The Worst Jamming Distribution for Securely Precoded OFDM

Yuan Liang Jian Ren Tongtong Li

Department of Electrical & Computer Engineering, Michigan State University
Email: {liangy11, renjian, tongli}@egr.msu.edu

Abstract—**In this paper, we address the problem of finding the worst jamming distribution in terms of channel capacity for the securely precoded OFDM (SP-OFDM) system, so as to evaluate the performance of SP-OFDM under destructive hostile jamming. We consider a practical communication scenario, where the transmitting symbols are uniformly distributed over a discrete and finite alphabet, and the jamming interference is subject to both average and peak power constraints. Using tools in functional analysis and complex analysis, first, we show the existence and uniqueness of the worst jamming distribution; then we further prove that the worst jamming distribution is discrete in amplitude with a finite number of mass points. Numerical examples are provided under different scenarios to demonstrate our theoretical results.**

***Index Terms*—hostile jamming, secure precoding, phase randomization, arbitrarily varying channel, channel capacity.**

I. Introduction

In wireless systems, one of the most commonly used techniques for limiting the effectiveness of an opponent's communication is referred to as jamming, in which the authorized user's signal is deliberately interfered by the adversary. Recently, it was found that disguised jamming [1]–[4], where the jamming is highly correlated with the signal, can reduce the system capacity to zero. Consider the following example:

$$R = S + J + N$$

where S is the authorized signal, J is the jamming interference, N is the noise independent of J and S, and R is the received signal. If the jammer is capable of eavesdropping on the symbol constellation and the codebook of the transmitter, it can simply replicate one of the sequences in the codebook of the legitimate transmitter, the receiver, then, would not be able to distinguish between the authorized sequence and the jamming sequence, resulting in a complete communication failure [5, ch 7.3].

To design the corresponding anti-jamming system against disguised jamming, the main task is to break the symmetricity between the authorized signal and the jamming interference, or make it impossible for the jammer to achieve this symmetricity [6]. In [7], we proposed a securely precoded OFDM (SP-OFDM) system by introducing a dynamic constellation through secure phase randomization. Using the arbitrarily varying channel (AVC) model, we showed that the AVC channel corresponding to SP-OFDM is not symmetrizable under disguised jamming, and hence SP-OFDM can achieve a positive deterministic coding capacity.

The channel capacity of SP-OFDM under hostile jamming is calculated from a maximin optimization problem, where the legitimate transmitter aims to maximize the mutual information (MI) between the transmitted signal and received signal, while the jammer aims to minimize it. In this paper, we focus on finding the worst jamming distribution which minimizes the channel capacity of SP-OFDM. The worst jamming distribution problem was also discussed in [8], where it was assumed that the receiver was *not* aware of the existence of jamming, and the jamming distribution was optimized to maximize the error rate. Unlike [8], in this paper, the SP-OFDM system is designed to have inherent anti-jamming feature, and we want to find the worst jamming for SP-OPDM and evaluate the channel capacity under it. We consider a practical wireless communication scenario, where the transmitting symbol is uniformly distributed in a discrete and finite alphabet, and the jamming interference is subject to finite average and peak power constraints.

By formulating the problem as a constrained functional optimization process [9], we derive the Kuhn-Tucker (KT) conditions that should be satisfied by the worst jamming distribution, from which we compute the worst jamming distribution numerically. The main conclusion of this paper is that: the worst jamming distribution should be discrete in amplitude with a finite number of mass points. In literature, constrained functional optimization was originally applied in seeking the capacity-achieving input distributions for different channels with different constraints, e.g., [10]–[13]. One common result of these work is that, the capacity-achieving input distributions for these channels are discretely distributed in amplitude, which is consistent with the result in this paper.

The rest of this paper is organized as follows. We briefly revisit the channel model of SP-OFDM and formulate the problem in Section II. The analysis on the worst jamming distribution is presented in Section III. Numerical results are provided in Section IV and we conclude in Section V.

II. Problem Formulation

In this section, we introduce the channel model we study and formulate the worst jamming distribution problem .

This work was partially supported by the National Science Foundation under Grants CNS-1217206 and ECCS-1744604.

978-1-5386-4727-1/18/$31.00 ©2018 IEEE

Consider the AWGN channel under hostile jamming

$$R = S + J + N, \tag{1}$$

where $S \in \Phi, J \in \mathbb{C}$, $N \sim \mathcal{CN}(0, \sigma^2)$, Φ is the constellation of the authorized signal, and $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a circularly symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$. Using the arbitrarily varying channel (AVC) model [14] [15], it can be shown that, given: (i) the constellation of the authorized signal, denoted by Φ, is fixed, (ii) S is uniformly distributed over Φ and (iii) no secure symbol-level precoding is involved, the corresponding AVC channel of (1) is symmetrizable under disguised jamming, resulting in zero deterministic coding capacity.

In [7], we proposed a securely precoded OFDM (SP-OFDM) system under disguised jamming, where the symbol-level precoding is achieved through secure phase randomization, whose equivalent channel model is

$$R = S + e^{j\Theta} J + N, \tag{2}$$

where Θ is a random phase shift controlled by the authorized user, uniformly distributed over $\{\frac{2\pi i}{M} \mid i = 0, 1, ..., M-1\}$ and M is a constant integer. The random phase shift Θ is encrypted such that it is not accessible to the jammer. In [7], we analyzed the SP-OFDM system under AVC model and showed that the channel (2) is not l-symmetrizable [15] with any finite average jamming power constraint, hence a positive channel capacity can be guaranteed. The result is summarized in the following lemma.

Lemma 1. *The AVC channel corresponding to SP-OFDM is not l-symmetrizable under any finite average power constraint, thus the deterministic coding capacity of SP-OFDM is positive. More specifically, given an average jamming power constraint P_J, the channel capacity C in this case equals*

$$\begin{aligned} C = \max_{\mathcal{P}_S} \min_{F_J} & \; I(S, R), \\ s.t. \quad & \int_{\mathbb{C}} |\boldsymbol{x}|^2 dF_J(\boldsymbol{x}) \le P_J, \end{aligned} \tag{3}$$

where $I(S, R)$ denotes the mutual information (MI) between R and S in (2), $\mathcal{P}_S$ denotes the probability distribution of S over Φ and $F_J(\cdot)$ denotes the CDF of J.

In this paper, we consider the maximin problem (3) in practical communication systems, where the transmitted symbol is uniformly distributed in the alphabet and the jamming interference is subject to a finite peak power constraint, denoted by a.

Let the transmitting alphabet $\Phi = \{\boldsymbol{s}_1, \boldsymbol{s}_2, ..., \boldsymbol{s}_{M_\Phi}\}$. The size of Φ is $|\Phi| = M_\Phi$. In realistic communication systems, S is uniformly distributed over Φ , i.e.,

$$\Pr\{S = \boldsymbol{s}_i\} = \frac{1}{M_\Phi}, \quad i \in \{1, 2, ..., M_\Phi\}. \tag{4}$$

Hence, the calculation of the capacity in (3) is reduced to finding the CDF F_J of the jamming signal that minimizes the MI $I(S, R)$.

The phase shift Θ is controlled by the authorized user, and is uniformly distributed over $\{\frac{2\pi m}{M} \mid m = 0, 1, ..., M-1\}$. When M is sufficiently large (e.g., $M \gg \frac{2\pi}{\sigma_\Theta}$, where σ_Θ^2 is the variance of phase noise existing in practical communication systems), taking the noise effect into consideration, we can approximately model Θ as a continuous RV uniformly distributed over $[0, 2\pi)$, and independent of J. In this way, the phase item within J can be completely absorbed into Θ. As a result, we only need to find the CDF of $|J|$ that minimizes the MI in (3). Without loss of generality, in the following, J is degraded to a RV over $\mathbb{R}$.

Let $\mathcal{F}_a$ be the set of all the possible CDFs $F_J(\cdot)$ of $J \in \mathbb{R}$ satisfying the peak-power constraint a, that is,

$$\begin{aligned} \mathcal{F}_a \quad = \quad & \{F(\cdot) \mid F(0^-) = 0, F(a) = 1, \\ & F(x) \text{ is nondecreasing and right-continuous}\}, \end{aligned}$$

and it is a convex set. Note that the MI $I(S, R) = H(S) - H(S|R)$ and S is uniformly distributed over Φ, so $I(S, R)$ is only a functional of $F_J(\cdot)$. Define functional $G(\cdot)$ as

$$G(F_J) \triangleq -H(S|R). \tag{5}$$

The maximin problem (3) can be reformulated as

$$\begin{aligned} \min_{F_J(\cdot) \in \mathcal{F}_a} & \; G(F_J) \\ s.t. \quad & \int_0^\infty x^2 \mathrm{d}F_J(x) \le P_J \end{aligned} \tag{6}$$

The optimal solution to (6) is the worst jamming distribution that results in the minimal capacity from the information theoretic point of view. In the following, we will discuss the property of the worst jamming distribution F_J to (6) following the approaches in [10], [11], [16]. The main result is that the worst jamming distribution should be discrete in amplitude with a finite number of mass points.

III. The Worst Jamming Distribution

In this section, we first verify the existence of the worst jamming distribution. Then we show that the worst F_J that minimizes $I(S, R)$ is discretely distributed over $[0, \infty)$ with a finite number of mass points.

A. The Existence of the Worst Jamming Distribution

In this subsection, we will prove the existence of the worst jamming distribution subject to the power constraints. First we derive the specific expression of $G(F_J)$ given $F_J \in \mathcal{F}_a$. Define function $u(x, y)$ as

$$u(x, y) \triangleq \frac{1}{\pi\sigma^2} e^{-\frac{x^2+y^2}{\sigma^2}} I_0(\frac{2xy}{\sigma^2}), \; x \ge 0, y \ge 0, \tag{7}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind with order 0. Then, for a given noise power σ^2, the conditional PDF of R given S and J can be calculated as

$$f_{R|S,J}(\boldsymbol{r} \mid \boldsymbol{s}_i, x) = u(|\boldsymbol{r} - \boldsymbol{s}_i|, x), \boldsymbol{r} \in \mathbb{C}, \boldsymbol{s}_i \in \Phi, x \ge 0. \tag{8}$$

The PDF of R and the posterior distribution of S are

$$f_R(\boldsymbol{r}) = \frac{\sum_i f_{R|S}(\boldsymbol{r} \mid \boldsymbol{s}_i)}{M_\Phi} = \frac{\sum_i \int_0^\infty u(|\boldsymbol{r} - \boldsymbol{s}_i|, x) \mathrm{d}F_J(x)}{M_\Phi}, \tag{9}$$

$$\Pr\{S = \boldsymbol{s}_i \mid R = \boldsymbol{r}\} = \frac{f_{R|S}(\boldsymbol{r} \mid \boldsymbol{s}_i)\Pr\{S = \boldsymbol{s}_i\}}{f_R(\boldsymbol{r})} = \frac{\int_0^\infty u(|\boldsymbol{r} - \boldsymbol{s}_i|, x)\mathrm{d}F_J(x)}{M_\Phi \cdot f_R(\boldsymbol{r})}. \quad (10)$$

Define functionals $Q_i(\boldsymbol{r}, F)$ and $L_i(\boldsymbol{r}, F)$ over $\mathbb{C} \times \mathcal{F}_a$ as

$$Q_i(\boldsymbol{r}, F) = \int_0^\infty u(|\boldsymbol{r} - \boldsymbol{s}_i|, x)\mathrm{d}F(x), \ \boldsymbol{r} \in \mathbb{C}, F \in \mathcal{F}_a. \quad (11)$$

$$L_i(\boldsymbol{r}, F) = \log\left(\frac{Q_i(\boldsymbol{r}, F)}{\sum_k Q_k(\boldsymbol{r}, F)}\right), \ \boldsymbol{r} \in \mathbb{C}, F \in \mathcal{F}_a. \quad (12)$$

Therefore $G(F_J)$ can be expressed as

$$G(F_J) = -H(S|R) = \frac{1}{M_\Phi}\sum_i \int_{\mathbb{C}} Q_i(\boldsymbol{r}, F_J)L_i(\boldsymbol{r}, F_J)\mathrm{d}\boldsymbol{r}. \quad (13)$$

Moreover, $I(S, R)$ is a convex function w.r.t. $F_J(\cdot)$, so is $G(\cdot)$, that is, $G((1-\lambda)F_1 + \lambda F_2) \leq (1-\lambda)G(F_1) + \lambda G(F_2)$ for any $\lambda \in [0, 1]$ and $F_1, F_2 \in \mathcal{F}_a$.

For any $P \geq 0$, Define functional $K_P(\cdot)$ as:

$$K_P(F) \triangleq \int_0^\infty x^2 \mathrm{d}F(x) - P. \quad (14)$$

Let $P = P_J$, the average jamming power constraint, and define

$$\Omega \triangleq \{F(\cdot) \mid F(\cdot) \in \mathcal{F}_a, K_{P_J}(F) \leq 0\}. \quad (15)$$

The optimization problem (6) can be expressed equivalently as

$$\min_{F \in \Omega} G(F). \quad (16)$$

In the following, the subscript J in the distribution function and the average power constraint is discarded for brevity. Next, we will show that functional $G(F)$ can achieve its minimum on Ω. More specifically, we have:

Theorem 1. *Given $\Omega = \{F(\cdot) \mid F(\cdot) \in \mathcal{F}_a, K_P(F) \leq 0\}$, where P is the average jamming power constraint, and $G(F) = -H(S \mid R)$ on set Ω. The real-valued functional $G(\cdot)$ can achieve its minimum on Ω.*

The proof of Theorem 1 is based on the following Lemma, which makes use of the weak* topology on the set of CDFs over $\mathbb{R}$.

Lemma 2. *[11] If f is a real-valued, weak* continuous functional [1] [9] on a weak* compact set [2] [9] Ω, then G achieves its minimum on Ω.*

Proof. According to Lemma 2, it is sufficient to prove that Ω is weak* compact and $G(\cdot)$ is weak* continuous. The weak* compactness of Ω was proved in [11]. The weak* continuity of $G(\cdot)$ can be proved by the Lebesgue dominated convergence theorem [12]. □

Given the existence of the worst jamming distribution over Ω, the worst jamming distribution is also unique.

Corollary 1. *The jamming distribution $F^* \in \Omega$ that minimizes functional $G(\cdot)$ is unique.*

Proof. By assuming two jamming distributions $F_0^*, F_1^* \in \Omega$ that can minimize $G(\cdot)$ equivalently, it can be proved that F_0^*, F_1^* render the same conditional distribution of R given S following [17, Theorem 2.6.3]. Then, by making use of the characteristic functions of R and J, it follows that F_0^* should be equivalent to F_1^*, which proves the uniqueness of the worst jamming distribution. □

As we are solving a constrained optimization problem, using the Kuhn-Tucker Theorem [9] [11], the following theorem can be obtained.

Theorem 2. *There exists a unique solution to the constrained convex optimization problem (6) on Ω. If $F_0(\cdot) \in \Omega$ is the solution to (6), then there exists a $\gamma \geq 0$ such that $F_0(\cdot)$ is also the solution to the Lagrangian dual problem*

$$\min_{F \in \mathcal{F}_a} G(F) + \gamma K_P(F) \quad (17)$$

and $\gamma K_P(F_0) = 0$.

Next we analyze the necessary and sufficient conditions for the minima in the dual problem. Using the definition of weak differentiability [18] in functional analysis, the following result can be obtained.

Theorem 3. *Define functional $g(\cdot\, ;\, \cdot)$ over $[0, \infty) \times \mathcal{F}_a$ as*

$$g(x; F) \triangleq \frac{\sum_i \int_{\mathbb{C}} u(|\boldsymbol{r} - \boldsymbol{s}_i|, x)L_i(\boldsymbol{r}, F)d\boldsymbol{r}}{M_\Phi}, \ \ x \geq 0, F \in \mathcal{F}_a. \quad (18)$$

A necessary and sufficient condition for $F_0(\cdot) \in \mathcal{F}_a$ to be the minima of the dual problem (17) is, $\forall F(\cdot) \in \mathcal{F}_a$,

$$\int_0^\infty [g(x; F_0) + \gamma x^2]dF(x) \geq G(F_0) + \gamma \int_0^\infty x^2 dF_0(x). \quad (19)$$

Moreover, if γ satisfies $\gamma K_P(F_0) = 0$, then F_0 is also the minima of the primal problem (6).

Furthermore, the following conditions on F_0 can be derived from (19) following the approach in [11].

Corollary 2. *Let $E_0 \subseteq [0, a]$ be the set of points of increase [3] of a distribution function $F_0(\cdot) \in \mathcal{F}_a$.*

$$\int_0^\infty [g(x; F_0) + \gamma x^2]dF(x) \geq G(F_0) + \gamma \int_0^\infty x^2 dF_0(x), \forall F(\cdot) \in \mathcal{F}_a, \quad (20)$$

iff

$$g(x; F_0) + \gamma x^2 \geq G(F_0) + \gamma \int_0^\infty t^2 dF_0(t), \ \ \forall x \in [0, a], \quad (21)$$

and

$$g(x; F_0) + \gamma x^2 = G(F_0) + \gamma \int_0^\infty t^2 dF_0(t), \ \ \forall x \in E_0. \quad (22)$$

[1] A functional G defined on X^* is weak* continuous iff for each $x^* \in X^*$ and each neighborhood V of $f(x^*)$, there is a neighborhood U of x^* such that $f(U) \subset V$. Here, X^* denotes the set of CDFs over $\mathbb{R}$.

[2] A set $K \subset X^*$ is said to be weak* compact if every infinite sequence from K contains a weak* convergent subsequence.

[3] A point $x_0 \in [0, a]$ is said to be a point of increase of the CDF function $F_0(\cdot)$ iff $\exists\delta > 0$, such that $\forall 0 < \varepsilon \leq \delta$, $\int_{x_0-\varepsilon}^{x_0+\varepsilon} \mathrm{d}F_0(x) > 0$

In the following, we refer to (21) and (22) as the KT conditions.

B. Discreteness of the Worst Jamming Distribution

In this subsection, we prove that in order for the optimal CDF function F^* (F_0) to satisfy (21) and (22), the corresponding optimal jamming distribution should be discrete. Recall that the peak power constraint is $a < \infty$, therefore there are two possibilities for the optimal CDF function F^*:

1) F^* has a finite number of points of increase on $[0, a]$.
2) There exists a bounded interval on $[0, a]$ where F^* has an infinite number of points of increase. Note that any continuous distribution of J is a special case of this.

Next we show that case 2 is impossible. The proof makes use of the facts in complex analysis, i.e., the identity theorem [19], to prove that condition (22) cannot be satisfied over an infinite number of points. To facilitate the application of identity theorem, we first need the following asymptotic lower bound of $g(x; F)$ on $x \in \mathbb{R}^+$ for any valid $F(\cdot) \in \mathcal{F}_a$.

Proposition 1. *For any $F(\cdot) \in \mathcal{F}_a$ and $x \in \mathbb{R}^+$ sufficiently large, $g(x; F)$ satisfies*

$$g(x; F) \geq -\mathcal{O}(x) \tag{23}$$

Proof. For any valid $\boldsymbol{r} \in \mathbb{C}$ and $F(\cdot) \in \mathcal{F}_a$, it can be derived that

$$\frac{Q_k(\boldsymbol{r}+\boldsymbol{s}_i, F)}{Q_i(\boldsymbol{r}+\boldsymbol{s}_i, F)} \leq \exp\left(\frac{2|\boldsymbol{s}_k - \boldsymbol{s}_i||\boldsymbol{r}| + 2a|\boldsymbol{r}| + 2a|\boldsymbol{s}_i - \boldsymbol{s}_k|}{\sigma^2}\right) \tag{24}$$

With (24), a lower bound of $L_i(\boldsymbol{r}+\boldsymbol{s}_i, F)$ can be derived as

$$\begin{aligned} L_i(\boldsymbol{r}+\boldsymbol{s}_i, F) &= -\log \frac{\sum_k Q_k(\boldsymbol{r}+\boldsymbol{s}_i, F)}{Q_i(\boldsymbol{r}+\boldsymbol{s}_i, F)} \geq \\ &-\log M_\Phi - \sum_k \frac{2|\boldsymbol{s}_k - \boldsymbol{s}_i||\boldsymbol{r}| + 2a|\boldsymbol{r}| + 2a|\boldsymbol{s}_i - \boldsymbol{s}_k|}{\sigma^2}, \end{aligned} \tag{25}$$

From (7), (18) and (25), we can derive a lower bound of $g(x; F)$ as

$$\begin{aligned} g(x; F) &= \frac{1}{M_\Phi} \sum_i \int_{\mathbb{C}} u(|\boldsymbol{r}|, x) L_i(\boldsymbol{r}+\boldsymbol{s}_i, F) \mathrm{d}\boldsymbol{r} \\ &\geq \int_{\mathbb{C}} u(|\boldsymbol{r}|, x) \left(C_0 + C_1|\boldsymbol{r}|\right) \mathrm{d}\boldsymbol{r} \\ &= \frac{1}{\pi\sigma^2} \int_{\mathbb{C}} e^{-\frac{|\boldsymbol{r}|^2 + x^2}{\sigma^2}} I_0\left(\frac{2x|\boldsymbol{r}|}{\sigma^2}\right)(C_0 + C_1|\boldsymbol{r}|) \mathrm{d}\boldsymbol{r} \end{aligned} \tag{26}$$

for some constants C_0 and C_1. Moreover,

$$\int_{\mathbb{C}} e^{-\frac{|\boldsymbol{r}|^2 + x^2}{\sigma^2}} I_0\left(\frac{2x|\boldsymbol{r}|}{\sigma^2}\right) \mathrm{d}\boldsymbol{r} = \pi e^{-\frac{x^2}{\sigma^2}} \int_0^{+\infty} e^{-\frac{\rho}{\sigma^2}} I_0\left(\frac{2x\sqrt{\rho}}{\sigma^2}\right) \mathrm{d}\rho \tag{27}$$

$$\int_{\mathbb{C}} e^{-\frac{|\boldsymbol{r}|^2 + x^2}{\sigma^2}} I_0\left(\frac{2x|\boldsymbol{r}|}{\sigma^2}\right) |\boldsymbol{r}| \mathrm{d}\boldsymbol{r} = \pi e^{-\frac{x^2}{\sigma^2}} \int_0^{+\infty} \sqrt{\rho} e^{-\frac{\rho}{\sigma^2}} I_0\left(\frac{2x\sqrt{\rho}}{\sigma^2}\right) \mathrm{d}\rho \tag{28}$$

From [20, 6.643], for any $Re(\mu + \frac{1}{2}) > 0$, we have

$$\int_0^\infty \frac{x^{\mu - \frac{1}{2}}}{e^{ax}} I_0(2b\sqrt{x}) \mathrm{d}x = \frac{\Gamma(\mu + \frac{1}{2})}{b} a^{-\mu} e^{\frac{b^2}{2a}} \mathcal{M}_{-\mu,0}\left(\frac{b^2}{a}\right) \tag{29}$$

where $\mathcal{M}_{\cdot,\cdot}(\cdot)$ denotes the Whittaker M function. For the special case of $\mu = \frac{1}{2}$, we have

$$\mathcal{M}_{-\frac{1}{2},0}(x) = \sqrt{x} e^{\frac{x}{2}}. \tag{30}$$

When x is sufficiently large, $\mathcal{M}_{\mu,0}(x)$ can be approximated by [21, 13.14.20]

$$\mathcal{M}_{\mu,0}(x) \approx \frac{x^{-\mu}}{\Gamma(1/2 - \mu)} e^{\frac{x}{2}} \tag{31}$$

Hence when x is sufficiently large, (27) and (28) scale as

$$\begin{gathered} e^{-\frac{x^2}{\sigma^2}} \int_0^{+\infty} e^{-\frac{\rho}{\sigma^2}} I_0\left(\frac{2x\sqrt{\rho}}{\sigma^2}\right) \mathrm{d}\rho = \mathcal{O}(1), \\ e^{-\frac{x^2}{\sigma^2}} \int_0^{+\infty} \sqrt{\rho} e^{-\frac{\rho}{\sigma^2}} I_0\left(\frac{2x\sqrt{\rho}}{\sigma^2}\right) \mathrm{d}\rho = \mathcal{O}(x). \end{gathered} \tag{32}$$

Following (26)-(32), we can conclude that for $x \in \mathbb{R}^+$ sufficiently large, $g(x; F)$ has $g(x; F) \geq -\mathcal{O}(x)$, which completes the proof of Proposition 1. □

With Proposition 1, we are able to prove that the "optimal" (i.e., the worst) jamming distribution should be discrete, which is stated in the following theorem.

Theorem 4. *Let E^* denote the set of points of increase for the optimal jamming distribution F^*, then E^* has a finite number of elements.*

Proof. We prove the theorem by contradiction. First we show that if $|E^*| = \infty$, then for the worst jamming distribution F^*,

$$g(x; F^*) + \gamma x^2 \equiv G(F^*) + \gamma \int_0^\infty t^2 \mathrm{d}F^*(t), \ \ \forall x \in \mathbb{R}^+. \tag{33}$$

In fact, since each point in E^* is bounded on $[0, a]$, using the Bolzano-Weierstrass theorem, we can find an infinite sequence $x_n, n = 1, 2, ..., \infty$ in E^* which has a limit point x^*, i.e.,

$$\lim_{n \to \infty} x_n = x^*, \text{and } x^* \in [0, a]. \tag{34}$$

From Corollary 2, we have $g(x_n; F^*) + \gamma x_n^2 \equiv G(F^*) + \gamma \int_0^\infty t^2 \mathrm{d}F^*(t), \forall x_n$.

It can be verified that the function defined in (21) and (22),

$$g(z; F^*) + \gamma z^2, \tag{35}$$

is analytic w.r.t. $z \in \mathbb{C}$ for any valid CDF function $F(\cdot) \in \mathcal{F}_a$, which implies its continuity on $[0, a]$. So we have

$$\begin{aligned} \lim_{n \to \infty} [g(x_n; F^*) + \gamma x_n^2] &= g(x^*; F^*) + \gamma (x^*)^2 \\ &= G(F^*) + \gamma \int_0^\infty t^2 \mathrm{d}F^*(t). \end{aligned} \tag{36}$$

From the Identity Theorem [19], if two analytic functions are identical on a infinite set of points (sequence x_n) in a region along with their limit points (x^*), these two functions are identical in the entire region. Therefore, we have

$$g(x; F^*) + \gamma x^2 - G(F^*) - \gamma \int_0^\infty t^2 \mathrm{d}F^*(t) \equiv 0, \forall x \in \mathbb{R}^+. \tag{37}$$

which is equal to (33).

Next we consider two cases: (1) $\gamma > 0$ and (2) $\gamma = 0$. We show that (33) cannot hold in either of them.

For $\gamma > 0$, note that $g(x;F^*) \geq -\mathcal{O}(x)$, the lower bound of $g(x;F^*) + \gamma x^2 - G(F^*) - \gamma \int_0^\infty t^2 \mathrm{d}F^*(t)$ scales quadratically with x as $x \to \infty$. That is, when x is sufficiently large, $g(x;F^*) + \gamma x^2 - G(F^*) - \gamma \int_0^\infty t^2 \mathrm{d}F^*(t) > 0$, which contradicts with (37). So for $\gamma > 0$, (33) cannot hold.

For $\gamma = 0$, the average power constraint is inactive. If for some finite a_0, there exists a distribution function ${F_{a_0}}^*(\cdot) \in \mathcal{F}_{a_0}$, which has an infinite number of points of increase, that can minimize functional $G(\cdot)$, then it follows from (37) that

$$g(x;{F_{a_0}}^*) - G({F_{a_0}}^*) \equiv 0, \forall x \in \mathbb{R}^+ \tag{38}$$

On the other hand, for any $\alpha > a_0$, ${F_{a_0}}^*(\cdot) \in \mathcal{F}_\alpha$. From (38), the KT conditions are always satisfied for distribution function ${F_{a_0}}^*$. Thus, for any $\alpha \geq a_0$, ${F_{a_0}}^*$ can minimize $G(\cdot)$ within $\mathcal{F}_\alpha$. However, in the next, we will show that as $\alpha \to \infty$, $\min_{F\in\mathcal{F}_\alpha} G(F) \to -\log M_\Phi$, i.e. the MI between the transmitted signal S and the received signal R approaches 0.

For any $\alpha > 0, \beta > 0$, define $\hat{F}_{\alpha,\beta}(x)$ to be a truncated Rayleigh distribution function as

$$\hat{F}_{\alpha,\beta}(x) \triangleq \begin{cases} 1 - e^{-\frac{x^2}{2\beta^2}}, & x < \alpha \\ 1, & x \geq \alpha \end{cases} \tag{39}$$

which converges to the Rayleigh distribution function as $\alpha \to \infty$ under the Levy metric [22, Section 2.3].

In [7], we have shown that if J follows a Rayleigh distribution of scale parameter β, and S has an average power constraint of P_S, then the MI between S and R is upper bounded by

$$I(S,R) \leq \log(1 + \frac{P_S}{\sigma^2 + 2\beta^2}). \tag{40}$$

Since $H(S) = \log M_\Phi$, then $-H(R|S) \leq \log(1 + \frac{P_S}{\sigma^2+2\beta^2}) - \log M_\Phi$ if J follows the Rayleigh distribution. Note that the third raw moment of a Rayleigh RV is bounded for any finite scale parameter α, from the weak continuity of $G(\cdot)$, we have

$$\lim_{\alpha\to\infty} G(\hat{F}_{\alpha,\beta}(x)) \leq \log(1 + \frac{P_S}{\sigma^2 + 2\beta^2}) - \log M_\Phi \tag{41}$$

$$\lim_{\alpha\to\infty} \min_{F\in\mathcal{F}_\alpha} G(F) \leq \lim_{\alpha\to\infty} G(\hat{F}_{\alpha,\beta}(x)) \leq \log(1 + \frac{P_S}{\sigma^2 + 2\beta^2}) - \log M_\Phi \tag{42}$$

for any $\beta > 0$. Since $G(F)$ is lower bounded by $-\log M_\Phi$, as $\beta \to \infty$, we have

$$\lim_{\alpha\to\infty} \min_{F\in\mathcal{F}_\alpha} G(F) = -\log M_\Phi \tag{43}$$

Since we have concluded that $G({F_{a_0}}^*) = \min_{F\in\mathcal{F}_\alpha} G(F)$ for any $\alpha > a_0$, we must have

$$G({F_{a_0}}^*) = -\log M_\Phi. \tag{44}$$

This indicates that the channel capacity becomes 0 when the jamming distribution is ${F_{a_0}}^*$, which is impossible since a positive channel capacity is ensured for a non-symmetrizable AVC channel as shown in Lemma 1. Therefore, the distribution function ${F_{a_0}}^*$ does not exist and (33) cannot hold.

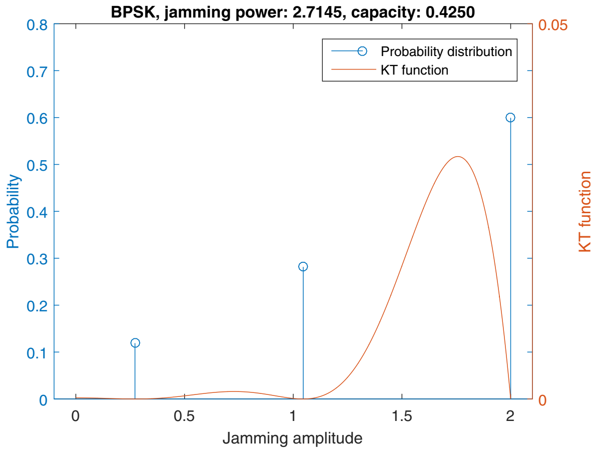

Fig. 1: The optimal jamming distribution and KT conditions for $a = 2, \gamma = 0$, BPSK alphabet.

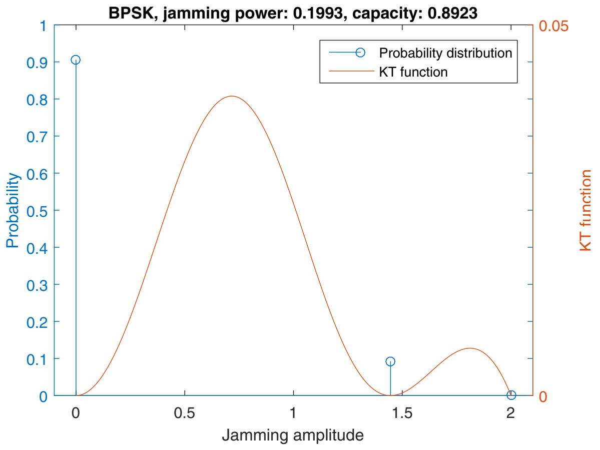

Fig. 2: The optimal jamming distribution and KT conditions for $a = 2, \gamma = 0.4$, BPSK alphabet.

As (33) cannot hold for any $\gamma \geq 0$, it implies that $|E^*| = \infty$ cannot hold. So E^* only has a finite number of elements, which completes the proof. □

Even though a necessary and sufficient condition on minima F^* has been given in Corollary 2 and we have proved that the optimal jamming distribution should be discrete, it is still difficult to derive a closed form expression of F^*. Therefore, we resort to numerical methods to evaluate the optimal distribution of J as well as the corresponding channel capacity in Section IV.

IV. Numerical Results

In this section, we compute the worst jamming distribution through numerical methods.

The KT conditions in Corollary 2 provide the gradients in the optimization problem [11], and the convexity of functional $G(\cdot)$ guarantees the convergence to the global optimal in optimization. Throughout the numerical results, the average power of the legitimate signal is normalized to 1 and we compute

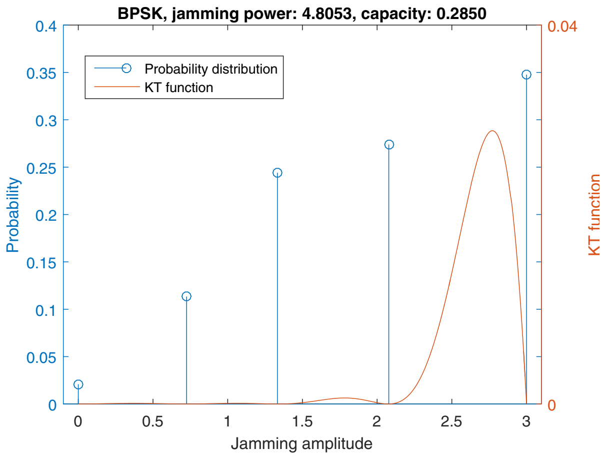

Fig. 3: The optimal jamming distribution and KT conditions for $a = 3, \gamma = 0$, BPSK alphabet.

the optimal jamming distribution assuming SNR = 6 dB. We vary the peak power constraint a and KT multiplier γ in the numerical computation. Fig. 1–3 show the worst jamming distributions and the KT functions[4] under different scenarios. The numerical results verify the theoretical analysis, where the worst jamming distributions are discrete. The corresponding KT conditions satisfy Corollary 2, which indicate that the obtained jamming distributions are indeed the global optimal. We can verify that the worst jamming distributions only contain a finite number of points. In addition, we can observe that: 1) the number and locations of the mass points, as well as the corresponding probabilities, vary with the parameters of the system. A constant jamming with the maximal peak power will not generally render the minimal channel capacity, since varying the jamming power level is able to add more ambiguity to the received signal. 2) Under strong average power constraints (jamming with low average power compared with that of the legitimate signal), the optimal strategy for the jammer is to launch jamming with a higher overall power level during a proportional of time and then keep silent during the other time intervals, instead of a constant jamming with a lower power level all the time. 3) Unlike the binary jamming distribution in [8] which maximizes the error rates of the legitimate receiver, the jamming distribution that minimizes the channel capacity tends to have more mass points.

V. Conclusions

In this paper, we discussed the problem of finding the worst jamming distribution in terms of channel capacity for the SP-OFDM system. We assumed that the transmitting symbol is uniformly distributed over the alphabet and the jamming interference is subject to both peak and average power constraints. We proved the existence and the uniqueness of the worst jamming distribution. By deriving the KT conditions for the worst jamming distribution, we further proved that the worst jamming distribution should be discrete in amplitude with a finite number of mass points. Numerical examples were provided to demonstrate our theoretical results. The in-depth analysis on the worst jamming distribution carried out here provided insightful information on the performance of SP-OFDM under destructive hostile jamming.

References

[1] A. Lapidoth and P. Narayan, “Reliable communication under channel uncertainty,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, Oct 1998.

[2] T. Song, K. Zhou, and T. Li, “Cdma system design and capacity analysis under disguised jamming,” *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2487–2498, Nov 2016.

[3] L. Zhang and T. Li, “Anti-jamming message-driven frequency hopping-part ii: Capacity analysis under disguised jamming,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 80–88, January 2013.

[4] T. Song, Z. Fang, J. Ren, and T. Li, “Precoding for ofdm under disguised jamming,” in *2014 IEEE Global Communications Conference*, Dec 2014, pp. 3958–3963.

[5] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge University Press, 2012.

[6] R. Ahlswede, “Elimination of correlation in random codes for arbitrarily varying channels,” *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 44, no. 2, pp. 159–175, 1978.

[7] Y. Liang, J. Ren, and T. Li, “Spectrally efficient ofdm system design under disguised jamming,” 2018, arXiv preprint, submitted to GLOBECOM 2018.

[8] S. Amuru and R. M. Buehrer, “Optimal jamming against digital modulation,” *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2212–2224, Oct 2015.

[9] D. G. Luenberger, *Optimization by Vector Space Methods*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc., 1997.

[10] S. Shamai and I. Bar-David, “The capacity of average and peak-power-limited quadrature gaussian channels,” *IEEE Transactions on Information Theory*, vol. 41, no. 4, pp. 1060–1071, Jul 1995.

[11] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, “The capacity of discrete-time memoryless rayleigh-fading channels,” *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[12] M. C. Gursoy, H. V. Poor, and S. Verdu, “The capacity of the noncoherent rician fading channel,” Princeton University, Tech. Rep., December 2002.

[13] A. ElMoslimany and T. M. Duman, “On the capacity of multiple-antenna systems and parallel gaussian channels with amplitude-limited inputs,” *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2888–2899, July 2016.

[14] I. Csiszar and P. Narayan, “The capacity of the arbitrarily varying channel revisited: positivity, constraints,” *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 181–193, Mar 1988.

[15] I. Csiszar, “Arbitrarily varying channels with general alphabets and states,” *IEEE Transactions on Information Theory*, vol. 38, no. 6, pp. 1725–1742, Nov 1992.

[16] T. H. Chan, S. Hranilovic, and F. R. Kschischang, “Capacity-achieving probability measure for conditionally gaussian channels with bounded inputs,” *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley-Interscience, 1991.

[18] J. G. Smith, “The information capacity of amplitude- and variance-constrained sclar gaussian channels,” *Information and Control*, vol. 18, no. 3, pp. 203 – 219, 1971.

[19] D. Sarason, *Complex Function Theory*. American Mathematical Society.

[20] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Elsevier/Academic Press, Amsterdam, 2007.

[21] “*NIST Digital Library of Mathematical Functions*,” http://dlmf.nist.gov/, Release 1.0.15 of 2017-06-01, f. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[22] P. Huber, *Robust Statistics*. Wiley, 2004.

[4] In the numerical results, we refer to KT function as $g(x;F) + \gamma x^2 - G(F) - \gamma \int_0^\infty t^2 \mathrm{d}F_0(t)$ given a jamming distribution F, where x is the jamming amplitude.