mu_{h_2}}}(Y_e|X,\theta_{\mu_{h_2}}) &= \frac{1}{\pi^T|\Sigma_N|} \\ &\times\exp\left(-\sum_{i=1}^{T}\lambda_i(|\hat{y}_{2i}|^2 + k_h^2\sigma_h^2|\hat{x}_i|^2)\right) \\ &\times\exp\left(2k_h\sigma_h|\hat{Y}_e^\dagger\Lambda_N\hat{X}|\cos(\theta_{\mu_{h_2}} - \beta_e)\right). \end{aligned} \tag{63}$$

Using (62) and (63) in (56) yields:

$$\begin{aligned} f_{\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\theta}_{\mu_h}}(Y|X,\theta_{\mu_h}) &= \frac{1}{\pi^{2T}|\Sigma_N|^2} \\ &\times\exp\left(-\sum_{i=1}^{T}\lambda_i(|\hat{y}_{2i-1}|^2 + |\hat{y}_{2i}|^2 + 2k_h^2\sigma_h^2|\hat{x}_i|^2)\right) \\ &\times\exp\left(2k_h\sigma_h|\hat{Y}_o^\dagger\Lambda_N\hat{X}|\cos(\theta_{\mu_{h_1}} - \beta_o)\right) \\ &\times\exp\left(2k_h\sigma_h|\hat{Y}_e^\dagger\Lambda_N\hat{X}|\cos(\theta_{\mu_{h_2}} - \beta_e)\right). \end{aligned} \tag{64}$$

To obtain $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$, we average the right hand side of (64) over $\boldsymbol{\theta}_{\mu_{h_1}}$ and $\boldsymbol{\theta}_{\mu_{h_2}}$, which yields:

$$\begin{aligned} f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) &= \frac{1}{\pi^{2T}|\Sigma_N|^2} \\ &\times\exp\left(-\sum_{i=1}^{T}\lambda_i(|\hat{y}_{2i-1}|^2 + |\hat{y}_{2i}|^2 + 2k_h^2\sigma_h^2|\hat{x}_i|^2)\right) \\ &\times I_0\left(2k_h\sigma_h|\hat{Y}_o^\dagger\;\Lambda_N\;\hat{X}|\right) I_0\left(2k_h\sigma_h|\hat{Y}_e^\dagger\;\Lambda_N\;\hat{X}|\right). \end{aligned} \tag{65}$$

It can be readily verified that:

$$\sum_{i=1}^{T}\lambda_i|\hat{y}_{2i-1}|^2 = Y_o\Sigma_N^{-1}Y_o, \tag{66}$$

$$\sum_{i=1}^{T}\lambda_i|\hat{y}_{2i}|^2 = Y_e^\dagger\;\Sigma_N^{-1}\;Y_e, \tag{67}$$

and

$$\sum_{i=1}^{T}\lambda_i|\hat{x}_i|^2 = X^\dagger\;\Sigma_N^{-1}\;X. \tag{68}$$

Using (66), (67), and (68) in (65) completes the proof.

Appendix C
Derivation of the Marginal pdf $f_{\boldsymbol{r}_1|\boldsymbol{X}}(r_1|X)$

For the case of $k_h = k_g = 0$, the distribution of the received signal vector, $\boldsymbol{Y}$, conditioned on the transmitted constellation symbol, $\boldsymbol{X}$, is given by $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$ in Eq. (15) on page 6 of the revised manuscript, that is,

$$\begin{aligned} &f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \\ &\frac{1}{\pi^{2T}\sigma_g^{2T}\sigma_v^{2T}|\Sigma|^2}\exp\left(-Y_o^\dagger\;\Sigma^{-1}\;Y_o - Y_e^\dagger\;\Sigma^{-1}Y_e\right) \\ &\times\int_0^1\int_0^1\frac{t\;(\;\bar{t}\;)^{T-3}}{|\Lambda t + \bar{t}\;I_T|}I_0\left(2\sqrt{u\bar{u}}t|Y_o^\dagger\;\Upsilon(\Sigma,t)\;Y_e|\right) \\ &\times\exp\left(\bar{u}\;Y_o^\dagger\;\Upsilon(\Sigma,t)\;Y_o + u\;Y_e^\dagger\;\Upsilon(\Sigma,t)\;Y_e - \frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right) \\ &\times du\;dt, \end{aligned} \tag{69}$$

where Y_o, Y_e, Σ, Λ, Υ, σ_g and σ_h are as defined in Theorem 1.

Using $T = 2$, $\Sigma^{-1} = Q\Lambda Q^\dagger$, and setting $\begin{bmatrix}\hat{Y}_o\\ \hat{Y}_e\end{bmatrix} = \begin{bmatrix}Q^\dagger & \mathbf{0}\\ \mathbf{0} & Q^\dagger\end{bmatrix}\begin{bmatrix}Y_o\\ Y_e\end{bmatrix}$. Since Q is a unitary matrix, the determinant of the Jacobian of the transformation from $\begin{bmatrix}Y_o\\ Y_e\end{bmatrix}$ to $\begin{bmatrix}\hat{Y}_o\\ \hat{Y}_e\end{bmatrix}$ is equal to 1, yielding $f_{\hat{\boldsymbol{Y}}|\boldsymbol{X}}(\hat{Y}|X) = f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$. Substituting for T, Σ^{-1} and $\hat{Y}$ in (69) yields:

$$f_{\hat{\boldsymbol{Y}}|\boldsymbol{X}}(\hat{Y}|X) = \frac{1}{\pi^4\sigma_g^4\sigma_v^4|\Sigma|^2}\exp\big(-\lambda_1\big(|\hat{y}_1|^2+|\hat{y}_2|^2\big)-\lambda_2\big(|\hat{y}_3|^2+|\hat{y}_4|^2\big)\big) \times\int_0^1\int_0^1 \frac{t\ (\ \bar{t}\ )^{-1}}{\prod_{i=1}^2(\lambda_i t+\bar{t})} I_0\left(2\sqrt{u\bar{u}t}\left|\sum_{i=1}^2\frac{\lambda_i^2}{\lambda_i t+\bar{t}}\,\hat{y}_{2i-1}\hat{y}_{2i}\right|\right) \times\exp\left(\frac{\lambda_1^2 t}{\lambda_1 t+\bar{t}}\big[|\hat{y}_1|^2\bar{u}+|\hat{y}_2|^2u\big]+\frac{\lambda_2^2 t}{\lambda_2 t+\bar{t}}\big[|\hat{y}_3|^2\bar{u}+|\hat{y}_4|^2u\big]\right) \times\exp\left(-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right)du\ dt. \quad (70)$$

Using the series expansion of the Bessel function $I_0(\cdot)$ in (70) yields:

$$f_{\hat{\boldsymbol{Y}}|\boldsymbol{X}}(\hat{Y}|X) = \frac{1}{\pi^4\sigma_g^4\sigma_v^4|\Sigma|^2}\exp\left(-\lambda_1\left(|\hat{y}_1|^2+|\hat{y}_2|^2\right)-\lambda_2\left(|\hat{y}_3|^2+|\hat{y}_4|^2\right)\right) \times\int_0^1\int_0^1\frac{t\ (\ \bar{t}\ )^{-1}}{\prod_{i=1}^2(\lambda_i t+\bar{t})}\sum_{k=0}^\infty\frac{1}{(k!)^2}\left(\left|\sum_{i=1}^2\frac{\lambda_i^2\sqrt{u\bar{u}t^2}y_{2i-1}y_{2i}}{\lambda_i t+\bar{t}}\right|^{2k}\right) \times\exp\left(\frac{\lambda_1^2 t}{\lambda_1 t+\bar{t}}\left[|\hat{y}_1|^2\bar{u}+|\hat{y}_2|^2u\right]+\frac{\lambda_2^2 t}{\lambda_2 t+\bar{t}}\left[|\hat{y}_3|^2\bar{u}+|\hat{y}_4|^2u\right]\right) \times\exp\left(-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right)dudt. \quad (71)$$

It can be shown that (71) is equivalent to:

$$f_{\hat{\boldsymbol{Y}}|\boldsymbol{X}}(\hat{Y}|X) = \frac{1}{\pi^4\sigma_g^4\sigma_v^4|\Sigma|^2}\exp\left(-\lambda_1\left(|\hat{y}_1|^2+|\hat{y}_2|^2\right)-\lambda_2\left(|\hat{y}_3|^2+|\hat{y}_4|^2\right)\right) \times\int_0^1\int_0^1\frac{t\ (\ \bar{t}\ )^{-1}}{\prod_{i=1}^2(\lambda_i t+\bar{t})}\exp\left(-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right)S(\hat{Y},u,t) \times\exp\left(\frac{\lambda_1^2 t}{\lambda_1 t+\bar{t}}\left[|\hat{y}_1|^2\bar{u}+|\hat{y}_2|^2u\right]+\frac{\lambda_2^2 t}{\lambda_2 t+\bar{t}}\left[|\hat{y}_3|^2\bar{u}+|\hat{y}_4|^2u\right]\right) \times dudt, \quad (72)$$

where

$$S(\hat{Y},t,u) = \sum_{k=0}^\infty\frac{1}{(k!)^2}\big((u\bar{u}t^2)^k\big(\alpha_1^2|\hat{y}_1|^2|\hat{y}_2|^2+\alpha_2^2|\hat{y}_3|^2|\hat{y}_4|^2+ 2\alpha_1\alpha_2|\hat{y}_1||\hat{y}_2||\hat{y}_3||\hat{y}_4|\cos(\theta_1-\theta_2-\theta_3+)\big)^k\big), \quad (73)$$

and

$$\alpha_i = \frac{\lambda_i^2}{\lambda_i t+\bar{t}}, \qquad i=1,2. \quad (74)$$

Using the trinomial expansion:

$$(a+b+c)^k = \sum_{m=0}^k\sum_{n=0}^{k-m}\frac{k!\ a^n\ b^m\ c^{k-n-m}}{m!n!(k-m-n)!}, \quad (75)$$

where $a = \alpha_1^2|\hat{y}_1|^2|\hat{y}_2|^2$, $b = 2\alpha_1\alpha_2|\hat{y}_1||\hat{y}_2||\hat{y}_3||\hat{y}_4|\cos(\theta_1-\theta_2-\theta_3+)$, and $c = \alpha_2^2|\hat{y}_3|^2|\hat{y}_4|^2$ yields:

$$S(\hat{Y},t,u) = \sum_{k=0}^\infty\frac{(u\bar{u}t^2)^k}{k!} \quad (76)$$

$$\times\sum_{m=0}^k\frac{(2\alpha_1\alpha_2|\hat{y}_1||\hat{y}_2||\hat{y}_3||\hat{y}_4|\cos(\theta_1-\theta_2-\theta_3+\theta_4))^m}{m!} \quad (77)$$

$$\times\sum_{n=0}^{k-m}\frac{\left(\alpha_1^2|\hat{y}_1|^2|\hat{y}_2|^2\right)^n\left(\alpha_2^2|\hat{y}_3|^2|\hat{y}_4|^2\right)^{k-n-m}}{n!(k-m-n)!}. \quad (78)$$

Using polar coordinates with $r_i = \sqrt{\Re\{\hat{y}_i\}+\Im\{\hat{y}_i\}}$, and $\theta_i = \arctan\left(\frac{\Im\{\hat{y}_i\}}{\Re\{\hat{y}_i\}}\right)$, $i=1,\dots,4$, in (71) yields:

$$f_{\boldsymbol{r},\boldsymbol{\theta}|\boldsymbol{X}}(r,\theta|X) = \frac{r_1r_2r_3r_4}{\pi^4\sigma_g^4\sigma_v^4|\Sigma|^2} \times\exp\left(-\lambda_1\left(r_1^2+r_2^2\right)-\lambda_2\left(r_3^2+r_4^2\right)\right)\int_0^1\int_0^1\frac{t\ (\ \bar{t}\ )^{-1}}{\prod_{i=1}^2(\lambda_i t+\bar{t})} \times\exp\Big(\alpha_1 t\left[r_1^2\bar{u}+r_2^2u\right]+\alpha_2 t\left[r_3^2\bar{u}+r_4^2u\right]-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\Big) \times\sum_{k=0}^\infty\sum_{m=0}^k\frac{(u\bar{u}t^2)^k(2\alpha_1\alpha_2r_1r_2r_3r_4)^m\cos^m(\theta_1-\theta_2-\theta_3+\theta_4)}{k!\ m!} \times\sum_{n=0}^{k-m}\frac{\left(\alpha_1^2r_1^2r_2^2\right)^n\left(\alpha_2^2r_3^2r_4^2\right)^{k-n-m}}{n!(k-m-n)!}dudt, \quad (79)$$

where $r = \begin{bmatrix}r_1 & r_2 & r_3 & r_4\end{bmatrix}^\dagger$ and $\theta = \begin{bmatrix}\theta_1 & \theta_2 & \theta_3 & \theta_4\end{bmatrix}^\dagger$. To obtain $f_{\boldsymbol{r}|\boldsymbol{X}}(r)$, we integrate (79) over all possible values of $\boldsymbol{\theta}$. We notice that [40]:

1) The integration over the period $[0,2\pi)$ of any odd power of $\cos(\beta)$ is zero, i.e.,
$$\int_0^{2\pi}\cos^m(\beta)d\beta = 0, \qquad \text{for } m \text{ odd}. \quad (80)$$
2) The integration over the period $[0,2\pi)$ of any even power of $\cos(\beta)$ is given by:
$$\int_0^{2\pi}\cos^m(\beta)d\beta = \frac{1}{2^m}\binom{m}{m/2}(2\pi). \quad (81)$$

Using these facts, it suffices to consider $m = 2\ell$. Using this and substituting for the values of α_1 and α_2 from (74) in (79) yields:

$$f_{\boldsymbol{r_1},\boldsymbol{r_2},\boldsymbol{r_3},\boldsymbol{r_4}|\boldsymbol{X}}(r_1,r_2,r_3,r_4|X) = \frac{1}{\sigma_g^4\sigma_v^4|\Sigma|^2} \times\sum_{k=0}^\infty\sum_{\ell=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{n=0}^{k-2\ell}\frac{1}{k!\ \ell!^2\ n!\ (k-2\ell-n)!}\int_0^1\int_0^1\frac{(u\bar{u}t^2)^k t(\bar{t})^{-1}}{\prod_{i=1}^2(\lambda_i t+\bar{t})} \times\exp\left(-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right)(2r_1)(r_1^2)^{n+\ell}\exp\left(-\left(\frac{\lambda_1^2tu+\lambda_1\bar{t}}{\lambda_1 t+\bar{t}}\right)r_1^2\right)$$

$$\times (2r_2)(r_2^2)^{n+\ell}\exp\left(-\left(\frac{\lambda_1^2 t\bar{u}+\lambda_1\bar{t}}{\lambda_1 t+\bar{t}}\right)r_2^2\right)$$
$$\times (2r_3)(r_3^2)^{k-n-\ell}\exp\left(-\left(\frac{\lambda_2^2 tu+\lambda_2\bar{t}}{\lambda_2 t+\bar{t}}\right)r_3^2\right)$$
$$\times (2r_4)(r_4^2)^{k-n-\ell}\exp\left(-\left(\frac{\lambda_2^2 t\bar{u}+\lambda_2\bar{t}}{\lambda_2 t+\bar{t}}\right)r_4^2\right)$$
$$\times\left(\frac{\lambda_1^2(\lambda_2 t+\bar{t})}{\lambda_2^2(\lambda_1 t+\bar{t})}\right)^{2(n+\ell)}\left(\frac{\lambda_2^2}{\lambda_2 t+\bar{t}}\right)^{2k} dudt. \quad (82)$$

To obtain $f_{\boldsymbol{r_1}|\boldsymbol{X}}(r_1|X)$, we proceed as follows:

$$f_{\boldsymbol{r_1}|\boldsymbol{X}}(r_1|X) = \int_0^\infty\int_0^\infty\int_0^\infty f_{\boldsymbol{r_1},\boldsymbol{r_2},\boldsymbol{r_3},\boldsymbol{r_4}|\boldsymbol{X}}(r_1,r_2,r_3,r_4|X)dr_2dr_3dr_4,$$
$$= \frac{2r_1}{\sigma_g^4\sigma_v^4|\Sigma|^2}\sum_{k=0}^{\infty}\sum_{\ell=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{n=0}^{k-2\ell}\frac{1}{k!\ \ell!^2\ n!\ (k-2\ell-n)!}$$
$$\times\int_0^1\int_0^1\frac{(u\bar{u}t^2)^k t(\bar{t})^{-1}}{\prod_{i=1}^2(\lambda_i t+\bar{t})}\exp\left(-\frac{t}{\sigma_g^2\sigma_v^2\bar{t}}\right)(r_1^2)^{n+\ell}$$
$$\times\exp\left(-\left(\frac{\lambda_1^2 tu+\lambda_1\bar{t}}{\lambda_1 t+\bar{t}}\right)r_1^2\right)$$
$$\times\int_0^\infty(r_2^2)^{n+\ell}\exp\left(-\left(\frac{\lambda_1^2 t\bar{u}+\lambda_1\bar{t}}{\lambda_1 t+\bar{t}}\right)r_2^2\right)2r_2dr_2$$
$$\times\int_0^\infty(r_3^2)^{k-n-\ell}\exp\left(-\left(\frac{\lambda_2^2 tu+\lambda_2\bar{t}}{\lambda_2 t+\bar{t}}\right)r_3^2\right)2r_3dr_3$$
$$\times\int_0^\infty(r_4^2)^{k-n-\ell}\exp\left(-\left(\frac{\lambda_2^2 t\bar{u}+\lambda_2\bar{t}}{\lambda_2 t+\bar{t}}r_4^2\right)\right)2r_4dr_4$$
$$\times\left(\frac{\lambda_1^2(\lambda_2 t+\bar{t})}{\lambda_2^2(\lambda_1 t+\bar{t})}\right)^{2(n+\ell)}\left(\frac{\lambda_2^2}{\lambda_2 t+\bar{t}}\right)^{2k}dudt.$$

Setting $r_i^2 = u, 2r_i dr_i = du$, $i\in\{2,3,4\}$, and using the fact that:

$$\int_0^\infty u^q\exp(-\gamma u)\,du = \frac{q!}{\gamma^{q+1}}, \quad (83)$$

where $\gamma>0$ and q is a positive integer, yields (17).

Appendix D
Proof of Lemma 3

Let $Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$ be the likelihood function based on the Gaussian approximation for which the covariance matrix is to be optimized. For the cases in which $k_g=0$, it is shown in Eq. (14) on page 5 of the revised manuscript that

$$Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \frac{1}{\pi^{2T}|\Xi|^2}\exp\left(-\left(Y_o^\dagger\Xi^{-1}Y_o+Y_e^\dagger\Xi^{-1}Y_e\right)\right), \quad (84)$$

where Ξ is the covariance matrix to be optimized.

For a given X, the Kullback-Leibler distance measure between the exact likelihood function, $f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$, and the likelihood function based on the Gaussian approximation, $Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$, is given by

$$D\big(f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)||Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\big) = \int_Y f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\log\left(\frac{f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)}{Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)}\right)dY, \quad (85)$$

i.e.,

$$D\big(f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)||Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\big) = \int_Y f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\log\big(f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\big)dY - \int_Y f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\log\big(Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\big)dY. \quad (86)$$

Our goal is to find an expression for

$$\Xi^\star(X) = \arg\min_{\Xi\succ 0} D\big(f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)||Q_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\big). \quad (87)$$

To do so, we note that the first term on the right hand side of (86) is independent of Ξ. Hence, to achieve our goal, it suffices to solve the following optimization problem:

$$\max_{\Xi\succ 0}\int_Y\big(-Y_o^\dagger\Xi^{-1}Y_o - Y_e^\dagger\Xi^{-1}Y_e - 2\log(|\Xi|)+2T\log(\pi)\big) \times f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)dY. \quad (88)$$

To solve this optimization problem, we let $R=\Xi^{-1}$, and note that because $\Xi\succ 0$, $R\succ 0$. Using this change of variables, we rewrite (88) as:

$$\max_{R\succ 0}\int_Y\left(2\log(|R|)-\left(Y_o^\dagger RY_o+Y_e^\dagger RY_e\right)\right)f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)dY. \quad (89)$$

The function $F(R)=2\log(|R|)-Y_o^\dagger RY_o-Y_e^\dagger RY_e$ is concave. Hence, differentiating (89) with respect to R and solving for R that makes the derivative equal to zero yields the optimal covariance for the given X, $R^\star(X)$, provided that the resulting $R^\star(X)\succ 0$. Differentiating (89) yields

$$2\big(R^\star(X)\big)^{-1} = \int_Y f_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)\left(Y_oY_o^\dagger+Y_eY_e^\dagger\right)dY = \int_{Y_o} f_{\boldsymbol{Y}_o|\boldsymbol{X}}(Y_o|X)\big(Y_oY_o^\dagger\big)dY_o + \int_{Y_e} f_{\boldsymbol{Y}_e|\boldsymbol{X}}(Y_e|X)\big(Y_eY_e^\dagger\big)dY_e. \quad (90)$$

To complete the proof, it remains to show that $R^\star(X)\succ 0$, but this can be readily seen from (90), since its right hand side is the statistical average of the sum of two rank-1 positive semidefinite matrices. Hence, $\big(R^\star(X)\big)^{-1}\succeq 0$. Since the noise components of the entries of Y are statistically independent of each other and of other components of Y, it can be readily seen that $|R^\star(X)|>0$, whereby $R^\star(X)\succ 0$, yielding the statement of the lemma.

References

[1] K. Pelechrinis, M. Iliofotou, and S. V. Krishnamurthy, "Denial of Service Attacks in Wireless Networks: The Case of Jammers," *IEEE Commun. Surv. Tutorial*, vol. 13, pp. 245–257, Feb. 2011.

[2] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *IEEE Proc.*, vol. 104, pp. 1727–1765, Sept. 2016.

[3] R. Krenz and S. Brahma, "Jamming LTE signals," in *IEEE Int. Black Sea Conf. Commun. and Ntwrking (BlackSeaCom)*, pp. 72–76, May 2015.

[4] R. Kumari and M. Mukhopadhyay, "Design of GPS antijamming algorithm using adaptive array antenna to mitigate the noise and interference," *Ambient Intelligence Humanized Computing*, Feb. 2019.

[5] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, "A survey on jamming attacks and countermeasures in WSNs," *IEEE Commun. Surv. Tutorial*, vol. 11, pp. 42–56, Apr. 2009.

[6] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, and W. Zhuang, "UAV Relay in VANETs Against Smart Jamming With Reinforcement Learning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4087–4097, 2018.

[7] R. D. Pietro and G. Oligeri, "Jamming mitigation in cognitive radio networks," *IEEE Network*, vol. 27, pp. 10–15, May 2013.

[8] M. Labib, S. Ha, W. Saad, and J. H. Reed, "A Colonel Blotto game for anti-jamming in the Internet of Things," in *Proc. IEEE Glob. Commun. Conf.*, (San Diego), pp. 1–6, 2015.

[9] T. Basar, "The Gaussian test channel with an intelligent jammer," *IEEE Trans. Inf. Theory*, vol. 29, pp. 152–157, Jan. 1983.

[10] J. Matuszewski, "Jamming efficiency of land-based radars by the airborne jammers," in *Int. Microwave Radar Conf. (MIKON)*, pp. 324–327, May 2018.

[11] M. Wilhelm, I. Martinovic, J. B. Schmitt, and V. Lenders, "Short paper: Reactive jamming in wireless networks: How realistic is the threat?," in *Proc. 4th ACM Conf. Wireless Ntwrk Security*, (New York), pp. 47–52, 2011.

[12] K. Grover, A. Lim, and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: A survey," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 17, pp. 197–215, Dec. 2014.

[13] T. C. Clancy, "Efficient OFDM denial: Pilot jamming and pilot nulling," in *Proc. IEEE Int. Conf. Commun. (ICC)*, (Kyoto), June 2011.

[14] M. Lichtman, J. D. Poston, S. Amuru, C. Shahriar, T. C. Clancy, R. M. Buehrer, and J. H. Reed, "A communications jamming taxonomy," *IEEE Security Privacy*, vol. 14, pp. 47–54, Jan. 2016.

[15] K. Luo, R. H. Gohary, and H. Yanikomeroglu, "The capacity of a broadcast channel with Gaussian jamming and a friendly eavesdropper," in *Proc. IEEE Inform. Theory Wkshp. (ITW)*, pp. 327–331, Oct. 2015.

[16] R. H. Gohary, Y. Huang, Z. Luo, and J. Pang, "A generalized iterative water-filling algorithm for distributed power control in the presence of a jammer," *IEEE Trans. Signal Processing*, vol. 57, pp. 2660–2674, July 2009.

[17] K. Pelechrinis, I. Broustis, S. V. Krishnamurthy, and C. Gkantsidis, "ARES: An anti-jamming reinforcement system for 802.11 networks," in *Proc. Int. Conf. Emerg. Ntwrk. Experiments and Technologies*, (Rome, Italy), 2009.

[18] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Game theoretic anti-jamming dynamic frequency hopping and rate adaptation in wireless systems," in *Int. Symp. Model. Opt. Mobile, Ad Hoc, and Wireless Ntwrks (WiOpt)*, (Hammamet, Tunisia), pp. 247–254, May 2014.

[19] M. Medard, "Capacity of correlated jamming channels," in *Proc. Allerton Conf. Commun, Control, Comput.*, pp. 1043—-1052, Oct. 1997.

[20] A. Bayesteh, M. Ansari, and A. K. Khandani, "Effect of jamming on the capacity of MIMO channels," in *Proc. Allerton Conf. Commun, Control, Comput.*, pp. 401–410, Oct. 2004.

[21] S. Shafiee and S. Ulukus, "Capacity of multiple access channels with correlated jamming," in *Proc. IEEE Military Commun. (MILCOM)*, pp. 218–224, Oct. 2005.

[22] C. A. Cole, C. Shahriar, and T. C. Clancy, "An anti-jam communications technique via spatial hiding precoding," in *Proc. IEEE Military Commun. (MILCOM)*, pp. 490–494, Oct. 2014.

[23] C. C. Craig, "On the frequency function of xy," *Ann. Math. Statist.*, vol. 7, pp. 1–15, Mar. 1936.

[24] K. S. Miller, *Multidimensional Gaussian distributions*. New York: John Wiley and Sons, 1964.

[25] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers, Scientists and Mathematicians*. Berlin: Springer-Verlag, 2006.

[26] R. K. Mallik and N. C. Sagias, "Distribution of inner product of complex Gaussian random vectors and its applications," *IEEE Trans. Commun.*, vol. 59, pp. 3353–3362, Dec. 2011.

[27] N. O'Donoughue and J. M. F. Moura, "On the product of independent complex Gaussians," *IEEE Trans. Signal Processing*, vol. 60, pp. 1050–1063, Mar. 2012.

[28] S. Guruacharya, B. K. Chalise, and B. Himed, "On the Product of Complex Gaussians With Applications to Radar," *IEEE Signal Processing Letters*, vol. 26, no. 10, pp. 1536–1540, 2019.

[29] Y. Li, Q. He, and R. S. Blum, "On the Product of Two Correlated Complex Gaussian Random Variables," *IEEE Signal Processing Letters*, vol. 27, pp. 16–20, 2020.

[30] T. Betlehem and A. J. Coulson, "Distribution of the sum of a complex Gaussian and the product of two complex Gaussians," in *Australian Commun. Theory Wkshp (AusCTW)*, (Wellington), pp. 126–129, Jan. 2012.

[31] K. A. Almahorg and R. H. Gohary, "Maximum likelihood detection in the presence of non-Gaussian jamming," *IEEE Trans. Signal Processing*, vol. 68, pp. 5722–5735, 2020.

[32] W. Shi, Y. Li, and Q. He, "Distribution of the Product of a Complex Gaussian Matrix and Vector and Its Sum with a Complex Gaussian Vector," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 6014–6018, 2020.

[33] S. Guruacharya, B. K. Chalise, and B. Himed, "Energy Distribution of Multiple Target Signal With Application to Target Counting," *IEEE Signal Processing Letters*, vol. 27, pp. 431–435, 2020.

[34] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill, 2008.

[35] W. Khawaja, I. Guvenc, D. W. Matolak, U. Fiebig, and N. Schneckenburger, "A Survey of Air-to-Ground Propagation Channel Modeling for Unmanned Aerial Vehicles," *IEEE Commun. Surv. Tutorial*, vol. 21, no. 3, pp. 2361–2391, 2019.

[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[37] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1962–1973, Sept. 2000.

[38] E. Hall and S. Wilson, "Design and analysis of turbo codes on Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 160–174, 1998.

[39] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, "Finite-length scaling for iteratively decoded LDPC ensembles," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 473–498, 2009.

[40] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products*. Academic Press, 2007.