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 from interference caused by 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 infinitely 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. We also show that for very large radar interference, and for signal to noise ratio equal to S, the capacity is equal to $\frac{1}{2}\log(1+\mathsf{S})$ and a proper-complex Gaussian input achieves it. It is concluded that the presence of the radar interference results in a loss of half of the degrees of freedom compared to an AWGN channel without radar interference.

I. Introduction

Shortage of spectrum resources, coupled 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 the release of portions of government radar bands (e.g., 3550-3700 MHz) to be shared with commercial wireless services. A new Citizens Broadband Radio Service (CBRS) for shared wireless broadband in the 3550-3700 MHz band has also been established in 2015. 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 channels in which two systems co-exist. In this work, the capacity of a white additive Gaussian noise communications channel, which in addition to noise, suffers interference from a radar transmission, is studied. This extends the authors' prior work in [1]. In the channel model considered, the interfering radar transmission is modeled to be additive, but not Gaussian. Rather, it is modeled as a constant (and known) amplitude signal, but with unknown and uniformly distributed phase at each channel use. The effect of the rectangular-shaped pulsed radar at a OFDM receiver is studied in [2] where it is shown that the marginal PMF of the radar amplitude is dominated by a single mass point and that the conditional PMF of the radar phase at the dominant amplitude mass point is uniform between $[0, 2\pi]$. Hence, the model we are considering here is a good approximation for the interference caused by a rectangular-shaped pulsed radar signal passing through an OFDM receiver.

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 under various input constraints has been studied.

In [3] Ihara bounds the capacity of additive, but not necessarily Gaussian, noise channels. 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-constrained Gaussian noise channel without the radar interference.

In [4, Theorem 1], it 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 the 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 our channel model however, we show that in the large radar interference regime, the capacity is indeed achieved by a proper-complex Gaussian input and hence the bound given in [4, Theorem 1] is loose for large radar interference.

In the seminal work by Smith [5], 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 the Gaussian input that achieves the capacity when the amplitude constraint is dropped. In [6], the authors also considered complex-valued Gaussian noise with average power and peak amplitude constraints and derived the optimal input distribution characteristics. In particular, they showed that the capacity of the complex-valued Gaussian noise channel under average power and peak amplitude constraints is 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.

Later, the optimality of a discrete input under peak amplitude constraint was shown to hold for a wide class of real-valued additive noise channels [7]. As for the complex-valued additive channels, [8] showed that for certain additive complex-valued channels with average power and peak amplitude constraints, the optimal input has discrete modulo. Moreover, recently it was shown in [9], that, under an average power constraint and certain 'smoothness' conditions on the noise distribution, the only real-valued additive noise channel whose capacity achieving input is continuous is the Gaussian noise channel. More recently, [10] studied the capacity of the additive quadrature Gaussian mixture noise under an average power constraint and showed that the capacity achieving input has discrete amplitude with only finitely many mass points and uniform phase between $[0, 2\pi]$. In addition, [11] proposes a method based on piecewise linear curve fitting to compute the achievable rates by a Gaussian input and an input with discrete amplitude and uniform phase in channels with additive Gaussian mixture noise.

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 (i.e., the kernel $K(x, y)$, which is formally defined in (10)) is no longer additive. For this equivalent non-additive channel, we can *not* proceed as per the steps preceding [9, eq.(4)]. This is so, because [9, 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] and we closely follow the steps within that. In particular, the trick used in [6] to reduce the two dimensional optimization problem into a one dimensional optimization problem helps us to avoid the use of an identity theorem for multiple dimensions. In fact, it was shown in [12] that the application of an identity theorem in several variables is not just a straightforward generalization of one variable.

Extensions of Smith's work [5] 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 [13], [14] a subset of the authors studied the uncoded average symbol error rate performance of the same channel model considered here. 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, part of the useful signal is also removed, and 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 the similar 'half degrees of freedom' loss 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 fundamental limits 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.

By upper bounding the output entropy by the cross entropy of the output distribution and an output distribution induced by Gaussian input, we show that very high radar interference results in a loss of half the degrees of freedom compared to an interference-free channel and that a Gaussian input is optimal in the high interference regime.

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 [3]. We numerically find some suboptimal inputs for the regime $0 < \alpha := \frac{\mathsf{I}_{(\mathrm{dB})}}{\mathsf{S}_{(\mathrm{dB})}} < 2$, where I and S are the interference and signal to noise ratios respectively, which perform better than the Gaussian input.

C. *Paper organization*

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

II. Problem formulation and preliminary results

Next, boldface letters indicate complex-valued random variables, while lightface letters real-valued ones. Capital letters represent the random variables and the lower case letters represent their realization. In addition, $\widetilde{X}$ and $\angle\mathbf{X}$ are respectively used to indicate the modulus squared, $|\mathbf{X}|^2$, and phase of the complex-valued random variable $\mathbf{X}$. We also define the following notation

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

$$\mathbb{C}_+ := \{z : z \in \mathbb{C},\ \Re(z) > 0\},$$

for the set of non-negative real line and right half plane in the complex domain, respectively.

A. *System model*

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 the 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 known. For later use, the distribution of the additive noise in (2) is given by

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

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

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

B. Channel Capacity

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

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

where $F_{\mathbf{X}}$ is the cumulative distribution function of $\mathbf{X}$, and $I(\mathbf{X};\mathbf{Y})$ denotes the mutual information between random variables $\mathbf{X}$ and $\mathbf{Y}$ in (1).

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 (1) (involving the real and the imaginary part of the input) in a way that allows optimization with respect to a univariate distribution only.

We now show that an optimal input distribution induces $\widetilde{Y}$ 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 $\widetilde{X}$ as follows.

We represent $\mathbf{X}$ and $\mathbf{Y}$ in their Polar forms as $\mathbf{X} = \sqrt{\widetilde{X}}e^{j\angle\mathbf{X}}$ and $\mathbf{Y} = \sqrt{\widetilde{Y}}e^{j\angle\mathbf{Y}}$ which are related as (for details please refer to VII-A)

$$f_{\widetilde{Y},\angle\mathbf{Y}|\widetilde{X},\angle\mathbf{X}}(y,\phi|x,\alpha) = \frac{e^{-\left(\mathsf{I}+y+x-2\sqrt{xy}\cos(\alpha-\phi)\right)}}{2\pi} I_0\left(2\sqrt{\mathsf{I}}\sqrt{y+x-2\sqrt{xy}\cos(\alpha-\phi)}\right), \tag{6}$$

and where the marginal distribution of $(\widetilde{Y},\angle\mathbf{Y})$ is given by

$$f_{\widetilde{Y},\angle\mathbf{Y}}(y,\phi) = \int_0^\infty \int_0^{2\pi} f_{\widetilde{Y},\angle\mathbf{Y}|\widetilde{X},\angle\mathbf{X}}(y,\phi|x,\alpha)\ df_{\widetilde{X},\angle\mathbf{X}}(x,\alpha). \tag{7}$$

Due to the change of variables we have

$$f_{\widetilde{Y},\angle\mathbf{Y}}(y,\phi) = \frac{1}{2} f_{\mathbf{Y}}(\sqrt{y}\cos\phi, \sqrt{y}\sin\phi)$$

and hence

$$\begin{aligned} h(\mathbf{Y}) &= h(\widetilde{Y}, \angle\mathbf{Y}) - \log(2) \\ &\leq h(\widetilde{Y}) + h(\angle\mathbf{Y}) - \log(2) && (8) \\ &\leq h(\widetilde{Y}) + \log(2\pi) - \log(2), && (9) \end{aligned}$$

where (8) holds with equality for independent $\widetilde{Y}$ and $\angle\mathbf{Y}$ and where (9) holds with equality for uniform $\angle\mathbf{Y}$. It can be seen from (6) and (7) that choosing $\widetilde{X}$ independent of $\angle\mathbf{X}$ and $\angle\mathbf{X}$ uniform over $[0, 2\pi]$ satisfies both (8) and (9) with equality and hence is the optimal choice for the input distribution.

Therefore, it is convenient for later use to denote the transition probability $f_{\widetilde{Y}|\widetilde{X}}(y|x)$ as the kernel $K(x, y)$ given by (see Appendix VII-A)

$$\begin{aligned} K(x,y) &:= f_{\widetilde{Y}|\widetilde{X}}(y|x) && (10a) \\ &= \int_{|\theta|\leq\pi} \frac{e^{-\mathsf{I}-\xi(\theta;x,y)}}{2\pi} I_0\left(2\sqrt{\mathsf{I}\ \xi(\theta;x,y)}\right)\ d\theta, && (10b) \end{aligned}$$

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

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

$$f_{\widetilde{Y}}(y; F_{\widetilde{X}}) := \int_{x\geq 0} K(x, y) dF_{\widetilde{X}}(x), \quad y \in \mathbb{R}_+. \qquad (11)$$

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

[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.

optimization over the distribution of a non-negative random variable $\widetilde{X}$

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

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

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

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

$$h(\widetilde{Y};F_{\widetilde{X}}) = \int_{y\ge 0} f_{\widetilde{Y}}(y;F_{\widetilde{X}}) \log \frac{1}{f_{\widetilde{Y}}(y;F_{\widetilde{X}})}\ dy. \tag{13}$$

The equivalence of (12a) and (12b) is by noting that the optimal input phase distribution is known to be uniform and the optimal input induces an output with independent modulo and phase, with uniform phase in $[0,2\pi]$. We express $h(\widetilde{Y};F_{\widetilde{X}})$ in (13) as

$$\begin{aligned} h(\widetilde{Y};F_{\widetilde{X}}) &= \int_{y\ge 0}\int_{x\ge 0} K(x,y) \log \frac{1}{f_{\widetilde{Y}}(y;F_{\widetilde{X}})}\ dF_{\widetilde{X}}(x)\ dy \\ &= \int_{x\ge 0} h(x;F_{\widetilde{X}})\ dF_{\widetilde{X}}(x), \end{aligned} \tag{14}$$

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

$$h(x;F_{\widetilde{X}}) := \int_{y\ge 0} K(x,y) \log \frac{1}{f_{\widetilde{Y}}(y;F_{\widetilde{X}})}\ dy,\ \ x\in\mathbb{R}_+, \tag{15}$$

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)

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

[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).

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

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

for the marginal entropy in (15)

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

where the conditional mean of $\widetilde{Y}$ given $\widetilde{X}$ is

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

C. Trivial Bounds

The Gaussian distribution maximizes the entropy for a given power and hence the Gaussian interference is the worst interference among all other interference distributions. Trivially, one can lower bound the capacity in (5) by treating the radar interference as a Gaussian noise. The capacity of the equivalent Gaussian interference channel with noise power $1+\mathsf{I}$ is a lower bound to the capacity of the channel with non-Gaussian interference and is given by $\log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right)$ and hence

$$\log\left(1 + \frac{\mathsf{S}}{1+\mathsf{I}}\right) \leq \mathsf{C}(\mathsf{S}, \mathsf{I}), \tag{21}$$

and upper bound it as

$$\mathsf{C}(\mathsf{S}, \mathsf{I}) \leq \max_{F_{\mathbf{X}}:\mathbb{E}[|\mathbf{X}|^2]\leq \mathsf{S}} I(\mathbf{X}; \mathbf{Y}, \Theta_{\mathsf{I}}) = \log\left(1 + \mathsf{S}\right), \tag{22}$$

or from Ihara's work [3] as

$$\mathsf{C}(\mathsf{S}, \mathsf{I}) \leq \log\left(\pi e(1 + \mathsf{S} + \mathsf{I})\right) - h(\mathbf{W}), \tag{23}$$

or from Zamir and Erez's work [4, Theorem 1], as

$$\mathsf{C}(\mathsf{S}, \mathsf{I}) \leq I(\mathbf{X}_G; \mathbf{Y}) + \log(2), \tag{24}$$

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

We shall use these bounds later to benchmark the achievable performance in Section V.

III. 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 (12c).

Theorem 1. *The optimal input distribution in* (5) *is 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 II-B, an optimal input distribution in (5) has $\angle \mathbf{X}$ uniformly distributed in $[0, 2\pi]$ and independent of $\widetilde{X}$. The modulo squared $\widetilde{X}$, solves the problem in (12c), whose supremum is attained by the *unique* input distribution $F_{\widetilde{X}}^{\text{opt}}$, because (see [15, Theorem 1]):

1) the space of input distributions $\mathcal{F}$ is compact and convex (see [15, Theorem 1]); $\mathcal{F}$ is given by

$$\mathcal{F} := \Big\{ F_{\widetilde{X}} : F_{\widetilde{X}}(x) = 0, \ \forall x < 0, \tag{25a}$$

$$dF_{\widetilde{X}}(x) \geq 0, \ \forall x \geq 0, \tag{25b}$$

$$\int_{x \geq 0} 1 \cdot dF_{\widetilde{X}}(x) = 1, \tag{25c}$$

$$L(F_{\widetilde{X}}) := \int_{x \geq 0} x \cdot dF_{\widetilde{X}}(x) - \mathsf{S} \leq 0 \Big\}, \tag{25d}$$

where the various constraints are: (25a) for non-negativity, (25b) and (25c) for a valid input distribution, and (25d) for the average power constraint; and

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

From this and by following Smith's approach [5], the solution of the optimization problem in (12c) is equivalent to the solution of

$$h'_{F_{\widetilde{X}}^{\text{opt}}}(\widetilde{Y}; F_{\widetilde{X}}) - \lambda L'_{F_{\widetilde{X}}^{\text{opt}}}(F_{\widetilde{X}}) \leq 0, \ \text{for all } F_{\widetilde{X}} \in \mathcal{F}, \tag{26a}$$

$$\lambda \geq 0 : L(F_{\widetilde{X}}^{\text{opt}}) = 0, \tag{26b}$$

where the functional $L(.)$ was defined in (25d), and where the prime sign along with the subscript $F_{\widetilde{X}}^{\text{opt}}$ denotes the weak* derivative of the function $h(\widetilde{Y}; F_{\widetilde{X}})$ at $F_{\widetilde{X}}^{\text{opt}}$ [5] (see Appendix VII-D).

The conditions in (26) can be equivalently expressed as the necessary and sufficient Karush-Kuhn-Tucker (KKT) condition (see Appendix VII-E): for some $\lambda \geq 0$

$$h(x; F_{\widetilde{X}}^{\text{opt}}) \leq h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}) + \lambda(x - \mathsf{S}), \quad \forall x \in \mathbb{R}_+, \tag{27a}$$

$$h(x; F_{\widetilde{X}}^{\text{opt}}) = h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}) + \lambda(x - \mathsf{S}), \quad \forall x \in \mathcal{E}^{\text{opt}}, \tag{27b}$$

where $\mathcal{E}^{\text{opt}}$ is the set of the points of increase of the optimal input distribution $F_{\widetilde{X}}^{\text{opt}}$. We can further restrict the feasible values of λ in (27) to $0 \leq \lambda < 1$ since by the Envelope Theorem [16] and the upper bound in (22), having $\lambda \geq 1$ is not possible.

At this point, as it is usual in these types of problems [5], the proof follows by ruling out different types of distributions. A distribution takes 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; or
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 ($F_{\widetilde{X}}^{\text{opt}}$ has an infinite number of mass points in some bounded interval). We prove that this case requires the inequality in (27) to hold with equality for all $x \geq 0$; we then prove this to be impossible.

We start by handling the case $0 < \lambda < 1$ and then tackle the case where $\lambda = 0$.

For the range $0 < \lambda < 1$, we re-write the KKT condition in (27) by following the recent work [9]. Given the conditional output power expressed as in (20), we can write

$$x - \mathsf{S} = \int_{y \geq 0} (y - (1 + \mathsf{I} + \mathsf{S}))\, K(x, y)\ dy,\ \forall x \in \mathbb{R}_+. \tag{28}$$

With (28), the KKT condition in (27) reads: there exists a constant $0 < \lambda < 1$ such that

$$g(x, \lambda) \leq h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}) = \text{constant for all } x \in \mathbb{R}_+, \tag{29}$$

with equality if $x \in \mathcal{E}^{\text{opt}}$ and where

$$g(x, \lambda) := \int_{y \geq 0} K(x, y) \log \left(\frac{\lambda e^{-\lambda y}}{f_{\widetilde{Y}}(y; F_{\widetilde{X}}^{\text{opt}})} \right)\ dy \tag{30a}$$

$$+ \log \frac{1}{\lambda} + \lambda(1 + \mathsf{I} + \mathsf{S}). \tag{30b}$$

We show next that (29) can not be satisfied if $F_{\widetilde{X}}^{\text{opt}}$ 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 (30) can be extended to the complex domain and that $g(z,\lambda),\ z \in \mathbb{C}_+$, is analytic.

Remark 1. *In this type of analysis, we only require the analyticity of the function $g(z,\lambda)$ over a region in the complex domain which contains the non-negative real line. Hence, it is sufficient to prove the analyticity of $g(z,\lambda)$ over a strip around the non-negative real line but we prove it over the entire right half plane (see Appendix VII-F).*

Since the analytical function $g(z,\lambda)$ is equal to a constant at the set of points of increase of $F_{\widetilde{X}}^{\text{opt}}$ and since the set of points of increase of $F_{\widetilde{X}}^{\text{opt}}$ has an accumulation point (by the Bolzano Weierstrass Theorem [17]), by the Identity Theorem [17], we conclude that $g(z,\lambda) =$ constant, $\forall z \in \mathbb{C}_+$. As the result $g(x,\lambda) =$ constant, $\forall x \in \mathbb{R}_+$. One solution, and the *only solution* due to invertibility of the integral transform with kernel $K(x,y)$ (see Appendix VII-G), for $g(x,\lambda)$ to be a constant and not to depend on x is that the function that multiplies the kernel in the integral in (30a) 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

$$f_{\widetilde{Y}}(y; F_{\widetilde{X}}^{\text{opt}}) = \lambda e^{-\lambda y},\ \forall y \in \mathbb{R}_+, \tag{31}$$

or in other words, we need the output $\mathbf{Y}$ to be a zero-mean proper-complex Gaussian random variable. By Cramer's decomposition Theorem, such an output in additive models is only possible if both the noise and the input are Gaussian. The noise is Gaussian if and only if $\mathsf{I} = 0$. Therefore, unless $\mathsf{I} = 0$, it is impossible for $F_{\widetilde{X}}^{\text{opt}}$ to have an infinite number of mass points in some bounded interval for some $0 < \lambda < 1$.

We now consider the case $\lambda = 0$. In this case, we can re-write the KKT conditions in (27) as

$$h(x; F_{\widetilde{X}}^{\text{opt}}) \leq h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}), \quad \forall x \in \mathbb{R}_+, \tag{32}$$

with equality if $x \in \mathcal{E}^{\text{opt}}$. With the proof of analyticity of $h(z; F_{\widetilde{X}}^{\text{opt}})$ (see Appendix VII-F) and by the same type of argument used for $0 < \lambda < 1$ (use of Bolzano Weiestrass Theorem and Identity Theorem), we conclude that (32) should hold with equality for all $x \in \mathbb{R}_+$ and hence

$$h(x; F_{\widetilde{X}}^{\text{opt}}) = h(\widetilde{Y}; F_{\widetilde{X}}^{\text{opt}}), \forall x \in \mathbb{R}_+. \tag{33}$$

The only solution to (33) due to invertibility of the integral transform with kernel $K(x, y)$ over the space of all functions (see Appendix VII-G) is

$$f_{\widetilde{Y}}(y; F_{\widetilde{X}}^{\text{opt}}) = e^{-h(\widetilde{Y}, F_{\widetilde{X}}^{\text{opt}})}.$$

This unique solution is not a valid probability distribution since it does not integrate to one. Hence we reached a contradiction, which implies that if $F_{\widetilde{X}}^{\text{opt}}$ has infinite number of mass points in some bounded interval then having $\lambda = 0$ is impossible.

Having revoked the possibility of $0 < \lambda < 1$ and case 1 and $\lambda = 0$ and case 1, we rule out the possiblity of case 1 altogether.

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

$$\begin{aligned} f_{\widetilde{Y}}(y; F_{\widetilde{X}}^{\text{opt}}) &= \sum_{i=1}^{M} p_i^\star K(x_i^\star, y) \\ &= \sum_{i=1}^{M} p_i^\star \int_{|\theta| \le \pi} \frac{e^{-(y + x_i^\star + \mathsf{I} + 2\sqrt{x_i^\star \mathsf{I}} \cos\theta)}}{2\pi} \quad \cdot I_0\left(2\sqrt{y(x_i^\star + \mathsf{I} + 2\sqrt{x_i^\star \mathsf{I}} \cos\theta)}\right) d\theta, \end{aligned} \tag{34}$$

where the expression in (34) is based on an equivalent way to write the kernel in (10) (see eq.(42b) in Appendix VII-A). With (34), one can bound the marginal entropy in (15) as

$$-h(x; F_{\widetilde{X}}^{\text{opt}}) = \int_{y \ge 0} K(x, y) \log f_{\widetilde{Y}}(y; F_{\widetilde{X}}^{\text{opt}}) dy \tag{35a}$$

$$\le -(x + \mathsf{I} + 1 + \log(2\pi)) + \log\left(\sum_{i=1}^{M} p_i^\star e^{-(\sqrt{x_i^\star} + \sqrt{\mathsf{I}})^2}\right) \tag{35b}$$

$$+ \int_{y \ge 0} K(x, y) \left(2\sqrt{y}(\sqrt{x_M^\star} + \sqrt{\mathsf{I}})\right) \, dy, \tag{35c}$$

where the second term in (35b) is independent of x and hence we only need to deal with (35c). The term in (35c) can be bounded as

$$\mathbb{E}\left[\sqrt{\widetilde{Y}} \middle| \widetilde{X} = x\right] \le \sqrt{\mathbb{E}\left[\widetilde{Y} \middle| \widetilde{X} = x\right]} = \sqrt{1 + x + \mathsf{I}}, \tag{36}$$

where (36) follows from Jensen's inequality and by (20). With the bound in (35) back into the KKT condition in (27) we get

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

or equivalently

$$c\sqrt{x} + \kappa_1 > (1-\lambda)x + \kappa_2, \tag{37}$$

for some finite constants $c > 0, \kappa_1, \kappa_2$ that are not functions of x. However, as $x \to \infty$, and since $0 \le \lambda < 1$, the right-hand-side of (37) 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. Capacity at high INR

In this section, we prove that in the high INR regime, the communication system has only 1/2 the degrees of freedom compared to the interference-free system, which is a substantial improvement from the zero rate achieved when communication in presence of radar signal is prohibited. We also show that the Gaussian input is asymptotically optimal as $\mathsf{I} \to \infty$.

Theorem 2. *The capacity of channel* (5) *as* $\mathsf{I} \to \infty$ *is given by*

$$\lim_{\mathsf{I}\to\infty} \mathsf{C}(\mathsf{S},\mathsf{I}) = \frac{1}{2}\log(1+\mathsf{S}).$$

Proof: We show that in the high INR regime, the mutual information between the input and the output is upper bounded by $\frac{1}{2}\log(1+\mathsf{S})$ for *any* input distribution subject to an average power constraint. We then show that the Gaussian input can asymptotically achieve this upper bound as $\mathsf{I} \to \infty$. We write:

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &= h(\mathbf{Y}) - \mathbf{h}(\mathbf{W}) \\ &= h(\widetilde{Y}) - h(\widetilde{W}) && \text{(38a)} \\ &\le \int_{y\ge 0} f_{\widetilde{Y}}(y)\log\frac{1}{R(y)}\, dy - h(\widetilde{W}), && \text{(38b)} \end{aligned}$$

where (38a) is because $\mathbf{Y}$ and $\mathbf{W}$ are circularly symmetric, (38b) is due to non-negativity of relative entropy and where $R(y)$ is an auxiliary output density function which is absolutely continuous with respect to $f_{\widetilde{Y}}(y)$. Take

$$R(y) = \frac{1}{\mathsf{S}+1} e^{-\left(\frac{y+\mathsf{I}}{\mathsf{S}+1}\right)} I_0\left(2\frac{\sqrt{y\mathsf{I}}}{\mathsf{S}+1}\right), \tag{39}$$

to be the auxiliary output distribution in (38b). The intuition behind this choice of $R(y)$ lies behind our conjecture that the Gaussian input is optimal for large INR and the fact that (39) is the induced distribution on $\widetilde{Y}$ by a proper-complex Gaussian input. Then by (38b) we have

$$\lim_{\mathsf{I}\to\infty} I(\mathbf{X};\mathbf{Y}) \le \lim_{\mathsf{I}\to\infty} \int_{x\ge 0, y\ge 0} K(x,y) \log\left(\frac{(\mathsf{S}+1)e^{\frac{y+\mathsf{I}}{\mathsf{S}+1}}}{I_0(2\frac{\sqrt{y\mathsf{I}}}{\mathsf{S}+1})}\right) \, dy \; dF_{\widetilde{X}}(x) - \frac{1}{2}\log(4\pi e\mathsf{I}) \tag{40a}$$

$$= \log(\mathsf{S}+1) + \lim_{\mathsf{I}\to\infty}\left\{\frac{\mathsf{S}+2\mathsf{I}+1}{\mathsf{S}+1} - \int_{x\ge 0, y\ge 0} K(x,y)\log\left(\frac{e^{\frac{2\sqrt{y\mathsf{I}}}{\mathsf{S}+1}}}{\sqrt{4\pi\frac{\sqrt{y\mathsf{I}}}{\mathsf{S}+1}}}\right) \, dy \; dF_{\widetilde{X}}(x)\right\} - \frac{1}{2}\log(4\pi e\mathsf{I}) \tag{40b}$$

$$= \frac{1}{2}\log(\mathsf{S}+1) + \lim_{\mathsf{I}\to\infty}\left\{\frac{\mathsf{S}+2\mathsf{I}+1}{\mathsf{S}+1} - 2\frac{\sqrt{\mathsf{I}}}{\mathsf{S}+1}\mathbb{E}\left[\sqrt{\widetilde{Y}}\right] + \frac{1}{4}\log(\mathsf{I}) + \frac{1}{4}\mathbb{E}\left[\log(\widetilde{Y})\right] - \frac{1}{2}\log(e\mathsf{I})\right\}$$

$$\le \frac{1}{2}\log(\mathsf{S}+1) + \frac{\mathsf{S}+1}{\mathsf{S}+1} - \frac{2}{\mathsf{S}+1}\left[\frac{\mathsf{S}+1}{4}\right] - \frac{1}{2} \tag{40c}$$

$$= \frac{1}{2}\log(\mathsf{S}+1), \tag{40d}$$

where (40a) is by [18, eq. (9)] and where (40b) and (40c) are proved in Appendix VII-H and Appendix VII-I, respectively. Next, a Gaussian input can achieve the upper bound given in (40d), as follows

$$\lim_{\mathsf{I}\to\infty} I(\mathbf{X}_G;\mathbf{Y}) = \lim_{\mathsf{I}\to\infty} h(\mathbf{Y}) - h(\mathbf{W}) \tag{41a}$$

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

$$= \lim_{\mathsf{I}\to\infty} \log(1+\mathsf{S}) + \frac{1}{2}\log\left(\frac{\mathsf{I}}{1+\mathsf{S}}\right) - \frac{1}{2}\log\left(\mathsf{I}\right) \tag{41b}$$

$$= \frac{1}{2}\log(1+\mathsf{S}) + \lim_{\mathsf{I}\to\infty}\frac{1}{2}\log\left((1+\mathsf{S})\left(\frac{\mathsf{I}}{1+\mathsf{S}}\right)\left(\frac{1}{\mathsf{I}}\right)\right)$$

$$= \frac{1}{2}\log(1+\mathsf{S}),$$

where $\mathbf{X}_G$ is the proper-complex Gaussian input and where (41b) is again by [18, eq. (9)]. ■

V. Numerical Evaluations

In this section, we numerically find a sub-optimal input for fixed $\mathsf{S} = 5$ and three different values of I in the regime $0 < \alpha := \frac{\mathsf{I}_{(\text{dB})}}{\mathsf{S}_{(\text{dB})}} < 2$ and we compare the achieved rates with that of a proper-complex Gaussian input. We also evaluate different achievable rates in the regime $-1 \leq \alpha \leq 2.5$ and compare them with the bound in Section II-C.

Numerically finding the optimal input for the channel considered in this paper is more challenging compared to channels with finite dimensional capacity achieving inputs such as the ones considered in [5] and [6]. In [5], for example, the optimization was initially performed for a very low SNR where an input with two mass points was proved to be optimal. As SNR increased, more mass points were added to the optimization problem in order to satisfy the KKT conditions and guarantee the optimality of the input. In the channel considered here however, a finite number of mass points is sub-optimal at *any* SNR. Hence, in the rest of this section, we find sub-optimal inputs with a finite number of mass points and solving the corresponding constraint optimization problem.

We used the constrained optimization toolbox in Matlab to produce our numerical results. Specifically, for each value of SNR, we start off by setting the number of input mass points to be equal to $n = 1$ and use Matlab's constrained optimization function to find the location and the weight of the mass points. This is a constrained optimization problem consisting of $2n$ parameters subject to the constraints given in (25). For each value of n we calculate the rate achieved by the result of the optimization problem. We increase the number of mass points n until the achieved rate remains unchanged after the 3rd digit after the decimal point. Figure 1 shows the location of the mass points for each sub-optimal input as a function of SNR. We note that we do not claim the rates achieved with these inputs to be optimal, nor do we claim that these input distributions are capacity-optimal. It is however interesting to note that they can outperform Gaussian inputs.

We find the achievable inputs for fixed $\mathsf{S} = 5$ and three different values of $\mathsf{I} = [3.6239, 9.5183, 25]$ which correspond to $\alpha = [0.8, 1.4, 2]$, by solving a finite dimensional constraint optimization problem. The achievable rates obtained by a Gaussian input and optimized finite dimensional inputs are given in Table I. As it can be seen, the optimized finite dimensional inputs achieve marginally better rates than the Gaussian input.

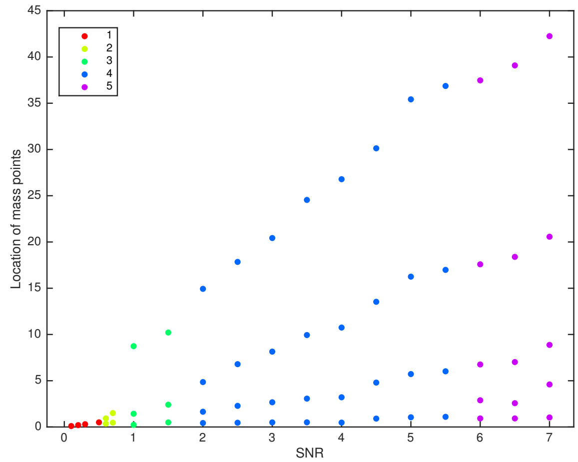


Fig. 1: Location of mass points for sub-optimal input as a function of SNR for fixed INR=5.

TABLE I: Achievable rates for Gaussian and optimized finite dimensional input, $\mathsf{S} = 5$.

Input \ INR	$\mathsf{S}^{0.8} = 3.6239$	$\mathsf{S}^{1.4} = 9.5183$	$\mathsf{S}^{2} = 25$
Gaussian	1.2905	1.1910	1.2470
Optimized finite dimension	1.2927	1.1922	1.2480

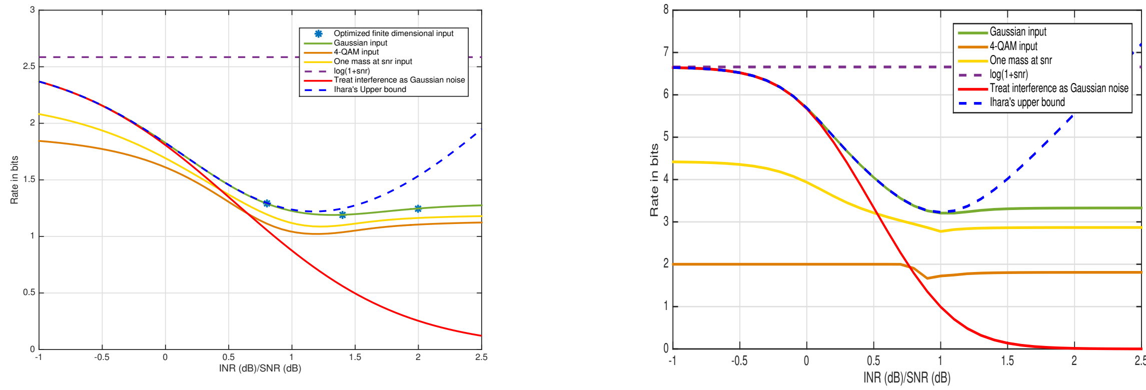
TABLE II: Achievable rates for Gaussian and optimized finite dimensional input, $\mathsf{S} = 10$.

Input \ INR	$\mathsf{S}^{0.8} = 6.3096$	$\mathsf{S}^{1.4} = 25.1189$	$\mathsf{S}^{2} = 100$
Gaussian	1.6986	1.6393	1.7100
Optimized finite dimension	1.7108	1.6398	1.7102

In Fig.2a we plot achievable rates as function of I for fixed $\mathsf{S} = 5$:

- (orange solid line) an equally likely 4-QAM constellation,
- (yellow solid line) a distribution with uniform phase and only one mass point at $\sqrt{\mathsf{S}}$ for the modulo,
- (green solid line) a proper-complex Gaussian input,
- (red solid line) treat the radar interference as Gaussian noise as given in (21), and
- (stared cyan points) optimized finite dimensional input.

We also show the outer bound in (23) (blue dashed line) and the one in (22) (purple dashed



(a) Lower and upper bounds to the capacity vs INR(dB)/SNR(dB) for fixed $\mathsf{S} = 5 = 6.9897$dB.

(b) Lower and upper bounds to the capacity vs INR(dB)/SNR(dB) for fixed $\mathsf{S} = 100 = 20$dB.

Fig. 2: Achievable rates and upper and lower bounds on capacity.

line).

The Gaussian input performs very well for $\alpha := \frac{\mathsf{I}_{(\text{dB})}}{\mathsf{S}_{(\text{dB})}} < 1$, where it closely follows the upper in (23), 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 $\mathsf{I} \leq \mathsf{S}$.

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 (24) for the simulated $\mathsf{S} = 5$ in Figure 2a (and the same trend is true for the simulated $\mathsf{S} = 100$ in Figure 2b), which shows that capacity can be well approximated by inputs with a finite number of masses. The rate achieved by optimized finite dimensional input at $\mathsf{I} = \mathsf{S}^{0.8}, \mathsf{S}^{1.4}$ and S^2 is only slightly higher than the rate achieved by a proper-complex Gaussian input in terms of bits per channel use. However, the difference in the performance of Gaussian inputs and the finite dimensional optimized inputs in terms of bits per second will be proportional to the difference between their achieved rates times the system bandwidth.

VI. 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 many mass points. We also argue that for large radar interference there is a loss of half the degrees of freedom compared to 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.

VII. Appendices

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

By (4) and by passing to Polar coordinates $\mathbf{Y} = \sqrt{\widetilde{Y}}e^{j\angle\mathbf{Y}}$ and $\mathbf{X} = \sqrt{\widetilde{X}}e^{j\angle\mathbf{X}}$ we have

$$\begin{aligned}
f_{\widetilde{Y},\angle\mathbf{Y}|\widetilde{X},\angle\mathbf{X}}(y,\phi|x,\alpha) &= \mathbb{E}_{\Theta}\left[\frac{e^{-\left|\sqrt{y}e^{j\phi}-\sqrt{x}e^{j\alpha}-\sqrt{\mathsf{I}}e^{j\theta}\right|}}{2\pi}\right] \\
&= \frac{1}{2\pi}\int_{\theta=0}^{2\pi}\frac{e^{-\left(\mathsf{I}+\left|\sqrt{y}e^{j\phi}-\sqrt{x}e^{j\alpha}\right|^2+2\sqrt{\mathsf{I}}\left|\sqrt{y}e^{j\phi}-\sqrt{x}e^{j\alpha}\right|\cos\left(\theta-\angle\left(\sqrt{y}e^{j\phi}-\sqrt{x}e^{j\alpha}\right)\right)\right)}}{2\pi}d\theta \\
&= \frac{e^{-\left(\mathsf{I}+y+x-2\sqrt{xy}\cos(\phi-\alpha)\right)}}{2\pi}I_0\left(2\sqrt{\mathsf{I}}\sqrt{y+x-2\sqrt{xy}\cos(\phi-\alpha)}\right)d\theta.
\end{aligned}$$

For the case that $\widetilde{X}$ and $\angle\mathbf{X}$ are independent and $\angle\mathbf{X}$ is uniform between $[0,2\pi]$, we have

$$\begin{aligned}
K(x,y) &:= f_{\widetilde{Y}\big|\widetilde{X}}(y|x) \\
&= \int_0^{2\pi}d\phi\int_0^{2\pi}\frac{d\alpha}{2\pi}\int_0^{2\pi}\frac{d\theta'}{2\pi}\frac{e^{-\left|\sqrt{y}e^{j\phi}-\sqrt{x}e^{j\alpha}-\sqrt{\mathsf{I}}e^{j\theta'}\right|^2}}{2\pi} \\
&= \int_{|\theta|\le\pi}\frac{e^{-(y+x+\mathsf{I}-2\sqrt{yx}\cos(\theta))}}{2\pi}I_0\left(2\sqrt{\mathsf{I}}\sqrt{y+x-2\sqrt{yx}\cos(\theta)}\right)d\theta \qquad (42a) \\
&= \int_{|\theta|\le\pi}\frac{e^{-(y+x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos(\theta))}}{2\pi}I_0\left(2\sqrt{y}\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos(\theta)}\right)d\theta, \qquad (42b)
\end{aligned}$$

where (42a) and (42b) correspond to solving for the two integrals in different orders and using the definition of the modified Bessel function.

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

The function $h(\widetilde{Y}; F_{\widetilde{X}})$ in (14) is weak* continuous if for any sequence of distribution functions $\{F_n\}_{n=1}^{\infty} \in \mathcal{F}$ such that $F_n \xrightarrow{w^*} F_{\widetilde{X}}$, then $h(\widetilde{Y}; F_n) \to h(\widetilde{Y}; F_{\widetilde{X}})$. In this regard, we have

$$\begin{aligned}
\lim_{n\to\infty} h(\widetilde{Y}; F_n) &= \lim_{n\to\infty} \int_{y\geq 0} f_{\widetilde{Y}}(y; F_n) \log \frac{1}{f_{\widetilde{Y}}(y; F_n)} dy \\
&= \int_{y\geq 0} \lim_{n\to\infty} f_{\widetilde{Y}}(y; F_n) \log \frac{1}{f_{\widetilde{Y}}(y; F_n)} dy \qquad (43a) \\
&= h(\widetilde{Y}; F_{\widetilde{X}}), \qquad (43b)
\end{aligned}$$

where the exchange of limit and integral in (43a) is due to the Dominated Convergence Theorem [19], and equality in (43b) 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 (43a), 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

$$\mid f_{\widetilde{Y}}(y; F_n) \log f_{\widetilde{Y}}(y; F_n) \mid < g(y), \;\; \forall y \in \mathbb{R}_+. \qquad (44)$$

We state the following Lemma which is a generalization of the one given in [20, Lemma A.2].

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

$$0 \leq -x \log x \leq \frac{e^{-1}}{\delta_1} x^{1-\delta_1}. \qquad (45)$$

Proof. Fix a $\delta_1 > 0$; the fuction $x \to -x^{\delta_1} \log x$ is concave in $0 < x \leq 1$, and is maximized at $x = e^{-1/\delta_1}$. Hence $-x^{\delta_1} \log x \leq \frac{e^{-1}}{\delta_1}$ and (45) follows. □

According to Lemma 1 we can write

$$\mid f_{\widetilde{Y}}(y; F_n) \log f_{\widetilde{Y}}(y; F_n) \mid \leq \frac{e^{-1}}{\delta_1} f_{\widetilde{Y}}(y; F_n)^{1-\delta_1}.$$

We next need to find $\Phi(y) : f_{\widetilde{Y}}(y; F_n) \leq \Phi(y)$ which would then lead to

$$g(y) = \frac{e^{-1}}{\delta_1} \Phi(y)^{1-\delta_1}, \qquad (46)$$

which is integrable for some $0 < \delta_1$. Similarly to [21, eq. A9] we can show that for any $\delta_2 > 0$

$$\Phi(y) = \begin{cases} 1 & y \leq 16\mathsf{I} \\ \frac{M}{y^{1.5-\delta_2}} & y > 16\mathsf{I} \end{cases}, \tag{47}$$

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

$$f_{\widetilde{Y}}(y; F_{\widetilde{X}}) = \int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2} K(x,y) dF_{\widetilde{X}}(x) + \int_{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}^{\infty} K(x,y) dF_{\widetilde{X}}(x). \tag{48}$$

The first term in (48) can be upper bounded as

$$\begin{aligned} &\int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2} K(x,y) dF_{\widetilde{X}}(x) \\ &\leq e^{-y} \int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2} \int_0^{2\pi} \frac{e^{-(x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta)}}{2\pi} \cdot I_0\left(2\sqrt{y}(\sqrt{x}+\sqrt{\mathsf{I}})\right) d\theta \; dF_{\widetilde{X}}(x) \\ &\leq e^{-y} I_0\left(2\sqrt{y}\frac{\sqrt{y}}{4}\right) \int_0^{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2} \int_0^{2\pi} \frac{e^{-(x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta)}}{2\pi} d\theta \; dF_{\widetilde{X}}(x) \\ &\leq e^{-y} I_0\left(y/2\right) \cdot 1 \leq e^{-y/2}, \end{aligned} \tag{49}$$

while the second term in (48) can be upper bounded as

$$\begin{aligned} \int_{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}^{\infty} K(x,y) dF_{\widetilde{X}}(x) &\leq \mathbb{P}[\widetilde{X} > (\sqrt{y}/4-\sqrt{\mathsf{I}})^2] \cdot \frac{e^{-y}}{2\pi} \int_0^{2\pi} \sup_{x_\theta > 0} \left\{ e^{-x_\theta} I_0\left(2\sqrt{y}\sqrt{x_\theta}\right) \right\} d\theta \\ &\leq \frac{e^{-y}}{2\pi} \int_0^{2\pi} \mathsf{S} \; \frac{\sup_{x_\theta>0} \left\{ e^{-x_\theta} I_0\left(2\sqrt{y}\sqrt{x_\theta}\right) \right\}}{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2} d\theta \qquad (50a) \\ &\leq e^{-y} \frac{3}{2} \frac{e^y}{\sqrt{4\pi y}} \left[1 + O(1/y)\right] \frac{\mathsf{S}}{(\sqrt{y}/4-\sqrt{\mathsf{I}})^2}, \qquad (50b) \end{aligned}$$

where $x_\theta := x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos(\theta)$, the inequality in (50a) is from Markov's inequality, and the one in (50b) is by [22, eq.(E.6)]. By (49) and (50b), we have

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

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

$$f_{\widetilde{Y}}(y; F_n) < \frac{M}{y^{1.5-\delta_2}}, \tag{51}$$

for all $y \geq y^*_{\delta_2}$. We fix δ_2 now. Due to continuity of the $f_{\widetilde{Y}}(y; F_n)$ for $y \in [16\mathsf{l}, y^*_{\delta_2}]$, there exists an $M < \infty$ such that (51) holds for all $y > 16\mathsf{l}$. The bound in (51) together with the one in (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 (47). Finally, one can find small enough δ_1 and δ_2 such that $g(y)$ given in (46) 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 strictly 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 linear function of $F_{\widetilde{X}}$ (due to invertibility of the kernel as proved in Appendix VII-G), we conclude that $h(\widetilde{Y}; F_{\widetilde{X}})$ is a strictly concave function of $F_{\widetilde{X}}$.

D. The functional $h(\widetilde{Y}; F_{\widetilde{X}}) - L(F_{\widetilde{X}})$, is weakly* differentiable at $F^{opt}_{\widetilde{X}}$

A function $h : \mathcal{F} \to \mathbb{R}$ for the convex space $\mathcal{F}$, is said to be weakly* differentiable at F_0 if

$$h'_{F_0}(F) = \lim_{\theta \to 0^+} \frac{h\left((1-\theta)F_0 + \theta F\right) - h(F_0)}{\theta}$$

exists for all $F \in \mathcal{F}$. By using the above definition, we show that $h'_{F^{\text{opt}}_{\widetilde{X}}}(\widetilde{Y}; F_{\widetilde{X}})$ and $L'_{F^{\text{opt}}_{\widetilde{X}}}(F_{\widetilde{X}})$ exist for all $F_{\widetilde{X}}, F^{\text{opt}}_{\widetilde{X}}$ and hence $h(\widetilde{Y}; F_{\widetilde{X}}) - L(F_{\widetilde{X}})$ is weakly* differentiable.

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

$$\begin{aligned}
h'_{F^{\text{opt}}_{\widetilde{X}}}(\widetilde{Y}; F_{\widetilde{X}}) &= \lim_{\theta \to 0^+} \frac{1}{\theta}\left[h(\widetilde{Y}; F_\theta) - h(\widetilde{Y}; F^{\text{opt}}_{\widetilde{X}})\right] \\
&= \lim_{\theta \to 0^+} \frac{1}{\theta} \int_{x \geq 0} \int_{y \geq 0} K(x, y) \log \frac{1}{f_{\widetilde{Y}}(y; F_\theta)} dy\ dF_\theta(x) \\
&- \lim_{\theta \to 0^+} \frac{1}{\theta} \int_{x \geq 0} \int_{y \geq 0} K(x, y) \log \frac{1}{f_{\widetilde{Y}}(y; F^{\text{opt}}_{\widetilde{X}})} dy\ dF^{\text{opt}}_{\widetilde{X}}(x) \\
&= \int_{x \geq 0} h(x; F^{\text{opt}}_{\widetilde{X}}) dF_{\widetilde{X}}(x) - h(\widetilde{Y}; F^{\text{opt}}_{\widetilde{X}}) \\
&- \int_{y \geq 0} \lim_{\theta \to 0^+} \frac{1}{\theta} f_{\widetilde{Y}}(y; F_\theta) \log \frac{f_{\widetilde{Y}}(y; F_\theta)}{f_{\widetilde{Y}}(y; F^{\text{opt}}_{\widetilde{X}})} dy, \qquad (52)
\end{aligned}$$

where the interchange of limit and integral in (52) is due to Dominated Convergence Theorem. By [23, Lemma 6], we can write

$$\Big|\frac{f_{\widetilde{Y}}(y;F_\theta)}{\theta}\log\frac{f_{\widetilde{Y}}(y;F_\theta)}{f_{\widetilde{Y}}(y;F_{\widetilde{X}}^{\mathrm{opt}})}\Big| \le f_{\widetilde{Y}}(y;F_{\widetilde{X}}) + f_{\widetilde{Y}}(y;F_{\widetilde{X}}^{\mathrm{opt}})$$

$$- f_{\widetilde{Y}}(y;F_{\widetilde{X}})\log f_{\widetilde{Y}}(y;F_{\widetilde{X}}) - f_{\widetilde{Y}}(y;F_{\widetilde{X}})\log f_{\widetilde{Y}}(y;F_{\widetilde{X}}^{\mathrm{opt}})$$

$$\le f_{\widetilde{Y}}(y;F_{\widetilde{X}}) + f_{\widetilde{Y}}(y;F_{\widetilde{X}}^{\mathrm{opt}}) + 2 f_{\widetilde{Y}}(y;F_{\widetilde{X}})(y+\mathsf{I}+\mathsf{S}), \tag{53}$$

where the right hand side of (53) is integrable. In addition, the term given in (52) is vanishing by L'Hospital's Rule. Hence, the weak* derivative is given by

$$h'_{F_{\widetilde{X}}^{\mathrm{opt}}}(\widetilde{Y};F_{\widetilde{X}}) = \int_{x\ge 0} h(x;F_{\widetilde{X}}^{\mathrm{opt}})dF_{\widetilde{X}}(x) - h(\widetilde{Y};F_{\widetilde{X}}^{\mathrm{opt}}). \tag{54}$$

It is also easy to show that

$$L'_{F_{\widetilde{X}}^{\mathrm{opt}}}(F_{\widetilde{X}}) = L(F_{\widetilde{X}}) - L(F_{\widetilde{X}}^{\mathrm{opt}}), \tag{55}$$

exists because of the linearity of the power constraint.

E. Equivalence of KKT conditions in (27) *to* (26)

We proceed as the proof of [15, Theorem 4]. Then

$$\int_{x\ge 0}\Big(h(x;F_{\widetilde{X}}^{\mathrm{opt}}) - \lambda x\Big)\, dF_{\widetilde{X}}(x) \le h(\widetilde{Y};F_{\widetilde{X}}^{\mathrm{opt}}) - \lambda\mathsf{S} \tag{56}$$

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

$$h(x;F_{\widetilde{X}}^{\mathrm{opt}}) \le h(\widetilde{Y};F_{\widetilde{X}}^{\mathrm{opt}}) + \lambda(x-\mathsf{S}),\ \forall x\in\mathbb{R}_+, \tag{57}$$

$$h(x;F_{\widetilde{X}}^{\mathrm{opt}}) = h(\widetilde{Y};F_{\widetilde{X}}^{\mathrm{opt}}) + \lambda(x-\mathsf{S}),\ \forall x\in\mathcal{E}^{\mathrm{opt}}. \tag{58}$$

The *if* direction is trivial since the derivative given in (54) has to be non-positive. To prove the *only if* direction, assume that (57) is false. Then there exists an $\widetilde{x}$ such that

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

If $F_{\widetilde{X}}^{\mathrm{opt}}$ is a unit step function at $\widetilde{x}$, then

$$\int_{x\ge 0}\Big(h(x;F_{\widetilde{X}}^{\mathrm{opt}}) - \lambda x\Big)\, dF_{\widetilde{X}}(x) = h(\widetilde{x},F_{\widetilde{X}}^{\mathrm{opt}}) - \lambda\widetilde{x} > h(\widetilde{Y};F_{\widetilde{X}}^{\mathrm{opt}}) - \lambda\mathsf{S},$$

which contradicts (56). Assume that (57) holds but (58) does not, i.e., there exists $\widetilde{x} \in \mathcal{E}^{\text{opt}}$:

$$h(\widetilde{x}; F^{\text{opt}}_{\widetilde{X}}) < h(\widetilde{Y}; F^{\text{opt}}_{\widetilde{X}}) + \lambda(\widetilde{x} - \mathsf{S}). \tag{59}$$

Since all functions in (59) 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., $\int_{E_{\widetilde{x}}} dF^{\text{opt}}_{\widetilde{X}}(x) = \delta > 0$; hence

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

which is a contradiction.

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

The analyticity of $g(z, \lambda)$, $z \in \mathbb{C}_+$, 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

$$h(z; F_{\widetilde{X}}) = \int_{y\geq 0} K(z, y) \log \frac{1}{f_{\widetilde{Y}}(y; F_{\widetilde{X}})} \ dy, \ z \in \mathbb{C}_+, \tag{60}$$

is analytic. Note that the integrand in (60) is a continuous function on $\{z \in \mathbb{C}_+\} \times \{y \in \mathbb{R}_+\}$ and analytic for each y so we use the Differentiation Lemma [17] to prove the analyticity by proving that $h(x; F_{\widetilde{X}})$ is uniformly convergent for any rectangle $K := \{z \in \mathbb{C} : 0 \leq a \leq \Re(z) \leq b, -b \leq \Im(z) \leq b\}$ (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

$$\begin{aligned} &|h(z; F_{\widetilde{X}})| \leq \int_{y\geq 0} |K(z, y)| \ |\log f_{\widetilde{Y}}(y; F_{\widetilde{X}})| \ dy \\ &\leq \int_{y\geq 0} \frac{1}{2\pi} \int_{|\theta|\leq\pi} \left|e^{-(z+y-2\sqrt{zy}\cos\theta+\mathsf{I})}\right| \cdot \left|I_0\left(2\sqrt{\mathsf{I}(z+y-2\sqrt{zy}\cos\theta)}\right)\right| \cdot \left|y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\right| \ d\theta \ dy \\ &\leq \int_{y\geq 0} \frac{1}{2\pi} \int_{|\theta|\leq\pi} e^{-\Re(z+y-2\sqrt{zy}\cos\theta+\mathsf{I})} \cdot I_0\left(2\Re\Big\{\sqrt{\mathsf{I}(z+y-2\sqrt{zy}\cos\theta)}\Big\}\right) \ \Big(y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\Big) \ d\theta \ dy \\ &\leq \int_{y\geq 0} \frac{1}{2\pi} \int_{|\theta|\leq\pi} e^{-\Re(y+z-2\sqrt{zy}\cos\theta+\mathsf{I})} \cdot e^{2\Re\{\sqrt{\mathsf{I}(z+y-2\sqrt{zy}\cos\theta)}\}} \left(y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\right) \ d\theta \ dy \\ &= \int_{y\geq 0} \frac{1}{2\pi} \int_{|\theta|\leq\pi} e^{-(\sqrt{\Re(y+z-2\sqrt{zy}\cos\theta)}-\sqrt{\mathsf{I}})^2} \cdot \left(y + \mathsf{I} + \beta_{F_{\widetilde{X}}}\right) \ d\theta \ dy. \end{aligned} \tag{61}$$

Since (61) is exponentially decreasing in $y \in \mathbb{R}_+$, the integral is bounded, concluding the proof.

G. Invertibility of the integral transform in (11)

In order to prove the invertibility of the integral transforms in (11) and (30a), we use the following general notation

$$\breve{h}(y) = \int_{x \geq 0} K(x, y) \, \mathrm{d}h(x), \;\; y \in \mathbb{R}_+, \tag{62}$$

in which $h(x)$ may represent the difference of two distribution functions or just a function. We will show that if $\breve{h}(y) \equiv 0$ for all $y \in \mathbb{R}_+$, then $\mathrm{d}h(x) \equiv 0$ for all $x \in \mathbb{R}_+$. From the invertibility of (62), also the integral transform $\int_{y \geq 0} K(x, y) \, \mathrm{d}h(y)$ is invertible due to the symmetry of the kernel $K(x, y)$ in x and y.

We first define the following two integrals [24, eq(6.633) and eq(6.684)]

$$\int_0^\infty e^{-\alpha y} I_\nu(\beta \sqrt{y}) J_\nu(\gamma \sqrt{y}) \; dy = \frac{1}{2\alpha} \exp\left(\frac{\beta^2 - \gamma^2}{4\alpha}\right) J_0\left(\frac{\beta\gamma}{2\alpha}\right), \Re\{\alpha\} > 0, \Re\{\nu\} > -1, \tag{63}$$

$$\int_0^\pi (\sin\theta)^{2\nu} \frac{J_\nu\left(\sqrt{\alpha^2 + \beta^2 - 2\alpha\beta\cos\theta}\right)}{\left(\sqrt{\alpha^2 + \beta^2 - 2\alpha\beta\cos\theta}\right)^\nu} \; d\theta = 2^\nu \sqrt{\pi} \Gamma\left(\nu + \frac{1}{2}\right) \frac{J_\nu(\alpha)}{\alpha^\nu} \frac{J_\nu(\beta)}{\beta^\nu}, \Re\{\nu\} > -\frac{1}{2}, \tag{64}$$

where $J_\nu(.)$ and $I_\nu(.)$ are the ν-th order Bessel function of the first kind and ν-th order modified Bessel function of the first kind, and where $\Gamma(.)$ is the Gamma function.

We next use (63) and (64) as follows. If $\breve{h}(y)=0$ for all $y\geq 0$, then for all $\gamma\geq 0$ we have

$$\begin{aligned}
&\int_0^\infty J_0(\gamma\sqrt{y})\breve{h}(y)\ dy = 0\\
\Longleftrightarrow &\int_0^\infty \mathrm{d}h(x)\int_0^\pi J_0\left(\gamma\sqrt{x+\mathsf{l}+2\sqrt{x\mathsf{l}}\cos\theta}\right)\ d\theta = 0 && \text{(65a)}\\
\Longleftrightarrow &\int_0^\infty J_0(\gamma\sqrt{x})J_0(\gamma\sqrt{\mathsf{l}})\ \mathrm{d}h(x) = 0 && \text{(65b)}\\
\Longleftrightarrow &\int_0^\infty J_0(\gamma z)\ \mathrm{d}h(z^2) = 0\\
\Longleftrightarrow &\int_0^\infty \mathrm{d}h(z^2)\int_0^\infty J_0(\gamma z)J_0(\gamma s)\gamma\ d\gamma = 0,\ \forall s\geq 0\\
\Longleftrightarrow &\int_0^\infty \frac{\delta(s-z)}{s}\mathrm{d}h(z^2) = 0,\ \forall s\geq 0 && \text{(65c)}\\
\Longleftrightarrow &\int_0^\infty \mathrm{d}h(z^2)\int_0^\infty e^{-st}\delta(s-z)\ ds = 0,\ \forall t\geq 0\\
\Longleftrightarrow &\int_0^\infty e^{-zt}\mathrm{d}h(z^2) = 0,\ \forall t\geq 0\\
\Longleftrightarrow &\ \mathrm{d}h(z^2) = 0,\ \forall z && \text{(65d)}\\
\Longleftrightarrow &\ \mathrm{d}h(x) = 0,\ \forall x\geq 0,
\end{aligned}$$

where (65a) follows by (63), (65b) by (64), and where (65c) is by orthogonality of Bessel functions given by

$$\int_0^\infty J_\nu(kr)J_\nu(k'r)r\ dr = \frac{\delta(k-k')}{k}.$$

Equation (65d) also follows by invertibility of the Laplace transform.

H. Justification of (40b)

In order to show that

$$\begin{aligned}
&\lim_{\mathsf{l}\to\infty}\int_{x\geq 0,y\geq 0} K(x,y)\log I_0\left(2\frac{\sqrt{y\mathsf{l}}}{\mathsf{S}+1}\right)\ dy\ \ dF_{\widetilde{X}}(x)\\
&=\lim_{\mathsf{l}\to\infty}\int_{x\geq 0,y\geq 0} K(x,y)\log\left(\frac{e^{2\frac{\sqrt{y\mathsf{l}}}{\mathsf{S}+1}}}{\sqrt{4\pi\frac{\sqrt{y\mathsf{l}}}{\mathsf{S}+1}}}\right)\ dy\ dF_{\widetilde{X}}(x),
\end{aligned}$$

we make the variable change $y\mathsf{I}=z$ and prove

$$\lim_{\mathsf{I}\to\infty}\int_{x\geq 0,z\geq 0} K\left(x,\frac{z}{\mathsf{I}}\right)\log\left(I_0\left(2\frac{\sqrt{z}}{\mathsf{S}+1}\right)e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}}\sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}}\right)\ \frac{dz}{\mathsf{I}}\ dF_{\widetilde{X}}(x)$$

$$=\lim_{\mathsf{I}\to\infty}\frac{1}{\mathsf{I}}\int_{z\geq 0} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{I}}\right)\log\left(I_0\left(2\frac{\sqrt{z}}{\mathsf{S}+1}\right)e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}}\sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}}\right)\ dz$$

$$=\lim_{\mathsf{I}\to\infty}\frac{1}{\mathsf{I}}\int_{0}^{\sqrt{\mathsf{I}}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{I}}\right)\log\left(I_0\left(2\frac{\sqrt{z}}{\mathsf{S}+1}\right)e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}}\sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}}\right)\ dz \tag{66}$$

$$+\lim_{\mathsf{I}\to\infty}\frac{1}{\mathsf{I}}\int_{\sqrt{\mathsf{I}}}^{\infty} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{I}}\right)\log\left(I_0\left(2\frac{\sqrt{z}}{\mathsf{S}+1}\right)e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}}\sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}}\right)\ dz=0 \tag{67}$$

We first claim that the limit in (66) is equal to zero since

$$\lim_{\mathsf{I}\to\infty}\frac{1}{\mathsf{I}}\left|\int_{z=0}^{\sqrt{\mathsf{I}}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{I}}\right)\log\left(I_0(2\frac{\sqrt{z}}{\mathsf{S}+1})e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}}\sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}}\right)\ dz\right|$$

$$\leq\lim_{\mathsf{I}\to\infty}\frac{1}{\mathsf{I}}\int_{z=0}^{\sqrt{\mathsf{I}}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{I}}\right)\left(4\frac{\sqrt{z}}{\mathsf{S}+1}+\left|\log\left(\sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}}\right)\right|\right)\ dz \tag{68a}$$

$$\leq\lim_{\mathsf{I}\to\infty}\frac{1}{\mathsf{I}}\int_{z=0}^{\sqrt{\mathsf{I}}}\left(4\frac{\sqrt{z}}{\mathsf{S}+1}+\frac{1}{2}\left|\log(\frac{4\pi}{\mathsf{S}+1})\right|+\frac{1}{4}\left|\log(z)\right|\right)\ dz \tag{68b}$$

$$\leq\frac{1}{\mathsf{I}}\left\{\frac{8}{3(\mathsf{S}+1)}\sqrt{\mathsf{I}}^{3/2}+\frac{1}{2}\left|\log(\frac{4\pi}{\mathsf{S}+1})\right|\sqrt{\mathsf{I}}-\frac{1}{4}\left(z\log z-z\right)\big|_0^1+\frac{1}{4}\left(z\log z-z\right)\big|_1^{\sqrt{\mathsf{I}}}\right\}=0,$$

where (68a) is due to the fact that $1\leq I_0(x)\leq e^x$ and (68b) is by (17).

To prove that the limit in (67) is also equal to zero, we note that for $x\gg 1$ the modified Bessel function of the first kind admits the following asymptotic expansion

$$I_0(x)\sim\frac{e^x}{\sqrt{2\pi x}}\left(1+\frac{1}{8x}+\frac{9}{128x^2}+\dots\right)$$

and hence we can write

$$I_0(x)e^{-x}\sqrt{2\pi x}-\left(1+\frac{1}{8x}\right)=O\left(\frac{1}{x^2}\right).$$

Consequently, we can claim that there exists I' and $c>0$ such that

$$-\frac{c}{z}\leq I_0\left(2\frac{\sqrt{z}}{\mathsf{S}+1}\right)e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}}\sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}}-\left(1+\frac{\mathsf{S}+1}{16\sqrt{z}}\right)\leq\frac{c}{z}, \tag{69}$$

for all $z \geq \sqrt{\mathsf{l}'}$. Moreover, there also exists l'' such that

$$\frac{\mathsf{S}+1}{16\sqrt{z}} - \frac{c}{z} \geq 0, \tag{70}$$

for all $z \geq \sqrt{\mathsf{l}''}$. We define the function

$$g(\mathsf{l}) := \int_{\sqrt{\mathsf{l}}}^{\infty} \frac{1}{\mathsf{l}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{l}}\right) \log\left(I_0\left(2\frac{\sqrt{z}}{\mathsf{S}+1}\right) e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}} \sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}} \right) \ dz.$$

Based on (69) and (70), for the regime $\mathsf{l} \geq \max\{\mathsf{l}', \mathsf{l}''\}$, we have

$$g(\mathsf{l}) \geq \int_{\sqrt{\mathsf{l}}}^{\infty} \frac{1}{\mathsf{l}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{l}}\right) \log\left(1 + \frac{\mathsf{S}+1}{16\sqrt{z}} - \frac{c}{z}\right) \ dz \geq 0, \tag{71}$$

and

$$\begin{aligned}
g(\mathsf{l}) &= \int_{\sqrt{\mathsf{l}}}^{\infty} \frac{1}{\mathsf{l}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{l}}\right) \log\left(I_0\left(2\frac{\sqrt{z}}{\mathsf{S}+1}\right) e^{-2\frac{\sqrt{z}}{\mathsf{S}+1}} \sqrt{4\pi\frac{\sqrt{z}}{\mathsf{S}+1}} \right) \ dz \\
&\int_{\sqrt{\mathsf{l}}}^{\infty} \frac{1}{\mathsf{l}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{l}}\right) \log\left(1 + \frac{\mathsf{S}+1}{16\sqrt{z}} + \frac{c}{z}\right) \ dz \qquad (72) \\
&\leq \int_{\sqrt{\mathsf{l}}}^{\infty} \frac{1}{\mathsf{l}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{l}}\right) \left(\frac{\mathsf{S}+1}{16\sqrt{z}} + \frac{c}{z}\right) \ dz \qquad (73) \\
&\leq \int_{\sqrt{\mathsf{l}}}^{\infty} \frac{1}{\mathsf{l}} f_{\widetilde{Y}}\left(\frac{z}{\mathsf{l}}\right) \left(\frac{\mathsf{S}+1}{16\sqrt{\sqrt{\mathsf{l}}}} + \frac{c}{\sqrt{\mathsf{l}}}\right) \ dz \qquad (74) \\
&= \int_{\frac{1}{\sqrt{\mathsf{l}}}}^{\infty} f_{\widetilde{Y}}(y) \left(\frac{\mathsf{S}+1}{16\sqrt{\sqrt{\mathsf{l}}}} + \frac{c}{\sqrt{\mathsf{l}}}\right) \ dy \qquad (75) \\
&\leq \left(\frac{\mathsf{S}+1}{16\sqrt{\sqrt{\mathsf{l}}}} + \frac{c}{\sqrt{\mathsf{l}}}\right) \int_{0}^{\infty} f_{\widetilde{Y}}(y) \ dy = \frac{\mathsf{S}+1}{16\mathsf{l}^{1/4}} + \frac{c}{\sqrt{\mathsf{l}}}, \qquad (76)
\end{aligned}$$

where in (73) we used the inequality $\log(1+x) < x, x \geq 0$, and in (74) the fact that $\frac{1}{\sqrt{z}}$ and $\frac{1}{z}$ are decreasing functions in z, in (75) we did the change of variable $y\mathsf{l} = z$, and in (76) we used that fact that $f_{\widetilde{Y}}(y)$ integrates to one. By (71) and (76), we conclude that

$$\begin{aligned}
&\liminf_{\mathsf{l}\to\infty} g(\mathsf{l}) \geq 0, \\
&\limsup_{\mathsf{l}\to\infty} g(\mathsf{l}) \leq \limsup_{\mathsf{l}\to\infty} \left(\frac{\mathsf{S}+1}{16\sqrt{\sqrt{\mathsf{l}}}} + \frac{c}{\sqrt{\mathsf{l}}}\right) = 0,
\end{aligned}$$

and hence $\lim_{\mathsf{l}\to\infty} g(\mathsf{l}) = 0$ and the proof is complete.

I. Justification of (40c)*: Lower bound on* $\lim_{\mathsf{I}\to\infty}\sqrt{\mathsf{I}}\left[\mathbb{E}[\sqrt{\widetilde{Y}}]-\sqrt{\mathsf{I}}\right]$

We claim that

$$\lim_{\mathsf{I}\to\infty}\sqrt{\mathsf{I}}\left[\mathbb{E}[\sqrt{\widetilde{Y}}]-\sqrt{\mathsf{I}}\right]\geq\frac{\mathsf{S}+1}{4}. \tag{77}$$

In this regard, we first find a lower bound on $\mathbb{E}[\sqrt{\widetilde{Y}}]$ as follows

$$\begin{aligned}
\mathbb{E}\left[\sqrt{\widetilde{Y}}\right] &= \int_{x\geq 0}\int_{|\theta|<\pi}\frac{1}{2\pi}\int_{y\geq 0}\sqrt{y}e^{-(y+x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta)}I_0\left(2\sqrt{y}\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}\right)\,dy\;d\theta\;dF_{\widetilde{X}}(x)\\
&= \int_{x\geq 0}\int_{|\theta|<\pi}\frac{1}{2\pi}e^{-(x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta)}\Gamma(3/2)\,{}_1F_1\left(\frac{3}{2},1,x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta\right)\;d\theta\;dF_{\widetilde{X}}(x) \qquad \text{(78a)}\\
&= \int_{x\geq 0}\int_{|\theta|<\pi-\delta}\frac{1}{2\pi}e^{-(x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta)}\Gamma(3/2)\,{}_1F_1\left(\frac{3}{2},1,x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta\right)\;d\theta\;dF_{\widetilde{X}}(x)\\
&\quad+\int_{x\geq 0}\int_{\pi-\delta}^{\pi+\delta}\frac{1}{2\pi}e^{-(x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta)}\Gamma(3/2)\,{}_1F_1\left(\frac{3}{2},1,x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta\right)\;d\theta\;dF_{\widetilde{X}}(x)\\
&\geq \int_{x\geq 0}\int_{|\theta|<\pi-\delta}\frac{1}{2\pi}\left[\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}+\frac{1}{4\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}}\right]\;d\theta\;dF_{\widetilde{X}}(x) \qquad \text{(78b)}\\
&\quad+\int_{x=0}^{A}\int_{\pi-\delta}^{\pi+\delta}\frac{1}{2\pi}\left[\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}+\frac{1}{4\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}}\right]\;d\theta\;dF_{\widetilde{X}}(x) \qquad \text{(78c)}\\
&\geq \int_{x\geq 0}\int_{|\theta|<\pi-\delta}\frac{1}{2\pi}\left[\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}+\frac{1}{4\left(\sqrt{x}+\sqrt{\mathsf{I}}\right)}\right]\;d\theta\;dF_{\widetilde{X}}(x)\\
&\quad+\int_{x=0}^{A}\int_{\pi-\delta}^{\pi+\delta}\frac{1}{2\pi}\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}\;d\theta\;dF_{\widetilde{X}}(x)\\
&= \int_{x\geq 0}\int_{|\theta|<\pi}\frac{1}{2\pi}\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}\;d\theta\;dF_{\widetilde{X}}(x)+\frac{1}{4\left(\sqrt{\mathsf{S}}+\sqrt{\mathsf{I}}\right)} \qquad \text{(78d)}\\
&\quad-\int_{x=A}^{\infty}\int_{\pi-\delta}^{\pi+\delta}\frac{1}{2\pi}\sqrt{x+\mathsf{I}+2\sqrt{x\mathsf{I}}\cos\theta}\;d\theta\;dF_{\widetilde{X}}(x), \qquad \text{(78e)}
\end{aligned}$$

where $A=\mathsf{I}^{1-\epsilon_1}$ and $\delta=\mathsf{I}^{-(1-\epsilon_2)}$ for some $\epsilon_1,\epsilon_2>0$. The equality in (78a) is due to [24, eq(6.631)]

$$\int_{x\geq 0}\sqrt{x}e^{-\alpha x}I_0(2\beta\sqrt{x})dx=\frac{\Gamma(\frac{3}{2})}{\alpha^{3/2}}\,{}_1F_1\left(3/2,1,\frac{\beta^2}{\alpha}\right),$$

where ${}_1F_1(a,b,x)$ is the confluent hypergeometric function [25, Chapter 13]. In (78b) and (78c) we have used the series expansion of ${}_1F_1(a,c,x)$ for $x \gg 1$ (which is admissible for $|\theta| < \pi - \delta$ and $x \ll \mathsf{I}$), which is given by [25, Section 13.7]

$$\begin{aligned} {}_1F_1(a,c,x) &= \frac{\Gamma(c)e^x x^{a-c}}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{(c-a)_n(1-a)_n}{n!} x^{-n} \\ &= \frac{\Gamma(c)e^x x^{a-c}}{\Gamma(a)} \left[1 + \frac{(c-a)(1-a)}{x} + O\left(\frac{1}{x^2}\right)\right], \end{aligned}$$

where $(a)_n := a(a+1)\dots(a+n-1)$. The second term in (78d) is derived by the Jensen's inequality.

The term in (78e) times $\sqrt{\mathsf{I}}$ is vanishing as $\mathsf{I} \to \infty$ by noting that

$$\begin{aligned} &\lim_{\mathsf{I}\to\infty} \sqrt{\mathsf{I}} \int_{x=A}^{\infty} \int_{\pi-\delta}^{\pi+\delta} \frac{1}{2\pi} \sqrt{x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos\theta}\ d\theta\ dF_{\widetilde{X}}(x) \\ &\le \lim_{\mathsf{I}\to\infty} \sqrt{\mathsf{I}} \int_{x=A}^{\infty} \frac{\delta}{\pi} \left(\sqrt{x} + \sqrt{\mathsf{I}}\right)\ dF_{\widetilde{X}}(x) \\ &\le \lim_{\mathsf{I}\to\infty} \left(\sqrt{\mathsf{I}}\delta\frac{\sqrt{\mathsf{S}}}{\pi} + \frac{\delta}{\pi}\mathsf{I}\mathbb{P}[X > A]\right) \le \lim_{\mathsf{I}\to\infty} \left(\sqrt{\mathsf{I}}\frac{\delta}{\pi}\sqrt{\mathsf{S}} + \frac{\delta}{\pi}\mathsf{I}\frac{\mathsf{S}}{A}\right) \to 0, \end{aligned}$$

for the chosen values of A and δ. In addition, we handle the first term in (78d) by using the fact that

$$\mathbb{E}_{\theta}\left[\sqrt{r + 1 + 2\sqrt{r}\cos\theta}\right] > \frac{r}{4} + 1,$$

for $r \le 4$ and hence

$$\begin{aligned} &\lim_{\mathsf{I}\to\infty} \sqrt{\mathsf{I}} \left[\int_{x\ge 0} \int_{|\theta|<\pi} \frac{1}{2\pi} \sqrt{x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos\theta}\ d\theta\ dF_{\widetilde{X}}(x) - \sqrt{\mathsf{I}}\right] \\ &= \lim_{\mathsf{I}\to\infty} \sqrt{\mathsf{I}} \left[\int_{x=0}^{4\mathsf{I}} \int_{|\theta|<\pi} \frac{1}{2\pi} \left[\sqrt{x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos\theta} - \sqrt{\mathsf{I}}\right]\ d\theta\ dF_{\widetilde{X}}(x)\right] \\ &+ \lim_{\mathsf{I}\to\infty} \sqrt{\mathsf{I}} \left[\int_{x=4\mathsf{I}}^{\infty} \int_{|\theta|<\pi} \frac{1}{2\pi} \left[\sqrt{x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos\theta} - \sqrt{\mathsf{I}}\right]\ d\theta\ dF_{\widetilde{X}}(x)\right] \\ &\ge \lim_{\mathsf{I}\to\infty} \sqrt{\mathsf{I}} \left[\int_{x=0}^{4\mathsf{I}} \left[\frac{x}{4\sqrt{\mathsf{I}}}\right]\ \ dF_{\widetilde{X}}(x)\right] + \lim_{\mathsf{I}\to\infty} \sqrt{\mathsf{I}} \left[\int_{x=4\mathsf{I}}^{\infty} \frac{1}{2\pi} \left[\sqrt{4\mathsf{I}} - \sqrt{\mathsf{I}} - \sqrt{\mathsf{I}}\right]\ dF_{\widetilde{X}}(x)\right] \\ &\ge \frac{\mathsf{S}}{4}. \end{aligned} \tag{79}$$

By (78d) and (79), the claim (77) is proved.

J. Justification of (40c)*: Calculation of* $\lim_{\mathsf{l}\to\infty}\left[\mathbb{E}_{\theta,\widetilde{X}}\left[\log(\widetilde{X}+\mathsf{l}+2\sqrt{\widetilde{X}\mathsf{l}}\cos\theta)\right]-\sqrt{\mathsf{l}}\right]$

We first find the expected value of $\log(\widetilde{X}+\mathsf{l}+2\sqrt{\widetilde{X}\mathsf{l}}\cos\theta)$ with respect to $\widetilde{X}$ and θ given by

$$\begin{aligned}\mathbb{E}_{\theta,\widetilde{X}}\left[\log(\widetilde{X}+\mathsf{l}+2\sqrt{\widetilde{X}\mathsf{l}}\cos\theta)\right] &= \int_{x\geq 0}\int_0^{2\pi}\log(x+\mathsf{l}+2\sqrt{x\mathsf{l}}\cos\theta)d\theta\ dF_{\widetilde{X}}(x)\\ &= C_1+C_2,\end{aligned}$$

where

$$\begin{aligned}C_1 &= \int_{x=0}^{\mathsf{l}}\int_0^{2\pi}\log(x+\mathsf{l}+2\sqrt{x\mathsf{l}}\cos\theta)\ d\theta\ dF_{\widetilde{X}}(x),\\ C_2 &= \int_{x\geq\mathsf{l}}\int_0^{2\pi}\log(x+\mathsf{l}+2\sqrt{x\mathsf{l}}\cos\theta)\ d\theta\ dF_{\widetilde{X}}(x).\end{aligned}$$

To calculate C_1, we state the following lemma.

Lemma 2.

$$\int_0^{2\pi}\log(1+2r\cos(x)+r^2)\ dx=0,\ \ 0\leq r\leq 1 \tag{80}$$

Proof. Based on Cauchy's integral formula

$$f(a)=\frac{1}{2\pi j}\oint_\gamma\frac{f(z)}{z-a}dz,$$

for $0\leq r<1$, we can write

$$\begin{aligned}\int_0^{2\pi}\log(1+2r\cos(x)+r^2)\ dx &= \int_0^{2\pi}\log(1+re^{jx})\ dx+\int_0^{2\pi}\log(1+re^{-jx})\ dx\\ &= 2\oint_\gamma\frac{\log(1+z)}{jz}\ dz=0.\end{aligned}$$

For $r=1$, we have

$$\begin{aligned}\int_0^{2\pi}\log(1+2r\cos(x)+r^2)\ dx &= 4\pi\log(2)+4\int_0^{\pi}\log\left(\cos(\theta)\right)\ d\theta \qquad (81a)\\ &= 4\pi\log(2)+4\int_0^{\frac{\pi}{2}}\log\left(\cos(\theta)\right)\ d\theta+4\int_0^{\frac{\pi}{2}}\log\left(\sin(\theta)\right)\ d\theta\\ &= 4\pi\log(2)+4\int_0^{\frac{\pi}{2}}\log\left(\frac{\sin(2\theta)}{2}\right)\ d\theta\\ &= 2\pi\log(2)+2\int_0^{\pi}\log\left(\sin(\theta)\right)\ d\theta\\ &= 2\pi\log(2)+2\int_0^{\pi}\log\left(\cos(\theta)\right)\ d\theta, \qquad (81b)\end{aligned}$$

which according to (81a) and (81b), results in

$$4\pi \log(2) + 4\int_0^{\pi} \log\left(\cos(\theta)\right)\ d\theta = 2\pi \log(2) + 2\int_0^{\pi} \log\left(\cos(\theta)\right)\ d\theta = 0.$$

$\square$

Based on lemma 2, we see that

$$\begin{aligned} C_1 &= \int_{x=0}^{\mathsf{I}} \int_0^{2\pi} \log(x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos\theta)\ d\theta\ dF_{\widetilde{X}}(x), \\ &= \int_{x=0}^{\mathsf{I}} \log(\mathsf{I})\ dF_{\widetilde{X}}(x) + \int_{x=0}^{\mathsf{I}} \int_0^{2\pi} \log(1 + 2\sqrt{\frac{x}{\mathsf{I}}}\cos\theta + \frac{x}{\mathsf{I}})\ d\theta\ dF_{\widetilde{X}}(x) \\ &= \int_{x=0}^{\mathsf{I}} \log(\mathsf{I})\ dF_{\widetilde{X}}(x). \end{aligned}$$

In addition,

$$\begin{aligned} C_2 &= \int_{x\geq \mathsf{I}} \int_0^{2\pi} \log(x + \mathsf{I} + 2\sqrt{x\mathsf{I}}\cos\theta)\ d\theta\ dF_{\widetilde{X}}(x) \\ &= \int_{x\geq \mathsf{I}} \log(x)\ dF_{\widetilde{X}}(x). \end{aligned}$$

As the result

$$\begin{aligned} &\lim_{\mathsf{I}\to\infty} \left[\mathbb{E}_{\theta,\widetilde{X}}\left[\log(\widetilde{X} + \mathsf{I} + 2\sqrt{\widetilde{X}\mathsf{I}}\cos\theta)\right] - \log(\mathsf{I})\right] \\ &= \lim_{\mathsf{I}\to\infty} \left[\int_{x=0}^{\mathsf{I}} \log(\mathsf{I})\ dF_{\widetilde{X}}(x) + \int_{x\geq \mathsf{I}} \log(x)\ dF_{\widetilde{X}}(x) - \int_{x=0}^{\infty} \log(\mathsf{I})\ dF_{\widetilde{X}}(x)\right] \\ &= \lim_{\mathsf{I}\to\infty} \int_{x\geq \mathsf{I}} \log(\frac{x}{\mathsf{I}})\ dF_{\widetilde{X}}(x) = 0, \end{aligned} \tag{82}$$

where the last step is by Dominated Convergence Theorem.

REFERENCES

[1] S. Shahi, D. Tuninetti, and N. Devroye, "On the capacity of the AWGN channel with additive radar interference," in *2016 54nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2016.

[2] A. Salim, D. Tuninetti, N. Devroye, and D. Erricolo, "Modeling the interference of pulsed radar signals at OFDM-based communications systems," *submitted for publication*, 2016.

[3] S. Ihara, "On the capacity of channels with additive non-Gaussian noise," *Information and Control*, vol. 37, no. 1, pp. 34 – 39, 1978.

[4] 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.

[5] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Information and Control*, vol. 18, no. 3, pp. 203 – 219, 1971.

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

[8] J. Huang and S. P. Meyn, "Characterization and computation of optimal distributions for channel coding," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2336–2351, July 2005.

[9] 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.

[10] H. V. Vu, N. H. Tran, M. C. Gursoy, T. Le-Ngoc, and S. I. Hariharan, "Capacity-achieving input distributions of additive quadrature gaussian mixture noise channels," *IEEE Transactions on Communications*, vol. 63, no. 10, pp. 3607–3620, Oct 2015.

[11] D. A. Le, H. V. Vu, N. H. Tran, M. C. Gursoy, and T. Le-Ngoc, "Approximation of achievable rates in additive gaussian mixture noise channels," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 5011–5024, Dec 2016.

[12] J. Sommerfeld, I. Bjelakovic, and H. Boche, "On the boundedness of the support of optimal input measures for Rayleigh fading channels," in *2008 IEEE International Symposium on Information Theory*, July 2008, pp. 1208–1212.

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

[14] ——, "On the error rate of a communication system suffering from additive radar interference," in *Proc. of IEEE Global Communicaitons Conference (GlobeCom)*, Washington, DC, Dec 2016.

[15] 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.

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

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

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

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

[20] M. C. Gursoy, H. V. Poor, and S. Verdu, "The capacity of the noncoherent Rician fading channel," http://ecs.syr.edu/faculty/gursoy/files/Techreport_revised.pdf, 2002, princeton University Technical Report.

[21] 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.

[22] 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.

[23] 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.

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

[25] F. W. Olver, *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.