On the Capacity of the AWGN Channel with Additive Radar Interference

Sara Shahi, Daniela Tuninetti, and Natasha Devroye
ECE, University of Illinois at Chicago, Chicago IL 60607, USA.
Email: sshahi7, danielat, devroye @uic.edu

***Abstract*—This paper investigates the capacity of a communications channel that, in addition to additive white Gaussian noise, also suffers the interference from a co-existing radar transmission. The radar interference (of short duty-cycle and of much wider bandwidth than the intended communication signal) is modeled as an additive term whose amplitude is known and constant, but whose phase is independent and identically uniformly distributed at each channel use. The capacity achieving input distribution, under the standard average power constraint, is shown to have independent modulo and phase. The phase is uniformly distributed in $[0, 2\pi]$. The modulo is discrete with countably infinite many mass points, but only finitely many in any bounded interval. From numerical evaluations, a proper-complex Gaussian input is seen to perform quite well for weak radar interference. Interestingly, for very large radar interference, a Gaussian input achieves $\frac{1}{2}\log(1+\mathsf{S})$. Since a Gaussian input is optimal to within one bit, it is concluded that the presence of the radar interference results in a loss of half degrees of freedom compared to an interference free channel.**

I. Introduction

Shortage of spectrum resources, specially with the ever increasing demand for commercial services, necessitates a more sensible bandwidth allocation policy. In 2012, the President's council of Advisors on Science and Technology published a report that recommended to release portions of governmental radar bands (e.g. 3550-3620 MHz) to be shared with commercial wireless services. Since then, several national funding agencies have launched research programs to encourage research in this area.

To understand how these two very different systems should best share the spectrum, it is useful to have an idea of the fundamental information theoretic performance limits of such spectrum sharing channels. In this work, we investigate the capacity of a white additive Gaussian noise communications channel, which in addition to noise, suffers interference from a radar transmission. In the channel model studied here, while the interfering radar transmission is modeled to be additive, it is not Gaussian. Rather, it is modeled as a constant (thus known) amplitude signal, but with unknown and uniformly distributed phase at each channel use. The reason is that radar systems usually operate with a high power, short duty-cycle waveform of much wider bandwidth than the intended communication signal. The capacity of an additive Gaussian noise channel under an average power constraint is well known: the optimal input is Gaussian of power equal to the power constraint. However, since the channel studied here is no longer Gaussian, several questions emerge: (i) what is the capacity of this channel and how does it differ from that of a Gaussian noise channel (without the radar interference), and (ii) what input achieves the capacity. In this paper we aim to address both these questions.

A. Past Work

The capacity of channels with additive noise and various input constraints has been studied before.

In [1] bounds on the capacity of additive, but not necessarily Gaussian, noise channel were given. Applying Ihara's upper bound to our channel model yields a bound that grows as the radar signal amplitude increases. This bound is not tight because the capacity of our channel is upper bounded by the capacity of the classical power-contained Gaussian noise channel without radar interference.

In [2, Theorem 1] is was shown that for any memoryless additive noise channel with a second moment/power constraint on the input, the rate achievable by using a white Gaussian input never incurs a loss of more than half a bit per real dimension with respect to capacity. This implies that one can obtain a capacity upper bound for a complex-valued additive noise channel by adding 1 bit to the rate attained with a proper-complex Gaussian input for the same channel.

In the seminal work by Smith [3], it was shown that the capacity of a real-valued white Gaussian noise channel with peak amplitude and average power constraints is achieved by a discrete input with a finite number of mass points. This is in sharp contrast to Gaussian inputs that achieve the capacity when the amplitude constraint is dropped. Later, the optimality of a discrete input under peak amplitude constraint was shown to hold for a wide class of additive noise channels [4]. Recently it was shown that, under average power constraint and certain 'smoothness' conditions on the noise distribution, the only additive noise channel whose capacity achieving input is continuous is the Gaussian noise channel [5].

The model considered in this paper, is a complex-valued additive noise channel with an average power constraint. When we transform the mutual information optimization problem over a bivariate (modulo and phase) input distribution into one over a univariate (modulo only) input distribution, the equivalent channel is no longer additive. For this equivalent non-additive channel, we can *not* proceed as per the steps preceding [5, eq.(4)]. This is so because [5, eq.(4)] heavily relies on certain integrals being convolution integrals and thus passing to the Fourier domain to study/infer certain properties of the optimal input distribution. In non-additive channels this is not possible. In this respect, our approach is similar to that of [6] where the capacity of the complex-valued Gaussian noise channel under average power and peak amplitude constraints was shown to be achieved by a complex-valued input with independent amplitude and phase; the optimal phase is uniformly distributed in the interval $[0, 2\pi]$, and the optimal amplitude is discrete with finitely many mass points. In this work we follow closely the steps in [6].

Extensions of Smith's work [3] to Gaussian channels with various fading models, possibly MIMO, are known in the literature but are not reported here because they are not directly relevant.

In [7], [8] a subset of the authors studied the uncoded average symbol error rate performance of the same channel model considered in this paper. Two regimes of operation emerged. In the low Interference to Noise Ratio (INR) regime, it was shown that the optimal decoder is a minimum Euclidean distance decoder, as for Gaussian noise only; while in the high INR regime, radar interference estimation and cancellation is optimal. Interestingly, in the process of canceling the radar interference at high INR, also part of the useful signal is removed, and that after cancellation the equivalent output is real-valued (one of the two real-valued dimensions of the original complex-valued output is lost). We shall observe a similar behavior for the capacity of this channel.

B. Contributions

The capacity of the channel model proposed here has not, to the best of our knowledge, been studied before and provides a new model for bounding the performance of a communication system in the presence of radar interference. Likewise, in the literature on the co-existence of radar and communications channels, we are not aware of any capacity results. Our contributions thus lie in the study of the capacity of this novel channel model, in which we show that the optimal input distribution has independent modulo and phase. The phase is uniformly distributed in $[0, 2\pi]$. The modulo is discrete with countably infinite many mass points, but only finitely many in any bounded interval.

We also show achievable rates. The Gaussian input is seen to perform very well for weak radar interference, where it closely follows the upper bound in [1], while for very large radar interference, it attains exactly half the interference-free capacity. This implies, based on [2, Theorem 1], that for large radar interference radar interference fundamentally limits the performance by entailing a loss of half the degrees of freedom compared to the interference free channel.

C. Paper organization

The paper is organized as follows. Section II introduces the channel model. Section III derives our main result. Section IV provides numerical results. Section V concludes the paper. Proofs can be found in the Appendix.

II. System Model

Next, boldface letters indicate complex-valued random variables, while lightface letters real-valued ones. In addition, we use $\mathbb{R}_+$ to represent the set

$$\mathbb{R}_+ = \{x : x \geq 0\}.$$

We model the effect of a high power, short duty cycle radar pulse at the receiver of a narrowband data communication system as

$$\mathbf{Y} = \mathbf{X} + \mathbf{W}, \tag{1}$$

$$\mathbf{W} = \sqrt{\mathsf{I}} e^{j\Theta_\mathsf{I}} + \mathbf{Z}, \tag{2}$$

where $\mathbf{Y}$ is the channel output, $\mathbf{X}$ is the input signal subject to the average power constraint $\mathbb{E}[|\mathbf{X}|^2] \leq \mathsf{S}$, Θ_I is the random phase of the radar interference uniformly distributed in $[0, 2\pi]$, and $\mathbf{Z}$ is a zero-mean proper-complex unit-variance Gaussian noise. The random variables $(\mathbf{X}, \Theta_\mathsf{I}, \mathbf{Z})$ are mutually independent. Θ_I and $\mathbf{Z}$ are independent and identically distributed over channel uses, that is, the channel is memoryless. Our normalizations imply that S is the average Signal to Noise Ratio (SNR) while I is the average Interference to Noise Ratio (INR). We assume I to be fixed and thus known. For later use, the distribution of the additive noise in (2) is given by

$$f_\mathbf{W}(\mathbf{w}) = \mathbb{E}_{\Theta_\mathsf{I}}\left[\frac{e^{-|\mathbf{w}-\sqrt{\mathsf{I}}e^{j\Theta_\mathsf{I}}|^2}}{\pi}\right] = \frac{e^{-|\mathbf{w}|^2-\mathsf{I}}}{\pi} I_0\left(2\sqrt{\mathsf{I}|\mathbf{w}|^2}\right), \tag{3}$$

where $I_0(\mathbf{w}) = \mathbb{E}[e^{\mathbf{w}\cos(\Theta_\mathsf{I})}] \in [1, e^{|\mathbf{w}|}]$ for $\mathbf{w} \in \mathbb{C}$ is the zero-order modified Bessel function of the first kind. The channel transition probability is thus

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = f_\mathbf{W}(\mathbf{y} - \mathbf{x}), \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{C}^2. \tag{4}$$

Our goal is to characterize the capacity of the memoryless channel in (1)-(2) given by

$$\mathsf{C} = \sup_{F_\mathbf{X}:\mathbb{E}[|\mathbf{X}|^2]\leq \mathsf{S}} I(\mathbf{X}; \mathbf{Y}), \tag{5}$$

where $F_\mathbf{X}$ is the cumulative distribution function of $\mathbf{X}$.

III. Channel Capacity

A. Trivial Bounds

Trivially, one can lower bound the capacity in (5) by treating the radar interference as a Gaussian noise and obtain

$$\log\left(1 + \frac{\mathsf{S}}{1 + \mathsf{I}}\right) \leq \mathsf{C}, \tag{6}$$

and upper bound it as

$$\mathsf{C} \leq \max_{F_\mathbf{X}:\mathbb{E}[|\mathbf{X}|^2]\leq \mathsf{S}} I(\mathbf{X}; \mathbf{Y}, \Theta_\mathsf{I}) = \log(1 + \mathsf{S}), \tag{7}$$

or from Ihara's work [1] as

$$\mathsf{C} \leq \log(\pi e(1 + \mathsf{S} + \mathsf{I})) - h(\mathbf{W}), \tag{8}$$

or from Zamir and Erez's work [2, Theorem 1], as

$$\mathsf{C} \leq I(\mathbf{X}_G; \mathbf{Y}) + \log(2), \tag{9}$$

where $I(\mathbf{X}_G; \mathbf{Y})$ is the achievable rate with a proper-complex Gaussian input that meets the power constraint with equality.

B. Equivalent Problem Formulation and Definitions

We aim to show that the supremum in (5) is actually attained by a *unique* input distribution, for which we want to derive its structural properties. Before we continue however, we rewrite the optimization for the original channel (involving the real and the imaginary part of the input) in a way that it is sufficient to optimize with respect to a univariate distribution only.

By following steps similar to those in [6, Section II.B], we can show that an optimal input distribution induces $\widetilde{Y} := |\mathbf{Y}|^2$ and $\angle\mathbf{Y}$ independent given $\mathbf{X}$, with $\angle\mathbf{Y}$ uniformly distributed over the interval $[0, 2\pi]$; such an output distribution can be attained by the uniform distribution on $\angle\mathbf{X}$ and by $\angle\mathbf{X}$ independent of $\widetilde{X} := |\mathbf{X}|^2$; therefore, it is convenient for later use to denote the channel transition probability $f_{\widetilde{Y}|\widetilde{X}}(y|x)$ as the kernel $K(x, y)$ given by (see Appendix VI-A)

$$K(x, y) := f_{\widetilde{Y}|\widetilde{X}}(y|x) \tag{10a}$$

$$= \int_{|b|\le\pi} \frac{e^{-\mathsf{I}-\xi(b;x,y)}}{2\pi} I_0\left(2\sqrt{\mathsf{I}\ \xi(b;x,y)}\right)\ db, \tag{10b}$$

$$\xi(b; x, y) := y + x - 2\sqrt{yx}\cos(b) \ge 0,\ (y, x) \in \mathbb{R}_+^2. \tag{10c}$$

Since the random variables $\widetilde{X} = |\mathbf{X}|^2$ and $\widetilde{Y} = |\mathbf{Y}|^2$ are connected through a channel with kernel $K(x, y)$, an input distributed as $F_{\widetilde{X}}$ results in an output with probability distribution function (pdf) given by[1]

$$f_{\widetilde{Y}}(y; F_{\widetilde{X}}) := \int_{x\ge0} K(x, y) dF_{\widetilde{X}}(x), \quad y \in \mathbb{R}_+. \tag{11}$$

We stress the dependence of the output pdf on the input distribution $F_{\widetilde{X}}$ by adding it as an 'argument' in $f_{\widetilde{Y}}(y; F_{\widetilde{X}})$.

Finding the channel capacity in (5) can thus be equivalently expressed as the following optimization over the distribution of a non-negative random variable $\widetilde{X}$

$$\mathsf{C} + h(|\mathbf{W}|^2) = \sup_{F_{\widetilde{X}}:\mathbb{E}[\widetilde{X}]\le\mathsf{S}} h(\widetilde{Y}; F_{\widetilde{X}}), \tag{12}$$

where $h(\widetilde{Y}; F_{\widetilde{X}})$ is the output differential entropy given by [2]

$$h(\widetilde{Y}; F_{\widetilde{X}}) = \int_{y\ge0} f_{\widetilde{Y}}(y; F_{\widetilde{X}}) \log \frac{1}{f_{\widetilde{Y}}(y; F_{\widetilde{X}})}\ dy. \tag{13}$$

We express $h(\widetilde{Y}; F_{\widetilde{X}})$ in (13) as

$$h(\widetilde{Y}; F_{\widetilde{X}}) = \int_{y\ge0}\int_{x\ge0} K(x, y) \log \frac{1}{f_{\widetilde{Y}}(y; F_{\widetilde{X}})}\ dF_{\widetilde{X}}(x)\ dy$$

$$= \int_{x\ge0} h(x; F_{\widetilde{X}})\ dF_{\widetilde{X}}(x), \tag{14}$$

where we defined the *marginal entropy* $h(x; F_{\widetilde{X}})$ as [3]

$$h(x; F_{\widetilde{X}}) := \int_{y\ge0} K(x, y) \log \frac{1}{f_{\widetilde{Y}}(y; F_{\widetilde{X}})}\ dy,\ x \in \mathbb{R}_+, \tag{15}$$

and where the order of integration in the line above (14) can be swapped by Fubini's theorem.

For later use, we note that the introduced functions can be bounded as follows: for the kernel in (10)

$$e^{-(y+x+\mathsf{I})} \le K(x, y) \le 1, \quad (x, y) \in \mathbb{R}_+^2; \tag{16}$$

for the output pdf in (11)

$$e^{-(y+\mathsf{I}+\beta_{F_{\widetilde{X}}})} \le f_{\widetilde{Y}}(y; F_{\widetilde{X}}) \le 1, \quad y \in \mathbb{R}_+, \tag{17}$$

where $\beta_{F_{\widetilde{X}}}$ is defined and bounded (by using Jensen's inequality together with the power constraint) as

$$0 \le \beta_{F_{\widetilde{X}}} := -\ln\left(\int_{x\ge0} e^{-x} dF_{\widetilde{X}}(x)\right) \le \mathsf{S}; \tag{18}$$

for the marginal entropy in (15)

$$0 \le h(x; F_{\widetilde{X}}) \le \mathbb{E}[\widetilde{Y}|\widetilde{X} = x] + \mathsf{I} + \beta_{F_{\widetilde{X}}}, \quad x \in \mathbb{R}_+, \tag{19}$$

where the conditional mean of $\widetilde{Y} = |\mathbf{Y}|^2$ given $\widetilde{X} = |\mathbf{X}|^2$ is

$$\mathbb{E}[\widetilde{Y}|\widetilde{X} = x] = x + \mathsf{I} + 1, \quad x \in \mathbb{R}_+, \tag{20}$$

C. Main Result

We are now ready to state our main result: a characterization of the structural properties of the optimal input distribution in (5), in relation to the problem in (12).

Theorem 1. *The optimal input distribution in* (5) *unique and has independent modulo and phase. The phase is uniformly distributed in* $[0, 2\pi]$. *The modulo is discrete with countably infinite many mass points, but only finitely many in any bounded interval.*

Proof: As argued in Section III-B, an optimal input distribution in (5) has $\angle\mathbf{X}$ uniformly distributed in $[0, 2\pi]$ and independent of $\widetilde{X} := |\mathbf{X}|^2$. The modulo $\widetilde{X}$ solves the problem in (12), whose supremum is attained by the *unique* input distribution $F^{\text{opt}}_{\widetilde{X}}$ that solves (12) because (see [9]):

1) the space of input distributions $\mathcal{F}$ is compact and convex (see [9, Theorem 1]); $\mathcal{F}$ is given by

$$\mathcal{F} := \Big\{F_{\widetilde{X}} : F_{\widetilde{X}}(x) = 0,\ \forall x < 0, \tag{21a}$$

$$dF_{\widetilde{X}}(x) \ge 0,\ \forall x \ge 0, \tag{21b}$$

$$\int_{x\ge0} 1 \cdot dF_{\widetilde{X}}(x) = 1, \tag{21c}$$

$$L(F_{\widetilde{X}}) := \int_{x\ge0} x \cdot dF_{\widetilde{X}}(x) - \mathsf{S} \le 0\Big\}, \tag{21d}$$

where the various constraints are: (21a) for non-negativity, (21b) and (21c) for a valid input distribution, and (21d) for the average power constraint; and

2) The differential entropy $h(\widetilde{Y}; F_{\widetilde{X}})$ in (14) is a weak* continuous (see Appendix VI-B) and strictly concave (Appendix VI-C) functional of the input distribution $F_{\widetilde{X}}$.

[1]The pdf $f_{\widetilde{Y}}(y; F_{\widetilde{X}})$ in (11) exists since the kernel in (10) is a continuous and bounded (see (16)) function and thus integrable.

[2]The entropy $h(\widetilde{Y}; F_{\widetilde{X}})$ in (13) exists since the output pdf in (11) is a continuous and bounded (see (17)) function and thus integrable.

[3]The marginal entropy $h(x; F_{\widetilde{X}})$ in (15) exists since the involved functions are integrable by (16) and (17).

From this and by Smith's approach [3], the solution of the optimization problem in (12) is equivalent to the solution of

$$h'_{F^{\text{opt}}_{\widetilde{X}}}(\widetilde{Y}; F_{\widetilde{X}}) - \lambda L'_{F^{\text{opt}}_{\widetilde{X}}}(F_{\widetilde{X}}) \leq 0, \text{ for all } F_{\widetilde{X}} \in \mathcal{F}, \tag{22a}$$

$$\lambda \geq 0 : L(F^{\text{opt}}_{\widetilde{X}}) = 0, \tag{22b}$$

where the functional $L(.)$ was defined in (21d), and where the prime denotes the weak* derivative (see Appendix VI-D).

The conditions in (22) can be equivalently expressed as the necessary and sufficient Karush-Kuhn-Tucker (KKT) condition: for some $\lambda \geq 0$

$$h(x; F^{\text{opt}}_{\widetilde{X}}) \leq h(\widetilde{Y}; F^{\text{opt}}_{\widetilde{X}}) + \lambda(x - \mathsf{S}), \quad \forall x \in \mathbb{R}_+, \tag{23}$$

where equality in (23) holds *only* at the points of increase of $F^{\text{opt}}_{\widetilde{X}}$ (see Appendix VI-E).

At this point, as it is usual in these types of problems [3], the proof follows by ruling out the other types of distributions different from the stated one. Generally speaking a distribution can have one of the following forms:

1) Its support contains an infinite number of mass points in some bounded interval;
2) It is discrete with finitely many mass points; and
3) It is discrete with countably infinitely many mass points but only a finite number of them in any bounded interval.

Next, we will rule out cases 1 and 2 by contradiction.

Rule out case 1 ($F^{\text{opt}}_{\widetilde{X}}$ has an infinite number of mass points in some bounded interval). We prove that this case corresponds to the situation where the inequality in (23) must hold with equality for all $x \geq 0$, which is impossible.

We start by noting that the optimal Lagrange multiplier $\lambda^{\text{opt}}(\mathsf{S})$, which represents the weak* derivative of the capacity C with respect to S, must satisfy $0 < \lambda^{\text{opt}}(\mathsf{S}) < 1$ for all $\mathsf{S} > 0$. This is so because, by the Envelope Theorem [10] and the upper bound in (7), the case $\lambda \geq 1$ is impossible. Also the case $\lambda^{\text{opt}}(\mathsf{S}) = 0$ is impossible; if otherwise, the unique solution of (23) (where uniqueness follows by invertibility of the integral transform in (11) as proven in Appendix VI-G) would induce the output pdf

$$f_{\widetilde{Y}}(y; F^{\text{opt}}_{\widetilde{X}}) = \exp\{-h(\widetilde{Y}; F^{\text{opt}}_{\widetilde{X}})\}, \ \forall y \in \mathbb{R}_+, \tag{24}$$

which is not a valid pdf since it does not integrate to one. Therefore we conclude that we must have $0 < \lambda^{\text{opt}}(\mathsf{S}) < 1$.

For the remaining case $0 < \lambda < 1$, we re-write the KKT condition in (23) by following the recent work [5]. Given the conditional output power expressed as in (20), we can write

$$x - \mathsf{S} = \int_{y \geq 0} \Big(y - (1 + \mathsf{I} + \mathsf{S})\Big) K(x, y) \ dy, \ \forall x \in \mathbb{R}_+. \tag{25}$$

With (25), the KKT condition in (23) reads: there exists a constant $0 < \lambda < 1$ such that

$$g(x, \lambda) \leq h(\widetilde{Y}; F^{\text{opt}}_{\widetilde{X}}) = \text{constant for all } x \in \mathbb{R}_+, \tag{26}$$

with equality only at the points of increase of $F^{\text{opt}}_{\widetilde{X}}$, and where

$$g(x, \lambda) := \int_{y \geq 0} K(x, y) \log \left(\frac{\lambda e^{-\lambda y}}{f_{\widetilde{Y}}(y; F^{\text{opt}}_{\widetilde{X}})} \right) dy \tag{27a}$$

$$+ \log \frac{1}{\lambda} + \lambda(1 + \mathsf{I} + \mathsf{S}). \tag{27b}$$

We show next that (26) can not be satisfied if $F^{\text{opt}}_{\widetilde{X}}$ contains an infinite number of mass points in some bounded interval. This step is accomplished by showing that the function $g(x, \lambda), \ x \in \mathbb{R}_+$, in (27) can be extended to the complex domain and that $g(z, \lambda), \ z \in \mathbb{C} : \Re(z) > 0$, is analytic. Note that it is sufficient to prove the analiticity of $g(z, \lambda)$ only for a strip around the non-negative real line but we prove it for all the right half plane in the complex domain (see Appendix VI-F). The analyticity of $g(z, \lambda), z \in \mathbb{C} : \Re(z) > 0$, and the existence of an accumulation point for the set of points of increase of $F^{\text{opt}}_{\widetilde{X}}$ (by Bolzano Weierstrass Theorem [11]) together with the Identity Theorem [11], implies that $g(z, \lambda) = \text{constant} \ \forall z \in \mathbb{C} : \Re(z) > 0$. From this we conclude that $g(x, \lambda) = \text{constant} \ \forall x \in \mathbb{R}_+$. One solution, and the *only solution* due to invertibility of the kernel $K(x, y)$ (see Appendix VI-G), for $g(x, \lambda)$ not to depend on x is that the function that multiplies the kernel in the integral in (27a) is a constant (in which case $\int_{y \geq 0} K(x, y) \ dy = 1$ for all $x \in \mathbb{R}_+$). For this to happen, we need

$$f_{\widetilde{Y}}(y; F^{\text{opt}}_{\widetilde{X}}) = \lambda e^{-\lambda y}, \ \forall y \in \mathbb{R}_+, \tag{28}$$

or in other words, we need that the output $\mathbf{Y}$ is a zero-mean proper-complex Gaussian random variable. Such an output in additive models is only possible if the noise is Gaussian, which is only possible if $\mathsf{I} = 0$. Therefore, for all $\mathsf{I} > 0$ is it impossible for $F^{\text{opt}}_{\widetilde{X}}$ to have an accumulation point and therefore $F^{\text{opt}}_{\widetilde{X}}$ must have finitely many masses in any bounded interval. Thus, we ruled out case 1.

Rule out case 2 ($F^{\text{opt}}_{\widetilde{X}}$ has a finite number of points). We again proceed by contradiction. We assume that the number of masse points is finite, say given by an integer $M < +\infty$, with masses located at $0 \leq x_1 < \ldots < x_M < \infty$ and each occurring with probability $p_1, \ldots, p_M$, respectively. Then the output pdf corresponding to this specific input distribution is

$$f_{\widetilde{Y}}(y; F^{\text{opt}}_{\widetilde{X}}) = \sum_{i=1}^{M} p_i K(x_i, y) \tag{29a}$$

$$= \sum_{i=1}^{M} \int_{|b| \leq \pi} \frac{e^{-(y + x_i + \mathsf{I} + 2\sqrt{x_i \mathsf{I}} \cos b)}}{2\pi} \tag{29b}$$

$$\cdot I_0 \left(2\sqrt{y(x_i + \mathsf{I} + 2\sqrt{x_i \mathsf{I}} \cos b)} \right) db, \tag{29c}$$

where the expression in (29) is based on an equivalent way to write the kernel (10) (see eq.(36) in Appendix VI-A).

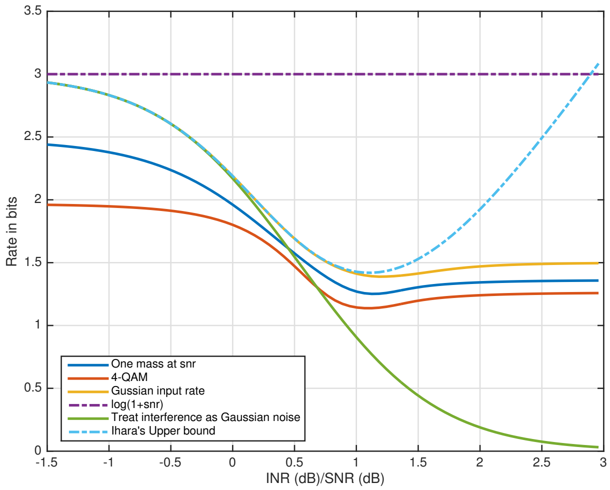

Fig. 1: Lower and upper bounds to the capacity vs I for $\mathsf{S} = 7 = 8.4510$dB.

With (29), one can bound the marginal entropy in (15) as

$$-h(x; F_{\widetilde{X}}^{\text{opt}}) = \int_{y\geq 0} K(x,y) \log f_{\widetilde{Y}}(y; F_{\widetilde{X}}^{\text{opt}}) dy \tag{30a}$$

$$\leq -\Big(x+\mathsf{I}+1+\log(2\pi)\Big)+\log\left(\sum_{i=1}^{M} p_i e^{-(\sqrt{x_i}+\sqrt{\mathsf{I}})^2}\right) \tag{30b}$$

$$+\int_{y\geq 0} K(x,y)\Big(2\sqrt{y}(\sqrt{x_M}+\sqrt{\mathsf{I}})\Big)\ dy, \tag{30c}$$

where the second term in (30b) is independent of x and hence we only need to deal with (30c). The term in (30c) can be bounded as

$$\mathbb{E}\left[\sqrt{\widetilde{Y}}\middle|\widetilde{X}=x\right] \leq \sqrt{\mathbb{E}\left[\widetilde{Y}\middle|\widetilde{X}=x\right]} = \sqrt{1+x+\mathsf{I}}, \tag{31}$$

where (31) follows from Jensen's inequality and by (20). With the bound in (30) back into the KKT condition in (23) we get

$$-x + c\sqrt{x} + \kappa_1 > -\lambda x + \kappa_2 \tag{32}$$

for some finite constants $c > 0, \kappa_1, \kappa_2$ that are not functions of x. However, as $x \to \infty$, and since we know that $\lambda < 1$, the right-hand-side of (32) grows faster than the left-hand-side, which is impossible. We reached a contradiction, which implies that the optimal number of mass points can not be finite. Thus, we ruled out case 2.

Having ruled out the possibility that $F_{\widetilde{X}}^{\text{opt}}$ has either infinitely many mass points in some bounded interval or is discrete with finitely many mass points, the only remaining option is that $F_{\widetilde{X}}^{\text{opt}}$ has countably infinitely many mass points, but only a finite number of masses in any bounded interval. This concludes the proof. ■

IV. Numerical Evaluations

In this section we evaluate achievable rates and compare them with the bound in Section III-A. In Fig.1 we plot achievable rates as function of I for fixed $\mathsf{S} = 7$:

- (red solid line) an equally likely 4-QAM constellation,
- (cyan solid line) a distribution with uniform phase and only one mass point at $\sqrt{\mathsf{S}}$ for the modulo,
- (orange solid line) a proper-complex Gaussian input, and
- (green solid line) treat the radar interference as Gaussian noise.

We also show the outer bound in (8) (blue dashed line) and the one in (7) (purple dashed line). We distinguish two regimes.

Low INR regime. The Gaussian input performs very well for $\alpha := \frac{\mathsf{I}_{(\text{dB})}}{\mathsf{S}_{(\text{dB})}} < 1$, where it follows closely the upper in (8), in comparison to the discrete 4-QAM input and a distribution with uniform phase and only one mass point at $\sqrt{\mathsf{S}}$ for the modulo. Although this behavior was expected for $\mathsf{I} \ll 1$ (actually a Gaussian input is optimal for $\mathsf{I} = 0$), it is very pleasing to see that it actually performs very well for the whole regime $\mathsf{I} \leq \mathsf{S}$.

High INR regime. For $\mathsf{I} \gg \mathsf{S}$, we see that the Gaussian input rate flattens at half the interference-free capacity because

$$\lim_{\mathsf{I}\to\infty} I(\mathbf{X}_G; \mathbf{Y}) = \lim_{\mathsf{I}\to\infty} h(\mathbf{Y}) - h(\mathbf{W})$$
$$= \lim_{\mathsf{I}\to\infty} \log(1+\mathsf{S}) + h\left(\sqrt{\frac{\mathsf{I}}{1+\mathsf{S}}}e^{j\Theta_{\mathsf{I}}} + \mathbf{Z}\right) - h\left(\sqrt{\mathsf{I}}e^{j\Theta_{\mathsf{I}}} + \mathbf{Z}\right)$$
$$= \lim_{\mathsf{I}\to\infty} \log(1+\mathsf{S}) + \frac{1}{2}\log\left(1+\frac{\mathsf{I}}{1+\mathsf{S}}\right) - \frac{1}{2}\log\left(1+\mathsf{I}\right)$$
$$= \frac{1}{2}\log(1+\mathsf{S}) + \lim_{\mathsf{I}\to\infty}\frac{1}{2}\log\left(1+\frac{\mathsf{S}}{1+\mathsf{I}}\right)$$
$$= \frac{1}{2}\log(1+\mathsf{S}),$$

since for $\mathsf{I} \gg 1$ the entropy of a non-central Chi-square random variable with 2 degress of freedom and non-centrality parameter I behaves as [12] $h(\mathbf{W}) = \frac{1}{2}\log(1+\mathsf{I}) + o(\mathsf{I})$. This, together with the upper bound in (9), implies that for $\mathsf{I} \gg 1$ the communication system has only $1/2$ degrees of freedom. Although this is a loss of $1/2$ degrees of freedom compared to the interference-free system, it is a substantial improvement from the zero rate achieved when communication in presence of radar signal is prohibited.

We note that the equally likely 4-QAM and the distribution with uniform phase and only one mass at $\sqrt{\mathsf{S}}$ for the modulo are only a 'constant gap' away from the the upper bound in (9) for the simulated $\mathsf{S} = 7$, which shows that capacity can be well approximated by inputs with a finite number of masses. The problem of designing efficient numerical routines to evaluate the capacity is currently under investigation. We note that the search for an optimal input is practically only needed in the regime $\mathsf{I} \geq \mathsf{S}$ where we could potentially gain at most 1 bit compared to the rate achieved by a proper-complex Gaussian input.

V. Conclusion

In this paper we studied the structural properties of the optimal (communication) input of a new channel model which models the impact of a high power, short duty cycle, wideband, radar interference on a narrowband communication signal. In particular, we showed that the optimal input distribution has uniform phase independent of the modulo, which is discrete with countably infinite mass points. We also argue that for large radar interference there is a loss of half the degrees of freedom compared to the the interference-free channel.

Acknowledgment

The work of the authors was partially funded by NSF under award 1443967. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

VI. Appendices

A. Derivation of the kernel $K(x,y)$ in (10)

By (4) and by passing to polar coordinates we have

$$K(x,y) := f_{\widetilde{Y}|\widetilde{X}}(y|x) \tag{33}$$

$$= \int_0^{2\pi} d\phi \int_0^{2\pi} \frac{d\alpha}{2\pi} \cdot \frac{e^{-|\sqrt{y}e^{j\phi}-\sqrt{x}e^{j\alpha}|^2-\mathsf{I}}}{2\pi} \cdot \cdot I_0\Big(2\sqrt{I}|\sqrt{y}e^{j\phi}-\sqrt{x}e^{j\alpha}|\Big) \tag{34}$$

$$= \int_{|b|\le\pi} \frac{e^{-(y+x+\mathsf{I}-2\sqrt{yx}\cos(b))}}{2\pi} \cdot I_0\Big(2\sqrt{\mathsf{I}}\sqrt{y+x-2\sqrt{yx}\cos(b)}\Big)db \tag{35}$$

$$= \int_{|b|\le\pi} \frac{e^{-(y+x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos(b))}}{2\pi} \cdot I_0\Big(2\sqrt{y}\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos(b)}\Big)db, \tag{36}$$

where the two expressions for the kernel $K(x,y)$ in (35) and (36) correspond to solving for the two integrals in different orders.

B. The map $F_{\widetilde{X}} \to h(\widetilde{Y};F_{\widetilde{X}})$ is weak* continuous

To prove the weak* continuity of the $h(\widetilde{Y};F_{\widetilde{X}})$ in (14), we have to show that for any sequence of distribution functions $\{F_n\}_{n=1}^\infty \in \mathcal{F}$ if $F_n \overset{w^*}{\to} F_{\widetilde{X}}$ then $h(\widetilde{Y};F_n) \to h(\widetilde{Y};F_{\widetilde{X}})$. In this regard we have

$$\lim_{n\to\infty} h(\widetilde{Y};F_n) = \lim_{n\to\infty}\int_{y\ge0} f_{\widetilde{Y}}(y;F_n)\log\frac{1}{f_{\widetilde{Y}}(y;F_n)}dy$$

$$= \int_{y\ge0}\lim_{n\to\infty} f_{\widetilde{Y}}(y;F_n)\log\frac{1}{f_{\widetilde{Y}}(y;F_n)}dy \tag{37}$$

$$= h(\widetilde{Y};F_{\widetilde{X}}), \tag{38}$$

where the exchange of limit and integral in (37) is due to the Dominated Convergence Theorem [13], and equality in (38) is due to continuity of the map $F_{\widetilde{X}} \to f_{\widetilde{Y}}(y;F_{\widetilde{X}})\log f_{\widetilde{Y}}(y;F_{\widetilde{X}})$. This last assertion is true by noting that $x \to x\log x$ is a continuous function of $x \in \mathbb{R}_+$ and $f_{\widetilde{Y}}(y;F_{\widetilde{X}})$ in (11) is a continuous function of $F_{\widetilde{X}}$ since $K(x,y)$ in (10) is a bounded continuous function of x for all $y \in \mathbb{R}_+$.

Back to (37), to satisfy the necessary condition required in the Dominated Convergence Theorem, we have to show that there exists an integrable function $g(y)$ such that

$$|\ f_{\widetilde{Y}}(y;F_n)\log f_{\widetilde{Y}}(y;F_n)\ | < g(y). \tag{39}$$

Lemma 1. *For any $\delta_1 > 0$ and $0 < x \le 1$*

$$0 \le -x\log x \le \frac{e^{-1}}{\delta_1}x^{1-\delta_1}. \tag{40}$$

Proof. Fix a $\delta_1 > 0$; the fuction $x \to -x^{\delta_1}\log x$ is concave in $0 < x \le 1$, and is maximized at $x = e^{-1/\delta_1}$. Hence $-x^{\delta_1}\log x \le \frac{e^{-1}}{\delta_1}$ and (40) follows. ☐

According to Lemma 1 we can write

$$|\ f_{\widetilde{Y}}(y;F_n)\log f_{\widetilde{Y}}(y;F_n)\ | \le \frac{e^{-1}}{\delta_1} f_{\widetilde{Y}}(y;F_n)^{1-\delta_1};$$

We next need to find $\Phi(y) : f_{\widetilde{Y}}(y;F_n) \le \Phi(y)$ such that

$$g(y) = \frac{e^{-1}}{\delta_1}\Phi(y)^{1-\delta_1} \tag{41}$$

is integrable for some $0 < \delta_1$. Similarly to [14, eq. A9] we can show that for any $\delta_2 > 0$

$$\Phi(y) = \begin{cases} 1 & y \le 16\mathsf{I} \\ \frac{M}{y^{1.5-\delta_2}} & y > 16\mathsf{I} \end{cases}, \tag{42}$$

is such a desirable upper bound for some $M < \infty$. The proof is as follows. For $y > 16\mathsf{I}$ we write

$$f_{\widetilde{Y}}(y;F_{\widetilde{X}}) = \int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2} K(x,y)dF_{\widetilde{X}}(x) \tag{43}$$

$$+ \int_{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}^{\infty} K(x,y)dF_{\widetilde{X}}(x). \tag{44}$$

The term in (43) can be upper bounded as

$$\int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2} K(x,y)dF_{\widetilde{X}}(x)$$

$$\le e^{-y}\int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}\int_0^{2\pi} \frac{e^{-(x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos b)}}{2\pi} \cdot \cdot I_0\Big(2\sqrt{y}(\sqrt{x}+\sqrt{\mathsf{I}})\Big)db\ dF_{\widetilde{X}}(x)$$

$$\le e^{-y}I_0\Big(2\sqrt{y}\frac{\sqrt{y}}{4}\Big) \cdot \int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}\int_0^{2\pi}\frac{e^{-(x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos b)}}{2\pi}db\ dF_{\widetilde{X}}(x)$$

$$\le e^{-y}I_0\big(y/2\big)\cdot 1 \le e^{-y/2}; \tag{45}$$

while in (44) can be upper bounded as

$$\int_{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}^{\infty} K(x,y)dF_{\widetilde{X}}(x)$$
$$\leq \mathbb{P}[\widetilde{X} > (\sqrt{y}/4-\sqrt{\mathsf{I}})^2]\cdot$$
$$\cdot \frac{e^{-y}}{2\pi}\int_0^{2\pi} \sup_{x_b>0}\left\{e^{-x_b} I_0\Big(2\sqrt{y}\sqrt{x_b}\Big)\right\}db$$
$$\leq \frac{e^{-y}}{2\pi}\int_0^{2\pi} \mathsf{S}\ \frac{\sup_{x_b>0}\left\{e^{-x_b} I_0\Big(2\sqrt{y}\sqrt{x_b}\Big)\right\}}{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}db \tag{46}$$
$$\leq e^{-y}\frac{3}{2}\frac{e^{y}}{\sqrt{4\pi y}}\Big[1+O(1/y)\Big]\frac{\mathsf{S}}{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}, \tag{47}$$

where $x_b := x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos(b)$, the inequality in (46) is from Markov's inequality, and the one in (47) is by [15, eq.(E.6)]. By (45) and (47), we have

$$f_{\widetilde{Y}}(y;F_n) \leq \frac{12\mathsf{S}}{\sqrt{\pi}}\Big[\frac{1}{y^{1.5}}+O(\frac{1}{y^{2.5}})\Big].$$

Hence, for any $0<\delta_2<1$ there exists some $M<\infty$ and $y^*_{\delta_2}$, such that

$$f_{\widetilde{Y}}(y;F_n) < \frac{M}{y^{1.5-\delta_2}}, \tag{48}$$

for all $y \geq y^*_{\delta_2}$. We fix δ_2 now. Due to continuity of the $f_{\widetilde{Y}}(y;F_n)$ in $[16\mathsf{I}, y^*_{\delta_2}]$, there exists an $M<\infty$ such that (48) holds for all $y>16\mathsf{I}$. (48) together with (17) gives

$$f_{\widetilde{Y}}(y;F_n) \leq \Phi(y),$$

for any $0<\delta_2<1$ and some $M<\infty$ and where $\Phi(y)$ was defined in (42). Finally, one can find small enough δ_1 and δ_2 such that $g(y)$ given in (41) is integrable.

C. The map $F_{\widetilde{X}} \to h(\widetilde{Y};F_{\widetilde{X}})$ is strictly concave

The function $h(\widetilde{Y};F_{\widetilde{X}})$ in (13) is concave in $f_{\widetilde{Y}}(y;F_{\widetilde{X}})$ in (11) (because $x \to -x\log(x)$ is). Since $f_{\widetilde{Y}}(y;F_{\widetilde{X}})$ is an injective function of $F_{\widetilde{X}}$ (due to invertibility of the kernel as proved in VI-G), we conclude that $h(\widetilde{Y};F_{\widetilde{X}})$ is a strictly concave function of $F_{\widetilde{X}}$.

D. Weak* differentiability of $F_{\widetilde{X}} \to h(\widetilde{Y};F_{\widetilde{X}}) - L(F_{\widetilde{X}})$

Using the definition of the functional derivative, we show that $h'_{F^{\rm opt}_{\widetilde{X}}}(\widetilde{Y};F_{\widetilde{X}})$ and $L'_{F^{\rm opt}_{\widetilde{X}}}(F_{\widetilde{X}})$ exist for all $F_{\widetilde{X}}, F^{\rm opt}_{\widetilde{X}}$ and hence $h(\widetilde{Y};F_{\widetilde{X}}) - L(F_{\widetilde{X}})$ is weak* differentiable.

First, for $\theta \in [0,1]$, we define $F_\theta := (1-\theta)F^{\rm opt}_{\widetilde{X}} + \theta F_{\widetilde{X}}$ and then we find the weak* derivative of $h(\widetilde{Y};F_{\widetilde{X}})$ at $F^{\rm opt}_{\widetilde{X}}$ as follows

$$h'_{F^{\rm opt}_{\widetilde{X}}}(\widetilde{Y};F_{\widetilde{X}}) = \lim_{\theta\to 0^+}\frac{1}{\theta}\big[h(\widetilde{Y};F_\theta) - h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}})\big]$$
$$= \lim_{\theta\to 0^+}\frac{1}{\theta}\int_{x\geq 0}\int_{y\geq 0} K(x,y)\log\frac{1}{f_{\widetilde{Y}}(y;F_\theta)}dy\ dF_\theta(x)$$
$$- \lim_{\theta\to 0^+}\frac{1}{\theta}\int_{x\geq 0}\int_{y\geq 0} K(x,y)\log\frac{1}{f_{\widetilde{Y}}(y;F^{\rm opt}_{\widetilde{X}})}dy\ dF^{\rm opt}_{\widetilde{X}}(x)$$
$$= \int_{x\geq 0} h(x;F^{\rm opt}_{\widetilde{X}})dF_{\widetilde{X}}(x) - h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}})$$
$$- \int_{y\geq 0}\lim_{\theta\to 0^+}\frac{1}{\theta}f_{\widetilde{Y}}(y;F_\theta)\log\frac{f_{\widetilde{Y}}(y;F_\theta)}{f_{\widetilde{Y}}(y;F^{\rm opt}_{\widetilde{X}})}dy, \tag{49}$$

where the interchange of limit and integral in (49) is due to Dominated Convergence Theorem. By [16, Lemma 6], we can write

$$\Big|\frac{f_{\widetilde{Y}}(y;F_\theta)}{\theta}\log\frac{f_{\widetilde{Y}}(y;F_\theta)}{f_{\widetilde{Y}}(y;F^{\rm opt}_{\widetilde{X}})}\Big| \leq f_{\widetilde{Y}}(y;F_{\widetilde{X}}) + f_{\widetilde{Y}}(y;F^{\rm opt}_{\widetilde{X}})$$
$$- f_{\widetilde{Y}}(y;F_{\widetilde{X}})\log f_{\widetilde{Y}}(y;F_{\widetilde{X}}) - f_{\widetilde{Y}}(y;F_{\widetilde{X}})\log f_{\widetilde{Y}}(y;F^{\rm opt}_{\widetilde{X}})$$
$$\leq f_{\widetilde{Y}}(y;F_{\widetilde{X}}) + f_{\widetilde{Y}}(y;F^{\rm opt}_{\widetilde{X}}) + 2f_{\widetilde{Y}}(y;F_{\widetilde{X}})(y+\mathsf{I}+\mathsf{S}), \tag{50}$$

where the right hand side of (50) is integrable. In addition, the term given in (49) is vanishing by L'Hospital's Rule. Hence, the weak* derivative is given by

$$h'_{F^{\rm opt}_{\widetilde{X}}}(\widetilde{Y};F_{\widetilde{X}}) = \int_{x\geq 0} h(x;F^{\rm opt}_{\widetilde{X}})dF_{\widetilde{X}}(x) - h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}}). \tag{51}$$

It is also easy to show that

$$L'_{F^{\rm opt}_{\widetilde{X}}}(F_{\widetilde{X}}) = L(F_{\widetilde{X}}) - L(F^{\rm opt}_{\widetilde{X}}), \tag{52}$$

exists because of the linearity of the power constraint.

E. KKT conditions

Let $\mathcal{E}^{\rm opt}$ be the set of points of increase of the optimal input distribution $F^{\rm opt}_{\widetilde{X}}$. Then

$$\int_{x\geq 0}\Big(h(x;F^{\rm opt}_{\widetilde{X}}) - \lambda x\Big)dF_{\widetilde{X}}(x) \leq h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}}) - \lambda\mathsf{S} \tag{53}$$

for all $F_{\widetilde{X}} \in \mathcal{F}$ if and only if

$$h(x;F^{\rm opt}_{\widetilde{X}}) \leq h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}}) + \lambda(x-\mathsf{S}),\ \forall x\in\mathbb{R}_+, \tag{54}$$
$$h(x;F^{\rm opt}_{\widetilde{X}}) = h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}}) + \lambda(x-\mathsf{S}),\ \forall x\in\mathcal{E}^{\rm opt}. \tag{55}$$

The *if* direction is trivial since the derivative given in (51) has to be non-positive. To prove the *only if* direction, assume that (54) is false. Then there exists an $\widetilde{x}$ such that

$$h(\widetilde{x};F^{\rm opt}_{\widetilde{X}}) > h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}}) + \lambda(\widetilde{x}-\mathsf{S}).$$

If $F^{\rm opt}_{\widetilde{X}}$ is a unit step function at $\widetilde{x}$, then

$$\int_{x\geq 0}\Big(h(x;F^{\rm opt}_{\widetilde{X}}) - \lambda x\Big)dF_{\widetilde{X}}(x) = h(\widetilde{x},F^{\rm opt}_{\widetilde{X}}) - \lambda\widetilde{x}$$
$$> h(\widetilde{Y};F^{\rm opt}_{\widetilde{X}}) - \lambda\mathsf{S},$$

which contradicts (53). Now assume that (54) holds but (55) does not, i.e., there exists $\widetilde{x} \in \mathcal{E}^{\text{opt}}$ such that

$$h(\widetilde{x}; F_{\widetilde{X}}^{\text{opt}}) < h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}) + \lambda(\widetilde{x} - \mathsf{S}). \tag{56}$$

Since all functions in (56) are continuous in x, the inequality is satisfied strictly on a neighborhood of $\widetilde{x}$ indicated as $E_{\widetilde{x}}$. Since $\widetilde{x}$ is a point of increase, the set $E_{\widetilde{x}}$ has nonzero measure, i.e., $\int_{E_{\widetilde{x}}} dF_{\widetilde{X}}^{\text{opt}}(x) = \delta > 0$; hence

$$\begin{aligned}
h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}) - \lambda\mathsf{S} &= \int_{x\geq 0} \left(h(x; F_{\widetilde{X}}^{\text{opt}}) - \lambda x\right) dF_{\widetilde{X}}^{\text{opt}}(x) \\
&= \int_{E_{\widetilde{x}}} \left(h(x; F_{\widetilde{X}}^{\text{opt}}) - \lambda x\right) dF_{\widetilde{X}}^{\text{opt}}(x) \\
&+ \int_{\mathcal{E}^{\text{opt}}\setminus E_{\widetilde{x}}} \left(h(x; F_{\widetilde{X}}^{\text{opt}}) - \lambda x\right) dF_{\widetilde{X}}^{\text{opt}}(x) \\
&< \delta(h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}) - \lambda\mathsf{S}) + (1-\delta)(h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}) - \lambda\mathsf{S}),
\end{aligned}$$

which is a contradiction.

F. The function $z \to g(z, \lambda)$ is analytic

The analyticity of $g(z, \lambda),\ z \in \mathbb{C} : \Re\{z\} > 0$ follows from the analyticity of $h(z; F_{\widetilde{X}})$ on the same domain, where $h(x; F_{\widetilde{X}})$ was defined in (15). In other words, we want to show that the function

$$h(z; F_{\widetilde{X}}) = \int_{y\geq 0} K(z, y) \log \frac{1}{f_{\widetilde{Y}}(y; F_{\widetilde{X}})}\ dy,\ \ z \in \mathbb{C}, \tag{57}$$

is analytic through the domain $\{z \in \mathbb{C} : \Re\{z\} > 0\}$. Note that the integrand in (57) is a continuous function on $\{z \in \mathbb{C} : \Re\{z\} > 0\} \times \{y \in \mathbb{R}_+\}$ and analytic for each y so we use the Differentiation Lemma [11] to prove the analyticity by proving that $h(x; F_{\widetilde{X}})$ is uniformly convergent for any rectangle $K := \{z \in \mathbb{C} : 0 < a \leq \Re(z) \leq b, -b \leq \Im(z) \leq b,\}$ (since any compact set $K \in \mathbb{C}$ is closed and bounded in the complex plane). By (17) we have

$$\left|\log f_{\widetilde{Y}}(y; F_{\widetilde{X}})\right| \leq y + \mathsf{I} + \beta_{F_{\widetilde{X}}},$$

and as a result we have

$$\begin{aligned}
&|h(z; F_{\widetilde{X}})| \leq \int_{y\geq 0} |K(z, y)|\ |\log f_{\widetilde{Y}}(y; F_{\widetilde{X}})|\ dy \\
&\leq \int_{y\geq 0} \frac{1}{2\pi} \int_{|b|\leq\pi} \left|e^{-(z+y-2\sqrt{zy}\cos b+\mathsf{I})}\right| \\
&\cdot \left|I_0\left(2\sqrt{\mathsf{I}(z+y-2\sqrt{zy}\cos b)}\right)\right| \cdot \left|y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\right|\ db\ dy \\
&\leq \int_{y\geq 0} \frac{1}{2\pi} \int_{|b|\leq\pi} e^{-\Re(z+y-2\sqrt{zy}\cos b+\mathsf{I})} \\
&\cdot I_0\left(2\Re\left\{\sqrt{\mathsf{I}(z+y-2\sqrt{zy}\cos b)}\right\}\right) \left(y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\right)\ db\ dy \\
&\leq \int_{y\geq 0} \frac{1}{2\pi} \int_{|b|\leq\pi} e^{-\Re(y+z-2\sqrt{zy}\cos b+\mathsf{I})} \\
&\qquad \cdot e^{2\Re\{\sqrt{\mathsf{I}(z+y-2\sqrt{zy}\cos b)}\}} \left(y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\right)\ db\ dy \\
&= \int_{y\geq 0} \frac{1}{2\pi} \int_{|b|\leq\pi} e^{-(\sqrt{\Re(y+z-2\sqrt{zy}\cos b)}-\sqrt{\mathsf{I}})^2} \\
&\qquad \cdot \left(y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\right)\ db\ dy.
\end{aligned} \tag{58}$$

Since (58) is exponentially decreasing in $y \in \mathbb{R}_+$, the integral is bounded. This concludes the proof.

G. Invertibility of integral transform in (11)

To prove the invertibility of transform

$$\breve{g}(y) = \int_{x\geq 0} K(x, y) g(x)\ dx,\ \ y \in \mathbb{R}_+, \tag{59}$$

we will show that if

$$\breve{g}(y) \equiv 0 \text{ for all } y \in \mathbb{R}_+,$$

then

$$g(x) \equiv 0 \text{ for all } x \in \mathbb{R}_+.$$

From the invertibility of (59), also the integral transform $\int_{y\geq 0} K(x, y) g(y) dy$ is invertible due to the symmetry of the kernel $K(x, y)$ in (10) in x and y.

We first define the following two integrals [17, eq(6.633) and eq(6.684)]

$$\int_0^\infty e^{-\alpha y} I_\nu(\beta\sqrt{y}) J_\nu(\gamma\sqrt{y})\ dy = \frac{1}{2\alpha} \exp\left(\frac{\beta^2 - \gamma^2}{4\alpha}\right) J_0\left(\frac{\beta\gamma}{2\alpha}\right),$$
$$\Re\{\alpha\} > 0, \Re\{\nu\} > -1 \tag{60}$$

where $J_0(.)$ is the zero order Bessel function of the first kind, and

$$\begin{aligned}
&\int_0^\pi (\sin\theta)^{2\nu} \frac{J_\nu\left(\sqrt{\alpha^2 + \beta^2 - 2\alpha\beta\cos\theta}\right)}{\left(\sqrt{\alpha^2 + \beta^2 - 2\alpha\beta\cos\theta}\right)^\nu}\ d\theta \\
&= 2^\nu \sqrt{\pi}\Gamma(\nu + \frac{1}{2}) \frac{J_\nu(\alpha)}{\alpha^\nu} \frac{J_\nu(\beta)}{\beta^\nu}, \quad \Re\{\nu\} > -\frac{1}{2},
\end{aligned} \tag{61}$$

where $\Gamma(.)$ is the Gamma function.

We next use (60) and (61) as follows. If $\breve{g}(y) = 0$ for all $y \geq 0$, then for all $\gamma \geq 0$ we have

$$\int_0^\infty J_0(\gamma\sqrt{y}) \breve{g}(y)\ dy = 0 \tag{62}$$
$$\iff \int_0^\infty g(x) dx \int_0^\pi J_0\left(\gamma\sqrt{x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos\theta}\right)\ d\theta = 0 \tag{63}$$
$$\iff \int_0^\infty g(x) J_0(\gamma\sqrt{x}) J_0(\gamma\sqrt{\mathsf{I}})\ dx = 0 \tag{64}$$
$$\iff \int_0^\infty g(z^2) J_0(\gamma z) z\ dz = 0 \tag{65}$$
$$\iff \mathcal{H}\{g(z^2)\} = 0, \tag{66}$$
$$\iff g(z^2) = 0,\ \forall z \in \mathbb{R}_+, \tag{67}$$
$$\iff g(x) = 0,\ \forall x \in \mathbb{R}_+, \tag{68}$$

where (63) follows by (60) and (64) by (61), and where $\mathcal{H}\{g(z)\}$ in (66) denotes the Hankel transform [18] of the function $g(z)$.

References

[1] S. Ihara, "On the capacity of channels with additive non-gaussian noise," *Information and Control*, vol. 37, no. 1, pp. 34 – 39, 1978. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0019995878904138

[2] R. Zamir and U. Erez, "A gaussian input is not too bad," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1362–1367, June 2004.

[3] 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. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0019995871903469

[4] A. Tchamkerten, "On the discreteness of capacity-achieving distributions," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2773–2778, Nov 2004.

[5] J. Fahs, N. Ajeeb, and I. Abou-Faycal, "The capacity of average power constrained additive non-gaussian noise channels," in *Telecommunications (ICT), 2012 19th International Conference on*, April 2012, pp. 1–6.

[6] 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.

[7] N. Nartasilpa, D. Tuninetti, N. Devroye, and D. Erricolo, "Lets Share CommRad: Effect of Radar Interference on Uncoded Data Communication Systems," in *IEEE Radar Conference (RadarCon)*, Philadelphia, May 2016.

[8] ——, "Lets Share CommRad: Effect of Radar Interference on Complex Constellations," in *submitted to IEEE Global Communications Conference (Globecom)*, Washington D.C., Dec. 2016.

[9] 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.

[10] D. G. Luenberger, *Optimization by vector space methods*. John Wiley & Sons, 1969.

[11] S. Lang, "Complex analysis (graduate texts in mathematics)," 1993.

[12] A. Lapidoth, "On phase noise channels at high snr," in *Information Theory Workshop, 2002. Proceedings of the 2002 IEEE*, Oct 2002, pp. 1–4.

[13] S. I. Resnick, *A probability path*. Springer Science & Business Media, 2013.

[14] M. Katz and S. Shamai, "On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent awgn channels," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2257–2270, Oct 2004.

[15] M. Katz, "On the capacity-achieving distribution of the discrete-time non-coherent and partially-coherent additive white gaussian noise channels," Master's thesis, Technion-Israel Institute of Technology, Haifa, Israel, 2002.

[16] L. Zhang, H. Li, and D. Guo, "Capacity of gaussian channels with duty cycle and power constraints," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1615–1629, March 2014.

[17] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic press, 2007.

[18] L. Debnath and D. Bhatta, *Integral transforms and their applications*. CRC press, 2014.