entropy

MDPI

Article

Amplitude Constrained Vector Gaussian Wiretap Channel: Properties of the Secrecy-Capacity-Achieving Input Distribution

Antonino Favano [1], Luca Barletta [1,*] and Alex Dytso [2]

[1] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milano, Italy
[2] Qualcomm, Bridgewater, NJ 08807, USA
* Correspondence: luca.barletta@polimi.it

Abstract: This paper studies the secrecy capacity of an n-dimensional Gaussian wiretap channel under a peak power constraint. This work determines the largest peak power constraint $\bar{\mathsf{R}}_n$, such that an input distribution uniformly distributed on a single sphere is optimal; this regime is termed the low-amplitude regime. The asymptotic value of $\bar{\mathsf{R}}_n$ as n goes to infinity is completely characterized as a function of noise variance at both receivers. Moreover, the secrecy capacity is also characterized in a form amenable to computation. Several numerical examples are provided, such as the example of the secrecy-capacity-achieving distribution beyond the low-amplitude regime. Furthermore, for the scalar case $(n = 1)$, we show that the secrecy-capacity-achieving input distribution is discrete with finitely many points at most at the order of $\frac{\mathsf{R}^2}{\sigma_1^2}$, where σ_1^2 is the variance of the Gaussian noise over the legitimate channel.

Keywords: wiretap channel; MIMO; amplitude constraints

Citation: Favano, A.; Barletta, L.; Dytso, A. Amplitude Constrained Vector Gaussian Wiretap Channel: Properties of the Secrecy-Capacity-Achieving Input Distribution. *Entropy* **2023**, *25*, 741. https://doi.org/10.3390/e25050741

Academic Editors: Syed A. Jafar and Eduard Jorswieck

Received: 13 February 2023
Revised: 23 April 2023
Accepted: 25 April 2023
Published: 30 April 2023

1. Introduction

Consider the vector Gaussian wiretap channel with outputs

$$\mathbf{Y}_1 = \mathbf{X} + \mathbf{N}_1, \tag{1a}$$
$$\mathbf{Y}_2 = \mathbf{X} + \mathbf{N}_2, \tag{1b}$$

where $\mathbf{X} \in \mathbb{R}^n$, $\mathbf{N}_1 \sim \mathcal{N}(\mathbf{0}_n, \sigma_1^2\mathbf{I}_n)$ and $\mathbf{N}_2 \sim \mathcal{N}(\mathbf{0}_n, \sigma_2^2\mathbf{I}_n)$, and with $(\mathbf{X}, \mathbf{N}_1, \mathbf{N}_2)$ being mutually independent. The output $\mathbf{Y}_1$ is observed by the legitimate receiver, whereas the output $\mathbf{Y}_2$ is observed by the malicious receiver. In this work, we are interested in the scenario where the input $\mathbf{X}$ is limited by a peak power constraint or amplitude constraint, and assume that $\mathbf{X} \in \mathcal{B}_0(\mathsf{R}) = \{\mathbf{x} : \|\mathbf{x}\| \leq \mathsf{R}\}$, i.e., $\mathcal{B}_0(\mathsf{R})$ is an n-ball centered at the origin and of radius R. For this setting, the secrecy capacity is given by

$$C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n) = \max_{\mathbf{X} \in \mathcal{B}_0(\mathsf{R})} I(\mathbf{X}; \mathbf{Y}_1) - I(\mathbf{X}; \mathbf{Y}_2) \tag{2}$$
$$= \max_{\mathbf{X} \in \mathcal{B}_0(\mathsf{R})} I(\mathbf{X}; \mathbf{Y}_1 | \mathbf{Y}_2), \tag{3}$$

where the last expression holds due to the (stochastically) degraded nature of the channel. It can be shown that for $\sigma_1^2 \geq \sigma_2^2$ the secrecy capacity is equal to zero. Therefore, in the remainder, we assume that $\sigma_1^2 < \sigma_2^2$.

We are interested in studying the input distribution $P_{\mathbf{X}^\star}$ that maximizes (3) in the low (but not vanishing) amplitude regime. Since closed-form expressions for secrecy capacity are rare, we derive the secrecy capacity in an integral form that is easy to evaluate. For the scalar case $(n = 1)$, we establish an upper bound on the number of mass points of $P_{X^\star}$, valid for any amplitude regime. We also argue in Section 2.3 that the solution to the secrecy

capacity can shed light on other problems seemingly unrelated to security. The paper also provides a number of numerical simulations of $P_{\mathbf{X}^\star}$ and C_s, the data for which are made available at [1].

1.1. Literature Review

The wiretap channel was introduced by Wyner in [2], who also established the secrecy capacity of the degraded wiretap channel. The results of [2] were extended to the Gaussian wiretap channel in [3]. The wiretap channel plays a central role in network information theory; the interested reader is referred to [4–8] and references therein for a detailed treatment of the topic. Furthermore, for an in-depth discussion on the wiretap fading channel, refer to [9–12].

In [3], it was shown that the secrecy-capacity-achieving input distribution of the Gaussian wiretap channel, under an average power constraint, is Gaussian. In [13], the authors investigated the Gaussian wiretap channel consisting of two antennas, both at the transmitter and receiver sides, and of a single antenna for the eavesdropper. The secrecy capacity of the MIMO wiretap channel was characterized in [14,15], where the Gaussian input was shown to be optimal. An elegant proof, using the I-MMSE relationship [16], of the optimality of Gaussian input, is given in [17]. Moreover, an alternative approach in the characterization of the secrecy capacity of a MIMO wiretap channel was proposed in [18]. In [19,20], the authors discuss the optimal signaling for secrecy rate maximization under average power constraints.

The secrecy capacity of the Gaussian wiretap channel under the peak power constraint has received far less attention. The secrecy capacity of the scalar Gaussian wiretap channel with an amplitude and power constraint was considered in [21], where the authors showed that the capacity-achieving input distribution $P_{X^\star}$ is discrete with finitely many support points.

The work of [21] was extended to noise-dependent channels by Soltani and Rezki in [22]. For further studies on the properties of the secrecy-capacity-achieving input distribution for a class of degraded wiretap channels, refer to [23–25].

The secrecy capacity for the vector wiretap channel with a peak power constraint was considered in [25], where it was shown that the optimal input distribution is concentrated on finitely many co-centric shells.

1.2. Contributions and Paper Outline

In Section 2, we introduce the mathematical tools, assumptions, and definitions used throughout the paper. Specifically, in Section 2.1, we introduce the oscillation theorem. In Section 2.2, we give a definition of low-amplitude regimes. Moreover, in Section 2.3, we show how the wiretap channel can be seen as a generalization of point-to-point channels and the evaluation of the largest minimum mean square error (MMSE), both under the assumption of amplitude-constrained input. In Section 2.4, we provide a definition of the Karush–Kuhn–Tucker (KKT) conditions for the wiretap channel.

In Section 3, we detail our main results. Theorem 2 provides a sufficient condition for the optimality of a single hypersphere. Theorem 3 and Theorem 4 give the conditions under which we can fully characterize the behavior of $\bar{\mathsf{R}}_n$, that is, the radius below which we are in the low-amplitude regime, i.e., the optimal input distribution is composed of a single shell. Furthermore, Theorem 5 gives an implicit and an explicit upper bound on the number of mass points of the secrecy-capacity-achieving input distribution when $n = 1$.

In Section 4, we derive the secrecy capacity expression for the low-amplitude regime in Theorem 6. We also investigate its behavior when the number of antennas n goes to infinity.

Section 5 extends the investigation of the secrecy capacity beyond the low-amplitude regime. We numerically estimate both the optimal input pmf and the resulting capacity via an algorithmic procedure based on the KKT conditions introduced in Lemma 2.

Sections 6–9 provide the proof for Theorem 3 and Theorem 4–6, respectively. Finally, Section 10 concludes the paper.

1.3. Notation

We use bold letters for vectors ($\mathbf{x}$) and uppercase letters for random variables (X). We denote by $\|\mathbf{x}\|$ the Euclidean norm of the vector $\mathbf{x}$. Given a vector $\mathbf{x} \in \mathbb{R}^n$ and a scalar a, with a little abuse of notation, we denote $\|a \cdot \mathbf{e}_1 + \mathbf{x}\|$ by $\|a + \mathbf{x}\|$, where $\mathbf{e}_1 = [1, 0, \ldots, 0]$ is the first vector in the standard basis of the Euclidean vector space $\mathbb{R}^n$. Given a random variable X, its probability density function (pdf), pmf, and cumulative distribution function are denoted by f_X, P_X, and F_X, respectively. The support set of $P_{\mathbf{X}}$ is denoted and defined as

$$\mathsf{supp}(P_{\mathbf{X}}) = \{\mathbf{x} : \text{ for every open set } \mathcal{D} \ni \mathbf{x} \text{ we have that } P_{\mathbf{X}}(\mathcal{D}) > 0\}. \tag{4}$$

We denote by $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ a multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. The pdf of a Gaussian random variable with zero mean and variance σ^2 is denoted by $\phi_\sigma(\cdot)$. We denote by $\chi^2_n(\lambda)$ the noncentral chi-square distribution with n degrees of freedom and with noncentrality parameter λ. We represent the $n \times 1$ vector of zeros by $\mathbf{0}_n$ and the $n \times n$ identity matrix by $\mathbf{I}_n$. Furthermore, we represent by D the relative entropy. The minimum mean squared error is denoted by

$$\mathrm{mmse}(\mathbf{X}|\mathbf{X} + \mathbf{N}) = \mathbb{E}\Big[\|\mathbf{X} - \mathbb{E}[\mathbf{X}|\mathbf{X} + \mathbf{N}]\|^2\Big]. \tag{5}$$

The modified Bessel function of the first kind of order $v \geq 0$ is denoted by $\mathsf{I}_v(x), x \in \mathbb{R}$. The following ratio of the Bessel functions is commonly used in this work:

$$\mathsf{h}_v(x) = \frac{\mathsf{I}_v(x)}{\mathsf{I}_{v-1}(x)}, \; x \in \mathbb{R}, \; v \geq 0. \tag{6}$$

Finally, the number of zeros (counted in accordance with their multiplicities) of a function $f\colon \mathbb{R} \to \mathbb{R}$ on the interval $\mathcal{I}$ is denoted by $\mathrm{N}(\mathcal{I}, f)$. Similarly, if $f\colon \mathbb{C} \to \mathbb{C}$ is a function on the complex domain, $\mathrm{N}(\mathcal{D}, f)$ denotes the number of its zeros within the region $\mathcal{D}$.

2. Preliminaries

2.1. Oscillation Theorem

In this work, we often need to upper bound the number of oscillations of a function, i.e., its number of sign changes. This is useful, for example, to bound the number of zeros of a function or the number of roots of an equation. To be more precise, let us define the number of sign changes as follows.

Definition 1 (Sign Changes of a Function)**.** *The number of sign changes of a function $\xi : \Omega \to \mathbb{R}$ is given by*

$$\mathscr{S}(\xi) = \sup_{m \in \mathbb{N}} \left\{ \sup_{y_1 < \cdots < y_m \subseteq \Omega} \mathscr{N}\{\xi(y_i)\}_{i=1}^m \right\}, \tag{7}$$

where $\mathscr{N}\{\xi(y_i)\}_{i=1}^m$ is the number of sign changes of the sequence $\{\xi(y_i)\}_{i=1}^m$.

Definition 2 (Totally Positive Kernel)**.** *A function $f : \mathbb{I}_1 \times \mathbb{I}_2 \to \mathbb{R}$ is said to be a totally positive kernel of order n if $\det\Big([f(x_i, y_j)]_{i,j=1}^m\Big) > 0$ for all $1 \leq m \leq n$, for all $x_1 < \cdots < x_m \in \mathbb{I}_1$, and $y_1 < \cdots < y_m \in \mathbb{I}_2$. If f is a totally positive kernel of order n for all $n \in \mathbb{N}$, then f is a strictly totally positive kernel.*

In [26], Karlin noticed that some integral transformations have a *variation-diminishing* property, which is described in the following theorem.

Theorem 1 (Oscillation Theorem). *Given domains $\mathbb{I}_1$ and $\mathbb{I}_2$, let $p\colon \mathbb{I}_1 \times \mathbb{I}_2 \to \mathbb{R}$ be a strictly totally positive kernel. For an arbitrary y, suppose $p(\cdot, y)\colon \mathbb{I}_1 \to \mathbb{R}$ is an n-times differentiable function. Assume that μ is a measure on $\mathbb{I}_2$, and let $\xi\colon \mathbb{I}_2 \to \mathbb{R}$ be a function with $\mathscr{S}(\xi) = n$. For $x \in \mathbb{I}_1$, define*

$$\Xi(x) = \int \xi(y) p(x, y) \mathrm{d}\mu(y). \tag{8}$$

If $\Xi\colon \mathbb{I}_1 \to \mathbb{R}$ is an n-times differentiable function, then either $\mathrm{N}(\mathbb{I}_1, \Xi) \leq n$, or $\Xi \equiv 0$.

The above theorem says that the number of zeros of a function Ξ, which is the output of the integral transformation, is less than the number of sign changes of the function ξ, which is the input to the integral transformation.

2.2. Low-Amplitude Regime

In this work, a low-amplitude regime is defined as follows.

Definition 3. *Let $\mathbf{X}_{\mathsf{R}} \sim P_{\mathbf{X}_{\mathsf{R}}}$ be uniform on $\mathcal{C}(\mathsf{R}) = \{\mathbf{x} : \|\mathbf{x}\| = \mathsf{R}\}$. The capacity in (3) is said to be in the low-amplitude regime if $\mathsf{R} \leq \bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$, where*

$$\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2) = \max\left\{\mathsf{R} : P_{\mathbf{X}_{\mathsf{R}}} = \arg \max_{P_{\mathbf{X}}\colon \mathbf{X} \in \mathcal{B}_0(\mathsf{R})} I(\mathbf{X}; \mathbf{Y}_1 | \mathbf{Y}_2)\right\}. \tag{9}$$

If the set in (9) is empty, then we assign $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2) = 0$.

The quantity $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$ represents the largest radius R, for which $P_{\mathbf{X}_{\mathsf{R}}}$ is secrecy-capacity-achieving.

One of the main objectives of this work is to characterize $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$.

2.3. Connections to Other Optimization Problems

The distribution $P_{\mathbf{X}_{\mathsf{R}}}$ occurs in a variety of statistical and information-theoretic applications. For example, consider the following two optimization problems:

$$\max_{P_{\mathbf{X}}\colon \mathbf{X} \in \mathcal{B}_0(\mathsf{R})} I(\mathbf{X}; \mathbf{X} + \mathbf{N}), \tag{10}$$

$$\max_{P_{\mathbf{X}}\colon \mathbf{X} \in \mathcal{B}_0(\mathsf{R})} \mathrm{mmse}(\mathbf{X} | \mathbf{X} + \mathbf{N}), \tag{11}$$

where $\mathbf{N} \sim \mathcal{N}(\mathbf{0}_n, \sigma^2 \mathbf{I}_n)$. The first problem seeks to characterize the capacity of the point-to-point channel under an amplitude constraint, and the second problem seeks to find the largest minimum mean squared error under the assumption that the signal has bounded amplitude; the interested reader is referred to [27–29] for a detailed background on both problems.

Similarly to the wiretap channel, we can define the low-amplitude regime for both problems as the largest R such that $P_{\mathbf{X}_{\mathsf{R}}}$ is optimal and denote these by $\bar{\mathsf{R}}_n^{\mathrm{ptp}}(\sigma^2)$ and $\bar{\mathsf{R}}_n^{\mathrm{MMSE}}(\sigma^2)$. We now argue that both $\bar{\mathsf{R}}_n^{\mathrm{ptp}}(\sigma^2)$ and $\bar{\mathsf{R}}_n^{\mathrm{MMSE}}(\sigma^2)$ can be seen as a special case of the wiretap solution. Hence, the wiretap channel provides an interesting unification and generalization of these two problems.

First, note that the point-to-point solution can be recovered from the wiretap by simply specializing the wiretap channel to the point-to-point channel, that is,

$$\bar{\mathsf{R}}_n^{\mathrm{ptp}}(\sigma^2) = \lim_{\sigma_2 \to \infty} \bar{\mathsf{R}}_n(\sigma^2, \sigma_2^2). \tag{12}$$

Second, to see that the MMSE solution can be recovered from the wiretap, recall that by the I-MMSE relationship [16] we have that

$$\max_{P_{\mathbf{X}}\colon \mathbf{X} \in \mathcal{B}_0(\mathsf{R})} I(\mathbf{X}; \mathbf{Y}_1) - I(\mathbf{X}; \mathbf{Y}_2)$$

$$= \max_{P_{\mathbf{X}}:\, \mathbf{X}\in\mathcal{B}_0(\mathsf{R})} \frac{1}{2}\int_{\sigma_1^2}^{\infty} \frac{\mathrm{mmse}(\mathbf{X}|\mathbf{X}+\sqrt{s}\mathbf{Z})}{s^2}\mathrm{d}s - \frac{1}{2}\int_{\sigma_2^2}^{\infty} \frac{\mathrm{mmse}(\mathbf{X}|\mathbf{X}+\sqrt{s}\mathbf{Z})}{s^2}\mathrm{d}s \tag{13}$$

$$= \max_{P_{\mathbf{X}}:\, \mathbf{X}\in\mathcal{B}_0(\mathsf{R})} \frac{1}{2}\int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathrm{mmse}(\mathbf{X}|\mathbf{X}+\sqrt{s}\mathbf{Z})}{s^2}\mathrm{d}s \tag{14}$$

where $\mathbf{Z}$ is standard Gaussian. Now, note that if we choose $\sigma_2^2 = \sigma_1^2 + \epsilon$, then by the mean value theorem we arrive at

$$\max_{P_{\mathbf{X}}:\, \mathbf{X}\in\mathcal{B}_0(\mathsf{R})} I(\mathbf{X};\mathbf{Y}_1) - I(\mathbf{X};\mathbf{Y}_2) = \max_{P_{\mathbf{X}}:\, \mathbf{X}\in\mathcal{B}_0(\mathsf{R})} \frac{\epsilon}{2}\frac{\mathrm{mmse}(\mathbf{X}|\mathbf{X}+\sqrt{\sigma_1^2}\mathbf{Z})}{\sigma_1^4} + o(\epsilon), \tag{15}$$

where $\lim_{\epsilon\to 0^+} o(\epsilon)/\epsilon = 0$. Consequently, for a small enough $\epsilon > 0$,

$$\bar{\mathsf{R}}_n^{\mathrm{MMSE}}(\sigma^2) = \bar{\mathsf{R}}_n(\sigma^2, \sigma^2+\epsilon). \tag{16}$$

2.4. KKT Conditions

Let us define the secrecy density for the vector Gaussian wiretap channel as

$$\Xi(\mathbf{x}; P_{\mathbf{X}^\star}) = \mathsf{D}(f_{\mathbf{Y}_1|\mathbf{X}}(\cdot|\mathbf{x})\|f_{\mathbf{Y}_1^\star}) - \mathsf{D}(f_{\mathbf{Y}_2|\mathbf{X}}(\cdot|\mathbf{x})\|f_{\mathbf{Y}_2^\star}), \tag{17}$$

where $\mathsf{D}(\cdot\|\cdot)$ is the relative entropy.

For the scalar case ($n = 1$), the KKT conditions are necessary and sufficient to ensure that $P_{X^\star}$ is capacity-achieving [21].

Lemma 1. *$P_{X^\star}$ maximizes (3) if, and only if,*

$$\Xi(x) = C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, 1), \qquad x \in \mathsf{supp}(P_{X^\star}), \tag{18}$$

$$\Xi(x) \le C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, 1), \qquad x \in [-\mathsf{R}, \mathsf{R}], \tag{19}$$

where for $x \in \mathbb{R}$

$$\Xi(x) = \mathsf{D}(f_{Y_1|X}(\cdot|x)\|f_{Y_1^\star}) - \mathsf{D}(f_{Y_2|X}(\cdot|x)\|f_{Y_2^\star}) \tag{20}$$

$$= \mathbb{E}[g(Y_1)|X = x] + \log\left(\frac{\sigma_2}{\sigma_1}\right), \tag{21}$$

and where

$$g(y) = \mathbb{E}\left[\log\frac{f_{Y_2^\star}(y+N)}{f_{Y_1^\star}(y)}\right], \qquad y \in \mathbb{R}, \tag{22}$$

with $N \sim \mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$.

Proof. The first part of Lemma 1 was shown in [21]. The proof of (21) goes as follows:

$$\mathsf{D}(f_{Y_1|X}(\cdot|x)\|f_{Y_1^\star}) - \mathsf{D}(f_{Y_2|X}(\cdot|x)\|f_{Y_2^\star}) - \log\left(\frac{\sigma_2}{\sigma_1}\right) \tag{23}$$

$$= \int_{-\infty}^{\infty} \log\frac{1}{f_{Y_1^\star}(y)}\phi_{\sigma_1}(y-x)\mathrm{d}y - \int_{-\infty}^{\infty} \log\frac{1}{f_{Y_2^\star}(y)}\mathbb{E}[\phi_{\sigma_1}(y-x-N)]\mathrm{d}y \tag{24}$$

$$= \int_{-\infty}^{\infty} \log\frac{1}{f_{Y_1^\star}(y)}\phi_{\sigma_1}(y-x)\mathrm{d}y - \int_{-\infty}^{\infty} \mathbb{E}\left[\log\frac{1}{f_{Y_2^\star}(y+N)}\right]\phi_{\sigma_1}(y-x)\mathrm{d}y \tag{25}$$

$$= \int_{-\infty}^{\infty} \mathbb{E}\left[\log\frac{f_{Y_2^\star}(y+N)}{f_{Y_1^\star}(y)}\right]\phi_{\sigma_1}(y-x)\mathrm{d}y \tag{26}$$

$$= \int_{-\infty}^{\infty} g(y)\phi_{\sigma_1}(y-x)\mathrm{d}y, \tag{27}$$

where $N \sim \mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$ and (24) hold by noticing that $\phi_{\sigma_2}(y-x)$ can be reformulated as the convolution of Gaussian pdfs $\mathbb{E}[\phi_{\sigma_1}(y-x-N)]$; in (25) we applied the change in variable $y \mapsto y+N$. This concludes the proof. □

The convexity of the optimization problem is also guaranteed for the vector wiretap model in (1) with $n > 1$. Then, the results of Lemma 1 can be extended to the vector case as follows.

Lemma 2. *$P_{\mathbf{X}^\star}$ maximizes (3) if, and only if,*

$$\Xi(\mathbf{x}; P_{\mathbf{X}^\star}) = C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n), \qquad \mathbf{x} \in \mathsf{supp}(P_{\mathbf{X}^\star}), \tag{28a}$$

$$\Xi(\mathbf{x}; P_{\mathbf{X}^\star}) \leq C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n), \qquad \mathbf{x} \in \mathcal{B}_0(\mathsf{R}), \tag{28b}$$

where for $\mathbf{x} \in \mathbb{R}^n$

$$\Xi(\mathbf{x}; P_{\mathbf{X}^\star}) = \mathsf{D}(f_{\mathbf{Y}_1|\mathbf{X}}(\cdot|\mathbf{x}) \| f_{\mathbf{Y}_1^\star}) - \mathsf{D}(f_{\mathbf{Y}_2|\mathbf{X}}(\cdot|\mathbf{x}) \| f_{\mathbf{Y}_2^\star}) \tag{29}$$

$$= \mathbb{E}[g(\mathbf{Y}_1)|\mathbf{X} = \mathbf{x}], \tag{30}$$

and where

$$g(\mathbf{y}) = \mathbb{E}\left[\log \frac{f_{\mathbf{Y}_2^\star}(\mathbf{y}+\mathbf{N})}{f_{\mathbf{Y}_1^\star}(\mathbf{y})}\right] + n \log\left(\frac{\sigma_2}{\sigma_1}\right), \ \mathbf{y} \in \mathbb{R}^n, \tag{31}$$

with $\mathbf{N} \sim \mathcal{N}(\mathbf{0}_n, (\sigma_2^2 - \sigma_1^2)\mathbf{I}_n)$.

Proof. This is a straightforward vector extension of Lemma 1. □

Thanks to the spherical symmetry of the additive noise distributions and of $P_{\mathbf{X}}$, the secrecy density $\Xi(\mathbf{x}; P_{\mathbf{X}})$ can be expressed as a function of $\|\mathbf{x}\|$ only. Therefore, we denote the secrecy density in spherical coordinates by $\widetilde{\Xi}(\|\mathbf{x}\|; P_{\|\mathbf{X}\|})$, and give a rigorous definition in (A9).

3. Main Results

3.1. A New Sufficient Condition on the Optimality of $P_{\mathbf{X}_\mathsf{R}}$

Our first main result provides a sufficient condition for the optimality of $P_{\mathbf{X}_\mathsf{R}}$.

Theorem 2. *If*

$$\mathsf{R} < \sigma_1^2 \sqrt{n\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)}, \tag{32}$$

then $P_{\mathbf{X}_\mathsf{R}}$ is secrecy-capacity-achieving.

Proof. Let us consider the equivalent definition of the secrecy density in spherical coordinates (A9). Note that if the derivative of $\widetilde{\Xi}(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|})$ makes at most one sign change, from negative to positive, then the maximum of $\|\mathbf{x}\| \mapsto \widetilde{\Xi}(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|})$ occurs at either $\|\mathbf{x}\| = 0$ or $\|\mathbf{x}\| = \mathsf{R}$.

From Lemma A1 in the Appendix B, the derivative of $\widetilde{\Xi}$ is as given below

$$\widetilde{\Xi}'(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|}) = \|\mathbf{x}\| \, \mathbb{E}\Big[\widetilde{M}_2(\sigma_1 Q_{n+2}) - M_1(\sigma_1 Q_{n+2})\Big] \tag{33}$$

where Q^2_{n+2} is a noncentral chi-square random variable with $n+2$ degrees of freedom and noncentrality parameter $\frac{\|\mathbf{x}\|^2}{\sigma_1^2}$, and

$$M_i(y) = \frac{1}{\sigma_i^2}\left(\frac{\mathsf{R}}{y}\mathsf{h}_{\frac{n}{2}}\left(\frac{\mathsf{R}}{\sigma_i^2}y\right) - 1\right), \qquad i \in \{1,2\} \tag{34}$$

$$\widetilde{M}_2(y) = \mathbb{E}[M_2(\|y + \mathbf{W}\|)], \tag{35}$$

where $\mathbf{W} \sim \mathcal{N}(\mathbf{0}_{n+2}, (\sigma_2^2 - \sigma_1^2)\mathbf{I}_{n+2})$. A calculation related to (33) was erroneously performed in [27]. However, this error does not change the results of [27] as only the sign of the derivative is important and not the value itself. Note that $\widetilde{\Xi}'(0; P_{\|\mathbf{X}_\mathsf{R}\|}) = 0$ and that $\widetilde{\Xi}'(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|}) > 0$ for a sufficiently large $\|\mathbf{x}\|$; in fact, we have

$$\widetilde{\Xi}'(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|}) > \|\mathbf{x}\|\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) - \frac{\|\mathbf{x}\|}{\sigma_1^2}\mathbb{E}\left[\frac{\mathsf{R}}{\sigma_1 Q_{n+2}}\right] \tag{36}$$

$$= \|\mathbf{x}\|\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) - \frac{\|\mathbf{x}\|}{\sigma_1^2}\mathbb{E}\left[\frac{\mathsf{R}}{\|\mathbf{x}\|}\mathsf{h}_{\frac{n}{2}}\left(\frac{\|\mathbf{x}\|}{\sigma_1}Q_n\right)\right] \tag{37}$$

$$\geq \|\mathbf{x}\|\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) - \frac{\mathsf{R}}{\sigma_1^2}, \tag{38}$$

where (36) follows from $0 \leq \mathsf{h}_{\frac{n}{2}}(x) \leq 1$ for $x \geq 0$; (37) follows by noticing that $\frac{\mathsf{R}}{\sigma_1\sqrt{t}} f_{Q^2_{n+2}}(t) = \frac{\mathsf{R}}{\|\mathbf{x}\|}\mathsf{h}_{\frac{n}{2}}\left(\frac{\|\mathbf{x}\|}{\sigma_1}\sqrt{t}\right) f_{Q^2_n}(t)$; and finally, (38) holds by $\mathsf{h}_{\frac{n}{2}}(x) \leq 1$.

Then, to show that $\widetilde{\Xi}(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|})$ is maximized in $\|\mathbf{x}\| = \mathsf{R}$, we need to prove that $\widetilde{\Xi}'(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|})$ changes sign at most once. To that end, we need Karlin's oscillation theorem presented in Section 2.1. By using (33), the fact that the pdf of a chi-square is a positive defined kernel [26], and Theorem 1, the number of sign changes of $\widetilde{\Xi}'(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|})$ is upper-bounded by the number of sign changes of

$$G_{\sigma_1,\sigma_2,\mathsf{R},n}(y) = \widetilde{M}_2(y) - M_1(y), \tag{39}$$

for $y \in \mathbb{R}^+$. Note that

$$G_{\sigma_1,\sigma_2,\mathsf{R},n}(y) \geq -\frac{1}{\sigma_2^2} + \frac{1}{\sigma_1^2} - \frac{\mathsf{R}}{\sigma_1^2 y}\mathsf{h}_{\frac{n}{2}}\left(\frac{\mathsf{R}}{\sigma_1^2}y\right) \tag{40}$$

$$\geq -\frac{1}{\sigma_2^2} + \frac{1}{\sigma_1^2} - \frac{\mathsf{R}^2}{\sigma_1^4 n}, \tag{41}$$

where the inequality in (40) follows from $\mathsf{h}_{\frac{n}{2}}(x) \geq 0$ for $x \geq 0$, and (41) follows from $\mathsf{h}_{\frac{n}{2}}(x) \leq \frac{x}{n}$ for $x \geq 0$ and $n \in \mathbb{N}$. We conclude by noting that (41) is nonnegative, hence has no sign change, for

$$\mathsf{R} < \sigma_1^2\sqrt{n\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)} \tag{42}$$

for all $y \in \mathbb{R}^+$, thus guaranteeing that $P_{\mathbf{X}_\mathsf{R}}$ is secrecy-capacity-achieving. □

Remark 1. *As a consequence of the proof of Theorem 2, for any $\mathsf{R} \geq 0, \sigma_2 \geq \sigma_1 \geq 0$ and $n \in \mathbb{N}$, if $G_{\sigma_1,\sigma_2,\mathsf{R},n}(y)$ has at most one sign change, then $P_{\mathbf{X}_\mathsf{R}}$ is secrecy-capacity-achieving if, and only if, for all $\|\mathbf{x}\| = \mathsf{R}$*

$$\Xi(\mathbf{0}; P_{\mathbf{X}_\mathsf{R}}) \leq \Xi(\mathbf{x}; P_{\mathbf{X}_\mathsf{R}}). \tag{43}$$

Because of the difficulty in evaluating analytical properties of (39), proving that $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ has at most one sign change does not seem easy. However, in Appendix A, we show via extensive numerical evaluations that $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ changes sign at most once for any $n, \mathsf{R}, \sigma_1, \sigma_2$ that we tried.

3.2. Characterizing the Low-Amplitude Regime

Let us characterize the low-amplitude regime as follows.

Theorem 3. *Consider a function*

$$f(\mathsf{R}) = \int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\sqrt{s}\mathbf{Z}\|}{s}\mathsf{R}\right) + \mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathsf{R}+\sqrt{s}\mathbf{Z}\|}{s}\mathsf{R}\right)\right] - 1}{s^2}\mathrm{d}s \tag{44}$$

where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$. If $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ of (39) has at most one sign change, the input $\mathbf{X}_\mathsf{R}$ is secrecy-capacity-achieving if, and only if, $\mathsf{R} \leq \bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$, where $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$ is given as the solution of

$$f(\mathsf{R}) = 0. \tag{45}$$

Remark 2. *Note that (45) always has a solution. To see this, observe that $f(0) = \frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} < 0$ and $f(\infty) = \frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} > 0$. Moreover, the solution is unique because $f(\mathsf{R})$ monotonically increases for $\mathsf{R} \geq 0$.*

The solution to (45) needs to be found numerically. To avoid any loss of accuracy in the numerical evaluation of $\mathsf{h}_v(x)$ for large values of x, we used the exponential scaling provided in the MATLAB implementation of $\mathsf{I}_v(x)$. Since evaluating $f(\mathsf{R})$ is rather straightforward and not time-consuming, we opted for a binary search algorithm.

In Table 1, we show the values of $\bar{\mathsf{R}}_n(1, \sigma_2^2)$ for some values of σ_2^2 and n. Moreover, we report the values of $\bar{\mathsf{R}}_n^{\mathrm{ptp}}(1)$ and $\bar{\mathsf{R}}_n^{\mathrm{MMSE}}(1)$ from [27] in the first and the last row, respectively. As predicted by (12), we can appreciate the close match of the $\bar{\mathsf{R}}_n^{\mathrm{ptp}}(1)$ row with the one of $\bar{\mathsf{R}}_n(1, 1000)$. Similarly, the agreement between the $\bar{\mathsf{R}}_n^{\mathrm{MMSE}}(1)$ row and the $\bar{\mathsf{R}}_n(1, 1.001)$ row is justified by (16).

Table 1. Values of $\bar{\mathsf{R}}_n^{\mathrm{MMSE}}(1)$, $\bar{\mathsf{R}}_n(1, \sigma_2^2)$, and $\bar{\mathsf{R}}_n^{\mathrm{ptp}}(1)$.

n	**MMSE**	σ_2^2				**ptp**
		1.001	**1.5**	**10**	**1000**	
1	1.057	1.057	1.161	1.518	1.664	1.666
2	1.535	1.535	1.687	2.221	2.450	2.454
3	1.908	1.909	2.098	2.768	3.061	3.065
4	2.223	2.224	2.444	3.229	3.575	3.580
5	2.501	2.501	2.750	3.634	4.026	4.031
6	2.751	2.752	3.025	3.999	4.432	4.438
7	2.981	2.982	3.278	4.334	4.805	4.811
8	3.195	3.196	3.513	4.646	5.151	5.158
9	3.395	3.396	3.733	4.937	5.475	5.483
10	3.585	3.586	3.941	5.213	5.781	5.789

Table 1. *Cont.*

n	MMSE	σ_2^2				ptp
		1.001	**1.5**	**10**	**1000**	
11	3.765	3.766	4.139	5.475	6.072	6.080
12	3.936	3.938	4.328	5.725	6.350	6.359
13	4.101	4.102	4.509	5.964	6.616	6.625
14	4.259	4.260	4.683	6.195	6.872	6.881
15	4.412	4.413	4.851	6.417	7.119	7.128
16	4.560	4.561	5.013	6.632	7.357	7.367
17	4.702	4.704	5.170	6.839	7.588	7.598
18	4.841	4.842	5.323	7.041	7.812	7.823
19	4.976	4.977	5.471	7.238	8.030	8.041
20	5.107	5.109	5.616	7.429	8.242	8.254
21	5.235	5.237	5.756	7.615	8.449	8.461
22	5.360	5.362	5.894	7.797	8.651	8.663
23	5.483	5.484	6.028	7.974	8.848	8.860
24	5.602	5.603	6.159	8.148	9.041	9.054
25	5.719	5.720	6.288	8.318	9.230	9.243
26	5.834	5.835	6.414	8.485	9.416	9.428
27	5.946	5.948	6.538	8.649	9.597	9.610
28	6.056	6.058	6.659	8.809	9.775	9.789
29	6.165	6.166	6.778	8.967	9.951	9.964
30	6.271	6.273	6.895	9.122	10.123	10.136
31	6.376	6.378	7.010	9.274	10.292	10.306
32	6.479	6.481	7.124	9.424	10.458	10.472
33	6.580	6.582	7.235	9.571	10.622	10.636
34	6.680	6.682	7.345	9.717	10.783	10.798
35	6.779	6.780	7.453	9.860	10.942	10.957

3.3. Large n Asymptotics

We now use the result in Theorem 3 to characterize the asymptotic behavior of $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$. In particular, it is shown that $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$ increases as $\sqrt{n}$.

Theorem 4. *For* $\sigma_1^2 \leq \sigma_2^2$

$$\lim_{n\to\infty} \frac{\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)}{\sqrt{n}} = c(\sigma_1^2, \sigma_2^2), \tag{46}$$

where $c = c(\sigma_1^2, \sigma_2^2)$ *is the solution of*

$$\int_{\sigma_1^2}^{\sigma_2^2} \frac{\frac{c^2}{\left(\frac{\sqrt{s}}{2}+\sqrt{\frac{s}{4}+c^2}\right)^2} + \frac{c^2(c^2+s)}{\left(\frac{s}{2}+\sqrt{\frac{s^2}{4}+c^2(c^2+s)}\right)^2} - 1}{s^2} \mathrm{d}s = 0. \tag{47}$$

Proof. See Section 7. □

In Figure 1, for $\sigma_1^2 = 1$ and $\sigma_2^2 = 1.001, 1.5, 10, 1000$, we show the behavior of $\bar{\mathsf{R}}_n(1, \sigma_2^2)/\sqrt{n}$ and how its asymptotic converges to $c(1, \sigma_2^2)$.

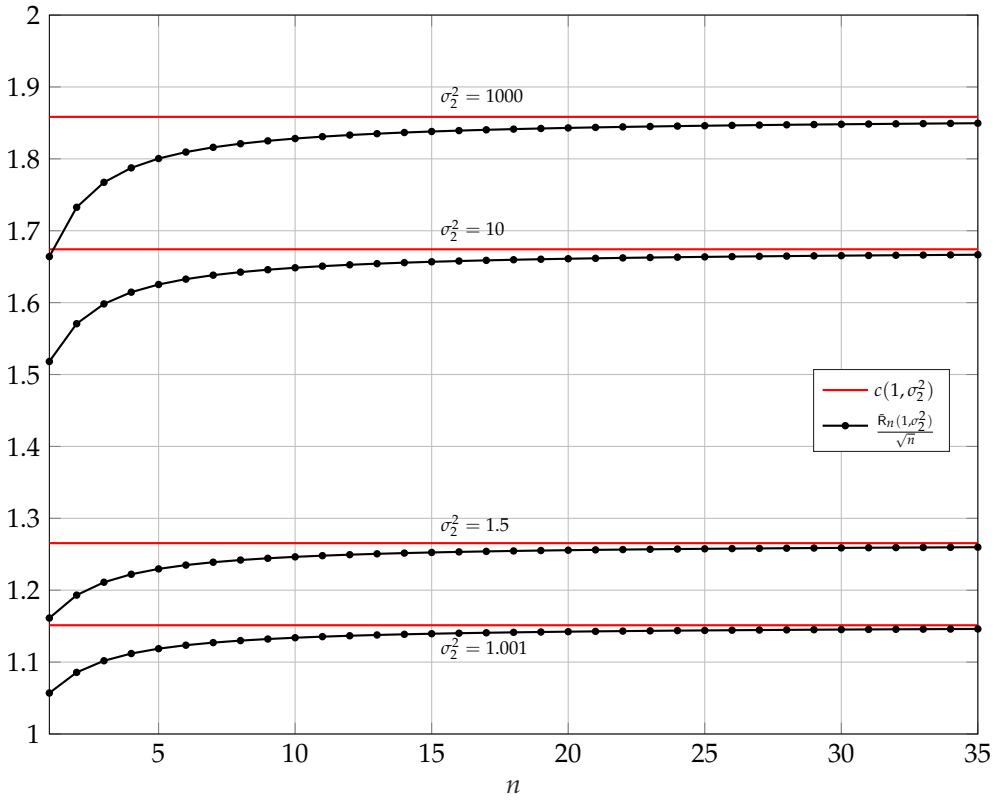

Figure 1. Asymptotic behavior of $\bar{\mathsf{R}}_n(1, \sigma_2^2)/\sqrt{n}$ versus n for $\sigma_1^2 = 1$ and $\sigma_2^2 = 1.001, 1.5, 10, 1000$. In red, we show $c(1, \sigma_2^2)$ defined in (46).

3.4. Scalar Case $(n = 1)$

For the scalar case, the optimal input distribution $P_{X^\star}$ is discrete. In this regime, we provide an implicit and an explicit upper bound on the number of support points of the optimal input probability mass function (pmf) $P_{X^\star}$.

Theorem 5. *Let $Y_1^\star$ and $Y_2^\star$ be the secrecy-capacity-achieving output distributions at the legitimate and malicious receivers, respectively, and let*

$$g(y) = \mathbb{E}\left[\log \frac{f_{Y_2^\star}(y+N)}{f_{Y_1^\star}(y)}\right], \qquad y \in \mathbb{R}, \tag{48}$$

with $N \sim \mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$. For $\mathsf{R} > 0$, an implicit upper bound on the number of support points of $P_{X^\star}$ is

$$|\mathsf{supp}(P_{X^\star})| \leq \mathrm{N}([-L, L], g(\cdot) + \kappa_1) < \infty \tag{49}$$

where

$$\kappa_1 = \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s, \tag{50}$$

$$L = \mathsf{R}\frac{\sigma_2 + \sigma_1}{\sigma_2 - \sigma_1} + \sqrt{\frac{\frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2} + 2C_s}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}}}. \tag{51}$$

Moreover, an explicit upper bound on the number of support points of $P_{X^\star}$ is obtained by using

$$\mathrm{N}([-L, L], g(\cdot) + \kappa_1) \leq \rho\frac{\mathsf{R}^2}{\sigma_1^2} + O(\log(\mathsf{R})), \tag{52}$$

where $\rho = (2\mathrm{e}+1)^2\left(\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}\right)^2 + \left(\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}+1\right)^2$.

The upper bounds in Theorem 5 are generalizations of the upper bounds on the number of points presented in [30] in the context of a point-to-point AWGN channel with an amplitude constraint. Indeed, if we let $\sigma_2 \to \infty$, while keeping σ_1 and R fixed, then the wiretap channel reduces to the AWGN point-to-point channel.

To find a lower bound on the number of mass points, a possible approach consists of the following steps:

$$\begin{aligned} C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, 1) &= I(X^\star; Y_1) - I(X^\star; Y_2) && (53)\\ &\leq H(X^\star) - I(X^\star; Y_2) && (54)\\ &\leq \log(|\mathsf{supp}(P_{X^\star})|) - I(X^\star; Y_2), && (55) \end{aligned}$$

where the above uses the nonnegativity of the entropy and the fact that entropy is maximized by a uniform distribution. Furthermore, by using a suboptimal uniform (continuous) distribution on $[-\mathsf{R}, \mathsf{R}]$ as an input and the entropy power inequality, the secrecy capacity is lower-bounded by

$$C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, 1) \geq \frac{1}{2}\log\left(1 + \frac{\frac{2\mathsf{R}^2}{\pi \mathrm{e}\sigma_1^2}}{1+\frac{\mathsf{R}^2}{\sigma_2^2}}\right). \tag{56}$$

Combining the bounds in (55) and (56), we arrive at the following lower bound on the number of points:

$$|\mathsf{supp}(P_{X^\star})| \geq \sqrt{1 + \frac{\frac{2\mathsf{R}^2}{\pi \mathrm{e}\sigma_1^2}}{1+\frac{\mathsf{R}^2}{\sigma_2^2}} \mathrm{e}^{I(X^\star;Y_2)}}. \tag{57}$$

At this point, one needs to determine the behavior of $I(X^\star; Y_2)$. A trivial lower bound on $|\mathsf{supp}(P_{X^\star})|$ can be found by lower-bounding $I(X^\star; Y_2)$ by zero. However, this lower bound on $|\mathsf{supp}(P_{X^\star})|$ does not grow with R, while the upper bound does increase with R. A possible way of establishing a lower bound that increases in R is by showing that $I(X^\star; Y_2) \approx \frac{1}{2}\log\left(1+\frac{\mathsf{R}^2}{\sigma_2^2}\right)$. However, because not much is known about the structure of the optimal input distribution $P_{X^\star}$, it is not immediately evident how one can establish such an approximation or whether it is valid.

4. Secrecy Capacity Expression in the Low-Amplitude Regime

The result in Theorem 3 can also be used to establish the secrecy capacity for all $\mathsf{R} \leq \bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$, as is performed next.

Theorem 6. *If* $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ *of* (39) *has at most one sign change and if* $\mathsf{R} \leq \bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$*, then*

$$C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n) = \frac{1}{2}\int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathsf{R}^2 - \mathsf{R}^2\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathsf{R}+\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right]}{s^2}\mathrm{d}s. \tag{58}$$

Proof. See Section 9. □

Large n Asymptotics

It is important to note that as $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$ grows as $\sqrt{n}$, according to Theorem 4, when we keep R constant and increase the number of antennas to infinity, the low-amplitude regime becomes the only regime. The next theorem characterizes the secrecy capacity in this 'massive-MIMO' regime (i.e., where R is fixed and n goes to infinity).

Theorem 7. *Consider the expression in* (58) *and fix* $\mathsf{R} \geq 0$ *and* $\sigma_1^2 \leq \sigma_2^2$*, then*

$$\lim_{n\to\infty} C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n) = \mathsf{R}^2 \left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_2^2} \right). \tag{59}$$

Proof. See Appendix C. □

Remark 3. *The result in Theorem 7 is reminiscent of the capacity in the wideband regime [31, Ch. 9], where the capacity increases linearly in the signal-to-noise ratio. Similarly, Theorem 7 shows that in the large antenna regime, the secrecy capacity grows linearly with the difference in the single-to-noise ratio between the legitimate user and the eavesdropper.*

In Theorem 7, R was held fixed. It is also interesting to study the case when R is a function of n. Specifically, it is interesting to study the case when $\mathsf{R} = c\sqrt{n}$ for some coefficient c.

Theorem 8. *Suppose that* $c \leq c(\sigma_1^2, \sigma_2^2)$*. Then,*

$$\lim_{n\to\infty} \frac{C_s(\sigma_1^2, \sigma_2^2, c\sqrt{n}, n)}{n} = \frac{1}{2} \log \left(\frac{1 + c^2/\sigma_1^2}{1 + c^2/\sigma_2^2} \right). \tag{60}$$

Proof. See Appendix D. □

Notice that (60) is equivalent to the secrecy capacity of a vector Gaussian wiretap channel subject to an average power constraint. Gaussian wiretap channels under average power constraints have been extensively investigated [3,32] and, for an average power constraint $\mathbb{E}[\|\mathbf{X}\|^2] \leq \mathsf{P}$, the resulting secrecy capacity is given by [3]

$$C_G(\sigma_1^2, \sigma_2^2, \mathsf{P}, n) = \frac{n}{2} \log \frac{1 + \mathsf{P}/\sigma_1^2}{1 + \mathsf{P}/\sigma_2^2}. \tag{61}$$

Thus, the result in (60) can be restated as

$$\lim_{n\to\infty} \frac{C_s(\sigma_1^2, \sigma_2^2, c\sqrt{n}, n)}{C_G(\sigma_1^2, \sigma_2^2, c^2, n)} = 1. \tag{62}$$

In other words, for the regime considered in Theorem 8, for a large enough n the secrecy capacity under the amplitude constraint $\mathsf{R}_n = c\sqrt{n}$ behaves as the secrecy capacity under the average power constraint c^2.

5. Beyond the Low-Amplitude Regime

To evaluate the secrecy capacity and find the optimal distribution $P_{\mathbf{X}^\star}$ beyond $\bar{\mathsf{R}}_n$ we rely on numerical estimations. We remark that, as pointed out in [25], the secrecy-capacity-achieving distribution is isotropic and consists of finitely many co-centric shells. Keeping this in mind, we can find the optimal input distribution $P_{\mathbf{X}^\star}$ by just optimizing over $P_{\|\mathbf{X}\|}$ with $\|\mathbf{X}\| \leq \mathsf{R}$.

5.1. Numerical Algorithm

In the case of scalar Gaussian wiretap channels, the secrecy capacity and the optimal input pmf can be estimated via the algorithm described in [33], i.e., a numerical procedure that takes inspiration from the deterministic annealing algorithm sketched in [34]. Let us denote by $\hat{C}_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$ the numerical estimate of the secrecy capacity, and by $\hat{P}_{\|\mathbf{X}^\star\|}$, the estimate of the optimal pmf on the input norm. To numerically evaluate $\hat{C}_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$ and $\hat{P}_{\|\mathbf{X}^\star\|}$, we extend to the vector case the algorithm in [33]. Our extension is defined in Algorithm 1. The input parameters of the main function are the noise variances σ_1^2 and

σ_2^2, the radius R, the vectors $\boldsymbol{\rho}$ and $\mathbf{p}$ being, respectively, the mass points positions and probabilities of a tentative input pmf, the number of iterations in the while loop N_c, and finally, a tolerance ε to set the precision of the secrecy capacity estimate.

Algorithm 1 Secrecy capacity and optimal input pmf estimation

procedure MAIN$(\sigma_1^2, \sigma_2^2, \mathsf{R}, \boldsymbol{\rho}, \mathbf{p}, N_c, \varepsilon)$
  repeat
    $k \leftarrow 0$
    while $k < N_c$ **do**
      $k \leftarrow k + 1$
      $\boldsymbol{\rho} \leftarrow$ GRADIENT ASCENT$(\boldsymbol{\rho}, \mathbf{p})$
      $\mathbf{p} \leftarrow$ BLAHUT–ARIMOTO$(\boldsymbol{\rho}, \mathbf{p})$
    end while
    valid $\leftarrow$ KKT VALIDATION$(\boldsymbol{\rho}, \mathbf{p}, \varepsilon)$
    if valid = False **then**
      $(\boldsymbol{\rho}, \mathbf{p}) \leftarrow$ ADD-POINT$(\boldsymbol{\rho}, \mathbf{p})$
    end if
  until valid = True
  $\hat{P}_{\|\mathbf{X}^\star\|} \leftarrow (\boldsymbol{\rho}, \mathbf{p})$
  $\hat{C}_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n) \leftarrow I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}^\star\|})$
  return $\hat{P}_{\|\mathbf{X}^\star\|}, \hat{C}_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$
end procedure

At its core, the numerical procedure iteratively refines its estimate of $P_{\|\mathbf{X}^\star\|}$ by running a gradient ascent algorithm to update the vector $\boldsymbol{\rho}$ and a variant of the Blahut–Arimoto algorithm [35] to update $\mathbf{p}$.

The GRADIENT ASCENT procedure uses the secrecy information as the objective function and stops either when $\boldsymbol{\rho}$ has reached convergence or at a given maximum number of iterations. Let us denote by $I_s(\|\mathbf{X}\|; P_{\|\mathbf{X}\|})$ the secrecy information as a function of the input norm. Notice that, given a tentative pmf $\hat{P}_{\|\mathbf{X}\|}$ of mass points $\boldsymbol{\rho}$, probabilities $\mathbf{p}$, and $|\mathsf{supp}(\hat{P}_{\|\mathbf{X}\|})| = K$, we have

$$I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}\|}) = \sum_{i=1}^{K} p_i \cdot \widetilde{\Xi}\Big(\rho_i\,; \hat{P}_{\|\mathbf{X}\|}\Big), \tag{63}$$

where $\widetilde{\Xi}(t; \hat{P}_{\|\mathbf{X}\|})$ is the secrecy density, with respect to the input norm, defined in (A9) and where p_i and ρ_i are, respectively, the ith element of $\mathbf{p}$ and $\boldsymbol{\rho}$. Then, the GRADIENT ASCENT updates are given by

$$\rho_i = \rho_i + \alpha \cdot \frac{\partial}{\partial \rho_i} I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}\|}), \quad i = 1, \ldots, K, \tag{64}$$

where the partial derivatives are defined in Appendix E and α is the step size in the gradient ascent. We remark that, to ensure convergence to a local maximum, we use the gradient ascent algorithm in a backtracking line search version [36]. By suitably adjusting the step size α at each iteration, the backtracking line search version guarantees us that each new update of $\boldsymbol{\rho}$ provides a nondecreasing associated secrecy information, compared to the previous update of $\boldsymbol{\rho}$.

The BLAHUT–ARIMOTO function runs a variant of the Blahut–Arimoto algorithm. For the scalar case, an example of the Blahut–Arimoto optimization, applied to wiretap channels, is given in [37]. Similar results can be extended to the case of vector wiretap channels. Given the current probabilities p_i's, the updates are obtained by evaluating

$$p_i' = p_i \exp\Big(\widetilde{\Xi}\Big(\rho_i; \hat{P}_{\|\mathbf{X}\|}\Big)\Big), \quad i = 1, \ldots, K, \tag{65}$$

and finally, by normalizing each p'_i and assigning them to the entries of the vector $\mathbf{p}$

$$p_i = \frac{p'_i}{\sum_{k=1}^{K} p'_i}, \quad i = 1, \dots, K. \tag{66}$$

Similarly to GRADIENT ASCENT, the BLAHUT–ARIMOTO procedure stops either when the values of $\mathbf{p}$ have reached a stable convergence or after a set number of updates.

Since the joint optimization of $\boldsymbol{\rho}$ and $\mathbf{p}$ is not numerically feasible, we need to reiterate both the BLAHUT–ARIMOTO and the GRADIENT ASCENT procedures a given number of times, namely N_c. The parameter N_c is chosen empirically in such a way that $\boldsymbol{\rho}$ and $\mathbf{p}$ become fairly stable, and therefore we can expect to have reached joint convergence for both of them.

Then, the KKT VALIDATION procedure ensures that the values of $\boldsymbol{\rho}$ and $\mathbf{p}$ are indeed close to the optimal ones. We check the optimality of $\hat{P}_{\|\mathbf{X}\|}$ by verifying whether the KKT conditions in Lemma 2 are satisfied. Since the algorithm has to verify the KKT conditions numerically, i.e., with finite precision, we find it more convenient to check the negated version of (28), where a tolerance parameter ε is introduced that trades off accuracy with computational burden. Specifically, $\hat{P}_{\|\mathbf{X}\|}$ is not an optimal input pmf if any of the following conditions are satisfied:

$$\left| \widetilde{\Xi}(t; \hat{P}_{\|\mathbf{X}\|}) - I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}\|}) \right| > \varepsilon, \quad \text{for some } t \in \mathsf{supp}(\hat{P}_{\|\mathbf{X}\|}) \tag{67a}$$

$$I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}\|}) + \varepsilon < \widetilde{\Xi}(t; \hat{P}_{\|\mathbf{X}\|}), \quad \text{for some } t \in [0, \mathsf{R}]. \tag{67b}$$

Note that in (67), in place of the secrecy capacity $C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$, which is unknown, we used the secrecy information given by the tentative pmf $\hat{P}_{\|\mathbf{X}\|}$, i.e., $I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}\|})$. Condition (67a) is derived by negating (28a): there exists a $t \in \mathsf{supp}(\hat{P}_{\|\mathbf{X}\|})$, such that $\widetilde{\Xi}(t; \hat{P}_{\|\mathbf{X}\|})$ is ε-away from the secrecy information $I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}\|})$. Condition (67b) is the negated version of (28b): there exists a $t \in [0, \mathsf{R}]$ such that $\widetilde{\Xi}(t; \hat{P}_{\|\mathbf{X}\|})$ is at least ε-larger than the secrecy information $I_s(\|\mathbf{X}\|; \hat{P}_{\|\mathbf{X}\|})$. With some abuse of notation, we refer to (67) as to the ε-KKT conditions. If the tentative pmf $\hat{P}_{\|\mathbf{X}\|}$ does not pass the check of the ε-KKT conditions, then the algorithm checks whether a new point has to be added to the pmf.

The ADD POINT procedure evaluates the position of the new mass point

$$\rho_{\text{new}} = \arg \max_{t \in [0, \mathsf{R}]} \widetilde{\Xi}(t; \hat{P}_{\|\mathbf{X}\|}). \tag{68}$$

The point ρ_{new} is appended to the vector $\boldsymbol{\rho}$ and the probabilities $\mathbf{p}$ are set to be equiprobable.

The whole procedure is repeated until KKT VALIDATION gives a positive outcome, and at that point the algorithm returns $\hat{P}_{\|\mathbf{X}^\star\|}$ as the optimal pmf estimate and $\hat{C}_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$ as the secrecy capacity estimate.

Remark 4. *In this work, we focus on the secrecy capacity and on the secrecy-capacity-achieving input distribution. However, it is possible to study other points of the rate-equivocation region of the degraded wiretap Gaussian channel by suitably changing the KKT conditions, as reported in [21], Equations (33) and (34). With the due modifications, the proposed optimization algorithm can find the optimal input distribution for any point of the rate-equivocation region.*

5.2. Numerical Results

In Figure 2, we show with black dots the numerical estimate $\hat{C}_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$ versus R, evaluated via Algorithm 1, for $\sigma_1^2 = 1$, $\sigma_2^2 = 1.5, 10$, $n = 2, 4$, and tolerance $\varepsilon = 10^{-6}$. For the same values of σ_1^2, σ_2^2, and n we also show, with the red lines, the analytical low-amplitude regime secrecy capacity $C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$ versus R from Theorem 6. In addition,

we show with blue dotted lines the secrecy capacity under the average power constraint $\mathbb{E}\left[\|\mathbf{X}\|^2\right] \leq \mathsf{R}^2$:

$$C_G(\sigma_1^2, \sigma_2^2, \mathsf{R}^2, n) = \frac{n}{2} \log \frac{1 + \mathsf{R}^2/\sigma_1^2}{1 + \mathsf{R}^2/\sigma_2^2} \geq C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n), \tag{69}$$

where the inequality follows by noting that the average power constraint $\mathbb{E}\left[\|\mathbf{X}\|^2\right] \leq \mathsf{R}^2$ is weaker than the amplitude constraint $\|\mathbf{X}\| \leq \mathsf{R}$. Finally, the dashed vertical lines show $\bar{\mathsf{R}}_n$, i.e., the upper limit of the low-amplitude regime, for the considered values of σ_1^2, σ_2^2, and n.

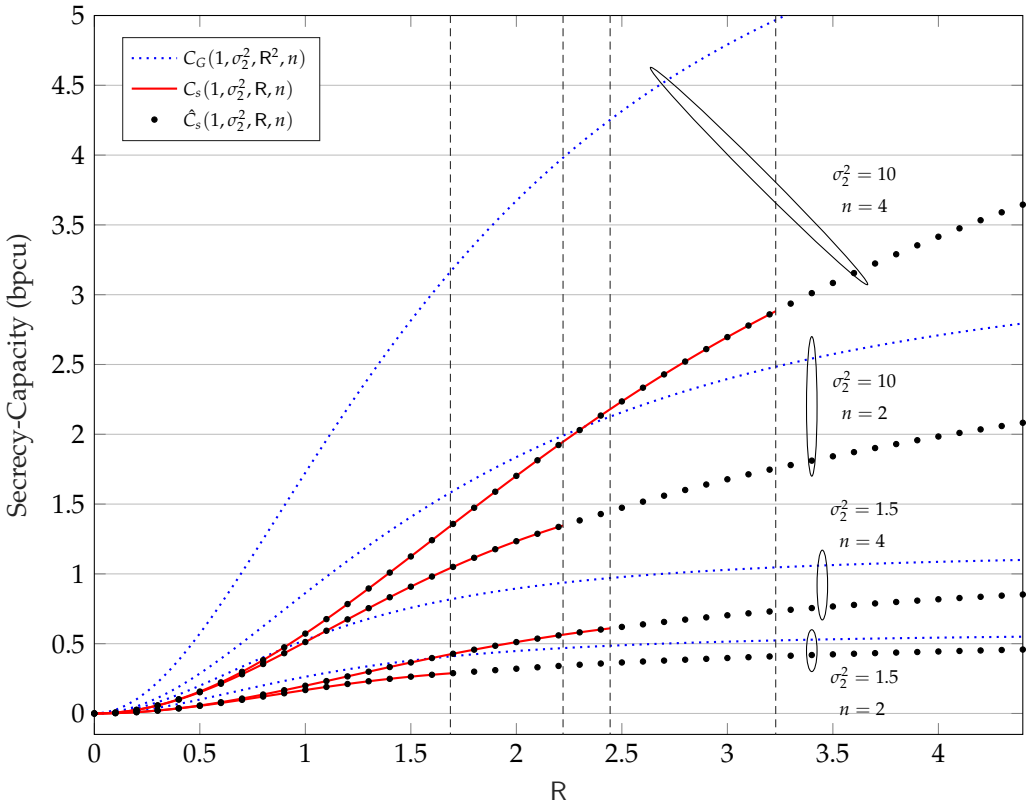

Figure 2. Secrecy capacity in bit per channel use (bpcu) versus R for $\sigma_2^2 = 1.5, 10$ and $n = 2, 4$. The secrecy capacity under average power constraints $C_G(\sigma_1^2, \sigma_2^2, \mathsf{R}^2, n)$ is defined in (69), while under peak power constraints, i.e., $C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n)$, is defined in (58).

In Figure 3, we consider discrete values for R and for each value of R we plot the corresponding estimated pmf $\hat{P}_{\|\mathbf{X}^\star\|}$, evaluated via Algorithm 1, for $\sigma_1^2 = 1$, $\sigma_2^2 = 1.5$, $n = 2, 8$, and tolerance $\varepsilon = 10^{-6}$. The figure shows, at each R, the normalized amplitude of support points in the estimated pmf, while the size of the circles qualitatively shows the probability associated with each support point. Similarly, Figure 4 shows the evolution of the pmf estimate for $\sigma_1^2 = 1$, $\sigma_2^2 = 10$, $n = 2, 8$, and $\varepsilon = 10^{-6}$. It is interesting to notice how in both Figures 3 and 4 when a new mass point is added to the pmf, it appears in zero. Moreover, the mass point of radius R always seems to be optimal.

Finally, Figure 5 shows the output distributions of the legitimate user and of the eavesdropper in the case of $\sigma_1^2 = 1$, $\sigma_2^2 = 10$, $n = 2$, and for two values of R. At the top of the figure, the distributions are shown for $\mathsf{R} = 2.25$, which is a value close to $\bar{\mathsf{R}}_2(1, 10)$. At the bottom of the figure, the distributions are shown for $\mathsf{R} = 7.5$. For both values of R, the legitimate user sees an output distribution where the co-centric rings of the input distribution are easily distinguishable. On the other hand, as expected, the output distribution seen by the eavesdropper is close to a Gaussian.

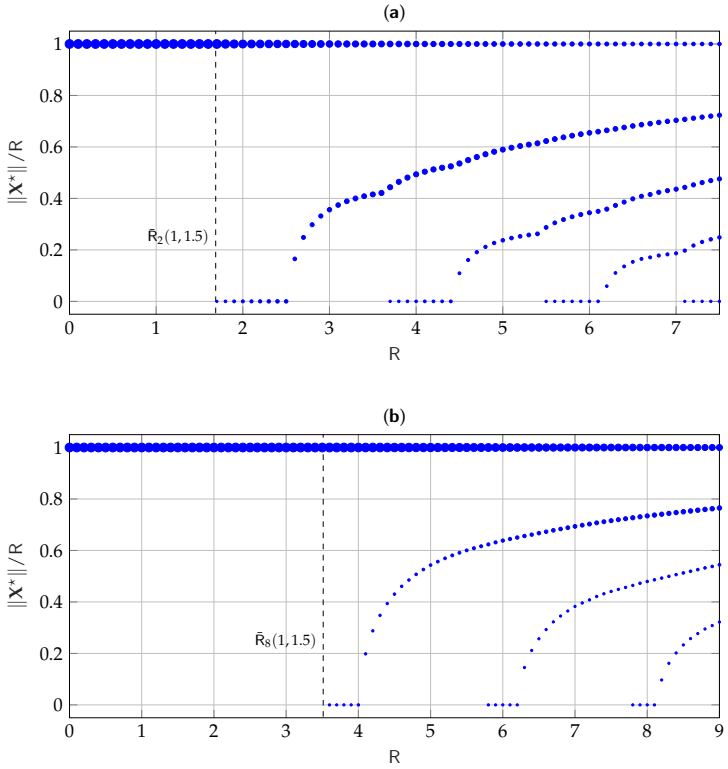

Figure 3. Evolution of the numerically estimated $\hat{P}_{\|\mathbf{X}^\star\|}$ versus R for $\sigma_1^2 = 1$, $\sigma_2^2 = 1.5$, (**a**) $n = 2$, and (**b**) $n = 8$.

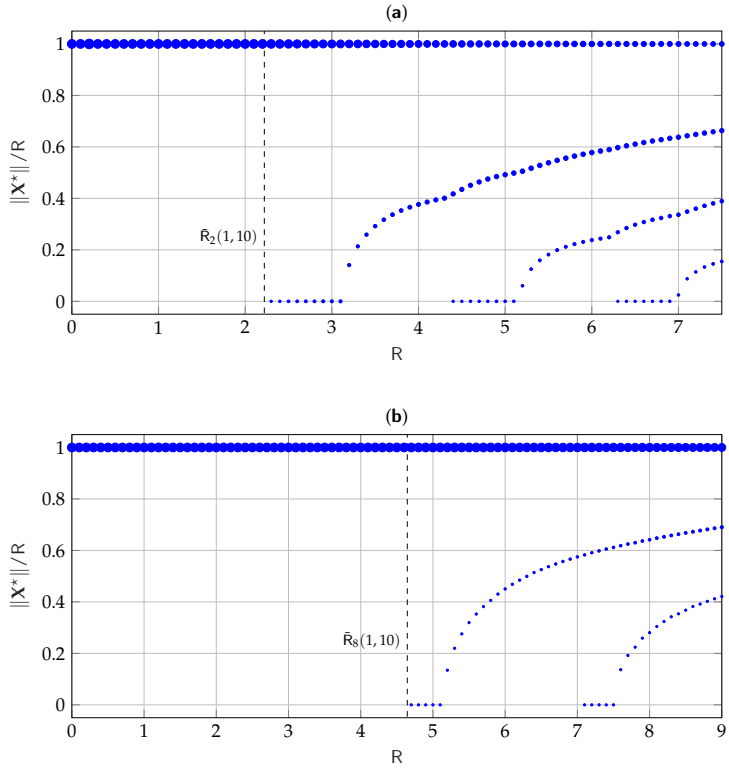

Figure 4. Evolution of the numerically estimated $\hat{P}_{\|\mathbf{X}^\star\|}$ versus R for $\sigma_1^2 = 1$, $\sigma_2^2 = 10$, (**a**) $n = 2$, and (**b**) $n = 8$.

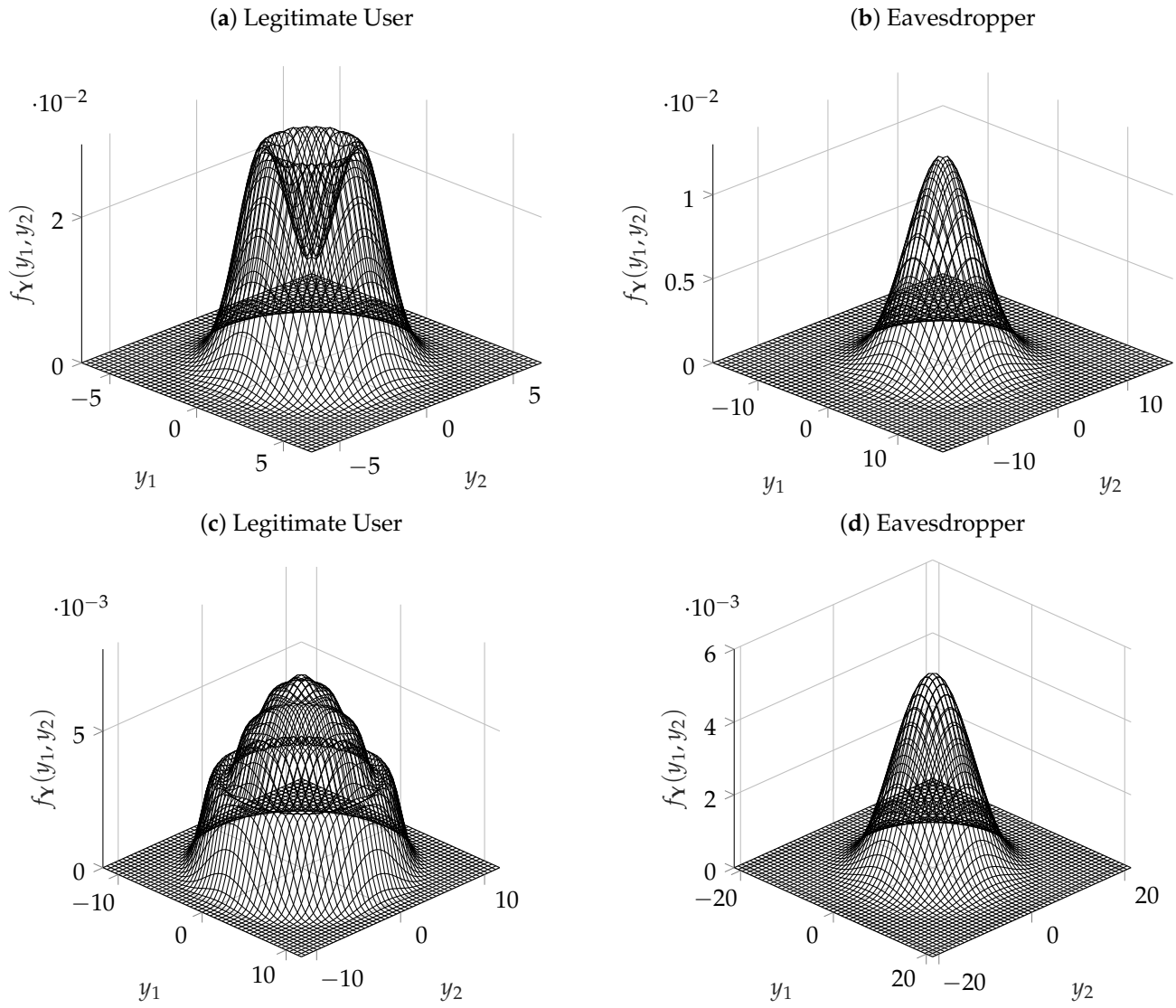

Figure 5. Output pdf of the legitimate user and of the eavesdropper for $\sigma_1^2 = 1$, $\sigma_2^2 = 10$, $n = 2$, (**a**,**b**) $\mathsf{R} = 2.25$, and (**c**,**d**) $\mathsf{R} = 7.5$. An animation showing the evolution of the output pdf as R varies can be found in [1].

6. Proof of Theorem 3

Estimation Theoretic Representation

By Remark 1, if $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ has at most one sign change, $P_{\mathbf{X}_\mathsf{R}}$ is secrecy-capacity-achieving if, and only if, for all $\|\mathbf{x}\| = \mathsf{R}$

$$\Xi(\mathbf{0}; P_{\mathbf{X}_\mathsf{R}}) \leq \Xi(\mathbf{x}; P_{\mathbf{X}_\mathsf{R}}). \tag{70}$$

We seek to re-write the condition (70) in the estimation theoretic form. To that end, we need the following representation of the relative entropy [38]:

$$\mathsf{D}(P_{\mathbf{X}_1+\sqrt{t}\mathbf{Z}} \| P_{\mathbf{X}_2+\sqrt{t}\mathbf{Z}}) = \frac{1}{2}\int_t^\infty \frac{g(s)}{s^2}\mathrm{d}s, \tag{71}$$

where

$$g(s) = \mathbb{E}\Big[\|\mathbf{X}_1 - \ell_2(\mathbf{X}_1 + \sqrt{s}\mathbf{Z})\|^2\Big] - \mathbb{E}\Big[\|\mathbf{X}_1 - \ell_1(\mathbf{X}_1 + \sqrt{s}\mathbf{Z})\|^2\Big] \tag{72}$$

and where

$$\ell_i(\mathbf{y}) = \mathbb{E}[\mathbf{X}_i|\mathbf{X}_i + \sqrt{s}\mathbf{Z} = \mathbf{y}] \tag{73}$$

$$= \int \mathbf{x}_i f_{\mathbf{X}_i|\mathbf{X}_i+\sqrt{s}\mathbf{Z}}(\mathbf{x}_i \mid \mathbf{y})\mathrm{d}\mathbf{x}_i,\ i \in \{1,2\}. \tag{74}$$

Another fact that will be important for our expression is

$$\mathbb{E}\left[\mathbf{X}_{\mathsf{R}} \mid \mathbf{X}_{\mathsf{R}} + \sqrt{s}\mathbf{Z} = \mathbf{y}\right] = \frac{\mathsf{R}\mathbf{y}}{\|\mathbf{y}\|}\mathsf{h}_{\frac{n}{2}}\left(\frac{\|\mathbf{y}\|\mathsf{R}}{s}\right), \tag{75}$$

see, for example [27], for the proof.

Next, using (71) and (75) note that for any $\|\mathbf{x}\| = \mathsf{R}$ we have that for $i \in \{1,2\}$

$$\mathsf{D}(P_{\mathbf{x}+\sqrt{\sigma_i^2}\mathbf{Z}}\|P_{\mathbf{X}_{\mathsf{R}}+\sqrt{\sigma_i^2}\mathbf{Z}}) = \frac{1}{2}\int_{\sigma_i^2}^{\infty} \frac{\mathbb{E}\left[\left\|\mathbf{x} - \frac{\mathsf{R}(\mathbf{x}+\sqrt{s}\mathbf{Z})}{\|\mathbf{x}+\sqrt{s}\mathbf{Z}\|}\mathsf{h}_{\frac{n}{2}}\left(\frac{\|\mathbf{x}+\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right\|^2\right]}{s^2}\mathrm{d}s \tag{76}$$

$$= \frac{1}{2}\int_{\sigma_i^2}^{\infty} \frac{\mathbb{E}\left[\|\mathbf{x}\|^2\right] - \mathbb{E}\left[\left\|\frac{\mathsf{R}(\mathbf{x}+\sqrt{s}\mathbf{Z})}{\|\mathbf{x}+\sqrt{s}\mathbf{Z}\|}\mathsf{h}_{\frac{n}{2}}\left(\frac{\|\mathbf{x}+\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right\|^2\right]}{s^2}\mathrm{d}s \tag{77}$$

$$= \frac{1}{2}\int_{\sigma_i^2}^{\infty} \frac{\mathsf{R}^2 - \mathsf{R}^2\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathbf{x}+\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right]}{s^2}\mathrm{d}s, \tag{78}$$

where (77) follows from

$$\mathrm{mmse}(\mathbf{X}_{\mathsf{R}}|\mathbf{Y}) = \mathbb{E}\left[\|\mathbf{X}_{\mathsf{R}} - \mathbb{E}[\mathbf{X}_{\mathsf{R}}|\mathbf{Y}]\|^2\right] \tag{79}$$

$$= \mathbb{E}\left[\|\mathbf{X}_{\mathsf{R}}\|^2\right] - \mathbb{E}\left[\|\mathbb{E}[\mathbf{X}_{\mathsf{R}}|\mathbf{Y}]\|^2\right]. \tag{80}$$

Moreover, for $\|\mathbf{x}\| = 0$, it holds

$$\mathsf{D}(P_{\mathbf{0}+\sqrt{\sigma_i^2}\mathbf{Z}}\|P_{\mathbf{X}_{\mathsf{R}}+\sqrt{\sigma_i^2}\mathbf{Z}}) = \frac{1}{2}\int_{\sigma_i^2}^{\infty} \frac{\mathsf{R}^2\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\mathsf{R}\|\mathbf{Z}\|}{s}\right)\right]}{s^2}\mathrm{d}s. \tag{81}$$

Now, note that by using the definition of $\Xi(\mathbf{x}; P_{\mathbf{X}_{\mathsf{R}}})$ in (30), (78), and (81) we have that for $\|\mathbf{x}\| = \mathsf{R}$

$$\Xi(\mathbf{x}; P_{\mathbf{X}_{\mathsf{R}}}) = \mathsf{D}(P_{\mathbf{x}+\sqrt{\sigma_1^2}\mathbf{Z}}\|P_{\mathbf{X}_{\mathsf{R}}+\sqrt{\sigma_1^2}\mathbf{Z}}) - \mathsf{D}(P_{\mathbf{x}+\sqrt{\sigma_2^2}\mathbf{Z}}\|P_{\mathbf{X}_{\mathsf{R}}+\sqrt{\sigma_2^2}\mathbf{Z}}) \tag{82}$$

$$= \frac{1}{2}\int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathsf{R}^2 - \mathsf{R}^2\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathbf{x}+\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right]}{s^2}\mathrm{d}s, \tag{83}$$

and

$$\Xi(\mathbf{0}; P_{\mathbf{X}_{\mathsf{R}}}) = \mathsf{D}(P_{\mathbf{0}+\sqrt{\sigma_1^2}\mathbf{Z}}\|P_{\mathbf{X}_{\mathsf{R}}+\sqrt{\sigma_1^2}\mathbf{Z}}) - \mathsf{D}(P_{\mathbf{0}+\sqrt{\sigma_2^2}\mathbf{Z}}\|P_{\mathbf{X}_{\mathsf{R}}+\sqrt{\sigma_2^2}\mathbf{Z}}) \tag{84}$$

$$= \frac{1}{2}\int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathsf{R}^2\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right]}{s^2}\mathrm{d}s \tag{85}$$

Consequently, the necessary and sufficient condition in Theorem 2 can be equivalently written as

$$\int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right) + \mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathbf{x}+\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right] - 1}{s^2}\mathrm{d}s \leq 0. \tag{86}$$

Now $\bar{\mathsf{R}}_n(\sigma_1^2, \sigma_2^2)$ will be the largest R that satisfies (86), which concludes the proof of Theorem 3.

7. Proof of Theorem 4

The objective of the proof is to understand how the condition in (45) behaves as $n \to \infty$. To study the large n behavior, we need to the following bounds on the h_ν [39,40]: for $\nu > \frac{1}{2}$

$$\mathsf{h}_\nu(x) = \frac{x}{\frac{2\nu-1}{2} + \sqrt{\frac{(2\nu-1)^2}{4} + x^2}} \cdot g_\nu(x), \tag{87}$$

where

$$1 \geq g_\nu(x) \geq \frac{\frac{2\nu-1}{2} + \sqrt{\frac{(2\nu-1)^2}{4} + x^2}}{\nu + \sqrt{\nu^2 + x^2}}. \tag{88}$$

Now let $\mathsf{R} = c\sqrt{n}$ for some $c > 0$. The goal is to understand the behavior of

$$\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right) + \mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathbf{x} + \sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right] \tag{89}$$

as n goes to infinity. First, let

$$V_n = \frac{\|\mathbf{Z}\|}{\sqrt{n}}, \tag{90}$$

and note that

$$\lim_{n\to\infty} \mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\sqrt{s}\mathbf{Z}\|c\sqrt{n}}{s}\right)\right] = \lim_{n\to\infty} \mathbb{E}\left[\left(\frac{\frac{cV_n}{\sqrt{s}}}{\frac{n-1}{2n} + \sqrt{\frac{(n-1)^2}{4n^2} + \left(\frac{cV_n}{\sqrt{s}}\right)^2}} \cdot g_{\frac{n}{2}}\left(\frac{cV_n}{\sqrt{s}}n\right)\right)^2\right] \tag{91}$$

$$= \mathbb{E}\left[\lim_{n\to\infty}\left(\frac{\frac{cV_n}{\sqrt{s}}}{\frac{n-1}{2n} + \sqrt{\frac{(n-1)^2}{4n^2} + \left(\frac{cV_n}{\sqrt{s}}\right)^2}} \cdot g_{\frac{n}{2}}\left(\frac{cV_n}{\sqrt{s}}n\right)\right)^2\right] \tag{92}$$

$$= \frac{c^2}{\left(\frac{\sqrt{s}}{2} + \sqrt{\frac{s}{4} + c^2}\right)^2}, \tag{93}$$

where (92) follows from the dominated convergence theorem, and (93) follows since, by the law of large numbers we have, almost surely,

$$\lim_{n\to\infty} V_n^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Z_i^2 = \mathbb{E}[Z^2] = 1. \tag{94}$$

Second, let

$$W_n = \frac{\|\mathbf{x} + \sqrt{s}\mathbf{Z}\|}{\sqrt{n}}, \tag{95}$$

where, without loss of generality, we take $\mathbf{x} = [\mathsf{R}, 0, \ldots, 0]$

$$\lim_{n\to\infty} \mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathbf{x} + \sqrt{s}\mathbf{Z}\|c\sqrt{n}}{s}\right)\right] = \lim_{n\to\infty} \mathbb{E}\left[\left(\frac{\frac{cW_n}{s} \cdot g_{\frac{n}{2}}\left(\frac{cW_n}{s}n\right)}{\frac{n-1}{2n} + \sqrt{\frac{(n-1)^2}{4n^2} + \left(\frac{cW_n}{s}\right)^2}}\right)^2\right] \tag{96}$$

$$= \mathbb{E}\left[\lim_{n\to\infty}\left(\frac{\frac{cW_n}{s}\cdot g_{\frac{n}{2}}\left(\frac{cW_n}{s}n\right)}{\frac{n-1}{2n}+\sqrt{\frac{(n-1)^2}{4n^2}+\left(\frac{cW_n}{s}\right)^2}}\right)^2\right] \tag{97}$$

$$= \frac{c^2(c^2+s)}{\left(\frac{s}{2}+\sqrt{\frac{s^2}{4}+c^2(c^2+s)}\right)^2}, \tag{98}$$

where (97) follows from the dominated convergence theorem and where (98) follows since, by the strong law of large numbers we have, almost surely,

$$\lim_{n\to\infty} W_n^2 = \lim_{n\to\infty}\frac{1}{n}(\sqrt{s}Z_1+c\sqrt{n})^2 + s\lim_{n\to\infty}\frac{1}{n}\sum_{i=2}^{n} Z_i^2 \tag{99}$$

$$= c^2+s. \tag{100}$$

Combining (93) and (98) with (45), we arrive at

$$\int_{\sigma_1^2}^{\sigma_2^2} \frac{\frac{c^2}{\left(\frac{\sqrt{s}}{2}+\sqrt{\frac{s}{4}+c^2}\right)^2}+\frac{c^2(c^2+s)}{\left(\frac{s}{2}+\sqrt{\frac{s^2}{4}+c^2(c^2+s)}\right)^2}-1}{s^2}\mathrm{d}s = 0. \tag{101}$$

8. Proof of Theorem 5

8.1. Implicit Upper Bound

A consequence of the KKT conditions of Lemma 1 is the inclusion

$$\mathsf{supp}(P_{X^\star}) \subseteq \{x\in[-\mathsf{R},\mathsf{R}] : \Xi(x) - C_s = 0\} \tag{102}$$

which suggests the following upper bound on the number of support points of $P_{X^\star}$:

$$|\mathsf{supp}(P_{X^\star})| \le \mathrm{N}\Big([-\mathsf{R},\mathsf{R}], \Xi(x) - C_s(\sigma_1^2,\sigma_2^2,\mathsf{R},1)\Big) \tag{103}$$

$$= \mathrm{N}\left([-\mathsf{R},\mathsf{R}], \mathbb{E}\left[g(Y_1)+\log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s \middle| X=x\right]\right) \tag{104}$$

$$\le \mathscr{S}\left(g(\cdot)+\log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \tag{105}$$

$$\le \mathrm{N}\left(\mathbb{R}, g(\cdot)+\log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \tag{106}$$

$$= \mathrm{N}\left([-L,L], g(\cdot)+\log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \tag{107}$$

$$< \infty, \tag{108}$$

where (104) follows from using (21); (105) follows from applying Karlin's oscillation Theorem 1 and the fact that the Gaussian pdf is a strictly totally positive kernel, which was shown in [26]; (107) is proved in Lemma A3 in the Appendix B; and (108) follows because $g(\cdot)$ is an analytic function in $(-L,L)$. The implicit upper bound (49) of Theorem 5 follows from (107) and (108).

8.2. Explicit Upper Bound

The key to finding an explicit upper bound on the number of zeros will be the following complex-analytic result.

Lemma 3 (Tijdeman's Number of Zeros Lemma [41])**.** *Let L, s, t be positive numbers, such that $s > 1$. For the complex valued function $f \neq 0$, which is analytic on $|z| < (st + s + t)L$, its number of zeros $\mathrm{N}(\mathcal{D}_L, f)$ within the disk $\mathcal{D}_L = \{z \colon |z| \leq L\}$ satisfies*

$$\mathrm{N}(\mathcal{D}_L, f) \leq \frac{1}{\log s}\left(\log \max_{|z|\leq(st+s+t)L} |f(z)| - \log \max_{|z|\leq tL} |f(z)|\right). \tag{109}$$

Furthermore, the following loosened version of the implicit upper bound in (49) will be useful.

Lemma 4.

$$|\mathsf{supp}(P_{X^\star})| \leq \mathrm{N}([-L, L], h(\cdot)) + 1 \tag{110}$$

where

$$\frac{h(y)}{\sigma_1^2 f_{Y_1}(y)} = \frac{\mathbb{E}_N[\mathbb{E}[X^\star|Y_2 = y + N]] - y}{\sigma_2^2} - \frac{\mathbb{E}[X^\star|Y_1 = y] - y}{\sigma_1^2} \tag{111}$$

$$= \frac{\mathbb{E}\big[N \log f_{Y_2}(y + N)\big]}{\sigma_2^2 - \sigma_1^2} - \frac{\mathbb{E}[X^\star|Y_1 = y] - y}{\sigma_1^2}, \tag{112}$$

and where $N \sim \mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$.

Proof. Starting from (107), we can write

$$|\mathsf{supp}(P_{X^\star})| \leq \mathrm{N}\left([-L, L], g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \tag{113}$$

$$\leq \mathrm{N}\big([-L, L], g'(\cdot)\big) + 1 \tag{114}$$

$$= \mathrm{N}\Big([-L, L], \sigma_1^2 f_{Y_1}(\cdot) g'(\cdot)\Big) + 1 \tag{115}$$

where in step (114), we applied Rolle's theorem, and in step (115), we used the fact that multiplying by a strictly positive function (i.e., $\sigma_1^2 f_{Y_1}$) does not change the number of zeros. The first derivative of g can be computed as follows:

$$g'(y) = \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_2}(y + N)\right] - \frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_1}(y) \tag{116}$$

$$= \frac{\mathbb{E}_N[\mathbb{E}[X^\star|Y_2 = y + N]] - y}{\sigma_2^2} - \frac{\mathbb{E}[X^\star|Y_1 = y] - y}{\sigma_1^2}, \tag{117}$$

where in the last step, we used the well-known Tweedie's formula (see for example [42,43]):

$$\mathbb{E}[X^\star|Y_i = y] = y + \sigma_i^2 \frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_i}(y). \tag{118}$$

An alternative expression for the first term in the right-hand side (RHS) of (116) is as follows:

$$\mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_2}(y + N)\right] = \int_{-\infty}^{\infty} f_N(n) \frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_2}(y + n) \mathrm{d}n \tag{119}$$

$$= -\int_{-\infty}^{\infty} \left(\frac{\mathrm{d}}{\mathrm{d}n} f_N(n)\right) \cdot \log f_{Y_2}(y + n) \mathrm{d}n \tag{120}$$

$$= \int_{-\infty}^{\infty} \frac{n}{\sigma_2^2 - \sigma_1^2} f_N(n) \cdot \log f_{Y_2}(y + n) \mathrm{d}n \tag{121}$$

$$= \frac{1}{\sigma_2^2 - \sigma_1^2} \mathbb{E}\big[N \log f_{Y_2}(y + N)\big], \tag{122}$$

where $f_N(n) = \phi_{\sqrt{\sigma_2^2-\sigma_1^2}}(n)$. The proof is concluded by letting

$$h(y) \triangleq \sigma_1^2 f_{Y_1}(y) g'(y). \tag{123}$$

□

To apply Tijdeman's number of zeros Lemma, upper and lower bounds to the maximum module of the complex analytic extension of h over the disk $\mathcal{D}_L = \{z : |z| \le L\}$ are proposed in Lemmas A4 and A5 in the Appendix B. Using those bounds, we can provide an upper bound on the number of mass points as follows:

$$\begin{aligned}
&\mathrm{N}([-L, L], h(\cdot)) \\
&\quad \le \mathrm{N}\left(\mathcal{D}_L, \check{h}(\cdot)\right) && (124)\\
&\quad \le \min_{s>1, t>0} \left\{ \frac{\log \frac{\max_{|z|\le (st+s+t)L} |\check{h}(z)|}{\max_{|z|\le tL} |\check{h}(z)|}}{\log s} \right\} && (125)\\
&\quad \le \log \frac{\frac{\mathrm{e}^{\frac{(2\mathrm{e}+1)^2 L^2}{2\sigma_1^2}}}{\sqrt{2\pi\sigma_1^2}} \left(a_1(2\mathrm{e}+1)^2 L^2 + a_2(2\mathrm{e}+1)L + a_3\right)}{(c_1 L - c_2 \mathsf{R}) \frac{\exp\left(-\frac{(L+\mathsf{R})^2}{2\sigma_1^2}\right)}{\sqrt{2\pi\sigma_1^2}}} && (126)\\
&\quad = \frac{(2\mathrm{e}+1)^2 L^2}{2\sigma_1^2} + \frac{(L+\mathsf{R})^2}{2\sigma_1^2} + \log \frac{a_1(2\mathrm{e}+1)^2 L^2 + a_2(2\mathrm{e}+1)L + a_3}{c_1 L - c_2 \mathsf{R}} && (127)\\
&\quad = \frac{(2\mathrm{e}+1)^2 (d_1 \mathsf{R} + d_2)^2}{2\sigma_1^2} + \frac{((d_1+1)\mathsf{R} + d_2)^2}{2\sigma_1^2} \\
&\qquad + \log \frac{a_1(2\mathrm{e}+1)^2 (d_1\mathsf{R}+d_2)^2 + a_2(2\mathrm{e}+1)(d_1\mathsf{R}+d_2) + a_3}{(c_1 d_1 - c_2)\mathsf{R} + c_1 d_2} && (128)\\
&\quad \le b_1 \frac{\mathsf{R}^2}{\sigma_1^2} + b_2 + \log \frac{b_3 \mathsf{R}^2 + b_4 \mathsf{R} + b_5}{b_6 \mathsf{R} + b_7} && (129)\\
&\quad \le b_1 \frac{\mathsf{R}^2}{\sigma_1^2} + O(\log(\mathsf{R})), && (130)
\end{aligned}$$

where (124) follows because extending to a larger domain can only increase the number of zeros; (125) follows from the Tijdeman's Number of Zeros Lemma; (126) follows from choosing $s = \mathrm{e}$ and $t = 1$ and using bounds in Lemmas A4 and A5; (128) follows from using the value of L in (A38); (129) using the bound $(a+b)^2 \le 2(a^2+b^2)$ and defining

$$\begin{aligned}
b_1 &= (2\mathrm{e}+1)^2 d_1^2 + (d_1+1)^2 && (131a)\\
&= (2\mathrm{e}+1)^2 \left(\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}\right)^2 + \left(\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}+1\right)^2 && (131b)\\
b_2 &= \frac{((2\mathrm{e}+1)^2+1) d_2^2}{\sigma_1^2} && (131c)\\
&= \frac{((2\mathrm{e}+1)^2+1)}{\sigma_1^2} \frac{\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2} + 2C_s}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}} && (131d)\\
&= ((2\mathrm{e}+1)^2+1)\left(1 + 2\frac{\sigma_2^2}{\sigma_2^2-\sigma_1^2} C_s\right) && (131e)\\
b_3 &= 2(2\mathrm{e}+1)^2 a_1 d_1^2 && (131f)
\end{aligned}$$

$$= 2(2e+1)^2 \frac{3\sigma_1^2}{\sigma_2^2\sqrt{\sigma_2^2-\sigma_1^2}} \left(\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}\right)^2 \tag{131g}$$

$$b_4 = (2e+1)d_1 a_2 \tag{131h}$$

$$= (2e+1)\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}\left(\frac{\sqrt{2}\sigma_1^2}{\sqrt{\sigma_2^2}\sqrt{\sigma_2^2-\sigma_1^2}}+2\right) \tag{131i}$$

$$b_5 = 2(2e+1)^2 a_1 d_2^2 + (2e+1)a_2 d_2 + a_3 \tag{131j}$$

$$= 2(2e+1)^2 \frac{3\sigma_1^2}{\sigma_2^2\sqrt{\sigma_2^2-\sigma_1^2}} \left(\frac{\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}+2C_s}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}}\right)$$
$$+ (2e+1)\left(\frac{\sqrt{2}\sigma_1^2}{\sqrt{\sigma_2^2}\sqrt{\sigma_2^2-\sigma_1^2}}+2\right)\sqrt{\frac{\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}+2C_s}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}}}$$
$$+ \frac{\sigma_1^2}{\sqrt{\sigma_2^2-\sigma_1^2}} \cdot \sqrt{|\log(2\pi\sigma_2^2)|^2 + \frac{24(\sigma_2^2-\sigma_1^2)^2}{\sigma_2^4}+\pi^2} \tag{131k}$$

$$b_6 = c_1 d_1 - c_2 \tag{131l}$$

$$= \frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1} - \frac{\sigma_2^2+\sigma_1^2}{\sigma_2^2} = 2\frac{\sigma_1}{\sigma_2} \tag{131m}$$

$$b_7 = c_1 d_2 \tag{131n}$$

$$= \frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}\sqrt{\frac{\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}+2C_s}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}}}; \tag{131o}$$

and (130) follows from the fact that the b_1, b_3, b_4, and b_6 coefficients do not depend on R and the fact that the coefficients b_2, b_5, and b_4, while they do depend on R through C_s, do not grow with R. The fact that C_s does not grow with R follows from the bound in (69).

Finally, the explicit upper bound on the number of support points of $P_{X^\star}$ in (52) is a consequence of (130).

9. Proof of Theorem 6

Using the KKT conditions in (28), we have that for $\mathbf{x} = [\mathsf{R}, 0, \ldots, 0]$

$$C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n) = \Xi(\mathbf{x}; P_{\mathbf{X}_\mathsf{R}}) \tag{132}$$

$$= \mathsf{D}(f_{\mathbf{Y}_1|\mathbf{X}}(\cdot|\mathbf{x})\|f_{\mathbf{Y}_1^\star}) - \mathsf{D}(f_{\mathbf{Y}_2|\mathbf{X}}(\cdot|\mathbf{x})\|f_{\mathbf{Y}_2^\star}) \tag{133}$$

$$= \frac{1}{2}\int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathsf{R}^2 - \mathsf{R}^2\mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2\left(\frac{\|\mathsf{R}+\sqrt{s}\mathbf{Z}\|\mathsf{R}}{s}\right)\right]}{s^2}\mathrm{d}s \tag{134}$$

where the last expression was computed in (83). This concludes the proof.

10. Conclusions

This paper has focused on the secrecy capacity of the n-dimensional vector Gaussian wiretap channel under the peak power (or amplitude constraint) in a so-called low (but not vanishing) amplitude regime. In this regime, the optimal input distribution $P_{\mathbf{X}_\mathsf{R}}$ is supported on a single n-dimensional sphere of radius R. The paper has identified the largest $\bar{\mathsf{R}}_n$, such that the distribution $P_{\mathbf{X}_\mathsf{R}}$ is optimal. In addition, the asymptotic of $\bar{\mathsf{R}}_n$ has been completely characterized as dimension n approaches infinity. As a by-product of the analysis, the capacity in the low-amplitude regime has also been characterized in a more or less closed form. The paper has also provided a number of supporting numerical examples.

Implicit and explicit upper bounds have been proposed on the number of mass points for the optimal input distribution $P_{X^\star}$ in the scalar case with $n = 1$.

There are several interesting future directions. For example, one interesting direction would be to determine a regime in which a mixture of a mass point at zero and $P_{\mathbf{X}_\mathsf{R}}$ is optimal. It would also be interesting to establish a lower bound on the number of mass points in the support of the optimal input distribution when $n = 1$. We note that such a lower bound was obtained for a point-to-point channel in [30]. We finally remark that the extension of the results of this paper to nondegraded wiretap channels is not trivial and also constitutes an interesting but ambitious future direction.

Author Contributions: A.F., L.B. and A.D. contributed equally to this work. All authors have read and agreed to the published version of the manuscript. Part of this work was presented at the 2021 IEEE Information Theory Workshop [44], at the 2022 IEEE International Symposium on Information Theory [45], at the 2022 IEEE International Mediterranean Conference on Communications and Networking [33], and in the PhD dissertation in [46].

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable

Data Availability Statement: Datasets for the numerical results provided in this work are available at [1].

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Examples of the Function $G_{\sigma_1,\sigma_2,\mathsf{R},n}$

In this section, we give supporting numerical arguments that the function $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ defined in (39) has at most one sign change. Figure A1 demonstrates the behavior of the function $G_{\sigma_1,\sigma_2,\mathsf{R},n}$. In addition, the code that generates the function $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ for various values of n, σ_1, and σ_2 is provided in [1].

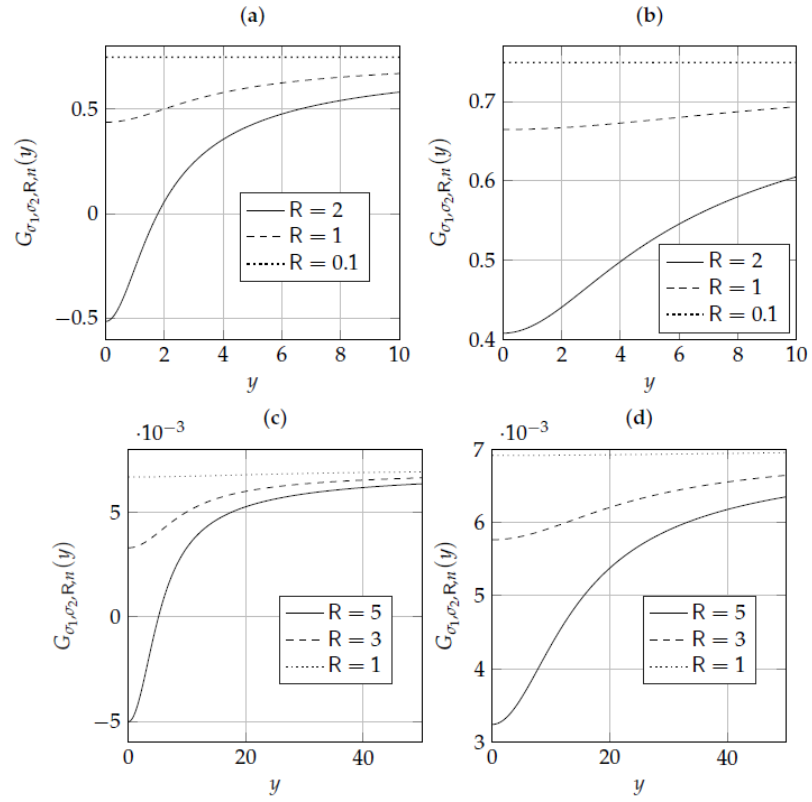

Figure A1. Examples of the function $G_{\sigma_1,\sigma_2,\mathsf{R},n}$ defined in (39). (**a**) $n = 3$, $\sigma_1 = 1$, and $\sigma_2 = 2$. (**b**) $n = 11$, $\sigma_1 = 1$, and $\sigma_2 = 2$. (**c**) $n = 4$, $\sigma_1 = 3$, and $\sigma_2 = 3.1$. (**d**) $n = 11$, $\sigma_1 = 3$, and $\sigma_2 = 3.1$.

Appendix B. Derivative of the Secrecy-Density

Lemma A1. *The derivative of the secrecy density for the input $P_{\mathbf{X}_\mathsf{R}}$ is*

$$\widetilde{\Xi}'(\|\mathbf{x}\|; P_{\|\mathbf{X}_\mathsf{R}\|}) = \|\mathbf{x}\|\, \mathbb{E}\Big[\widetilde{M}_2(\sigma_1 Q_{n+2}) - M_1(\sigma_1 Q_{n+2})\Big] \tag{A1}$$

where Q^2_{n+2} is a noncentral chi-square random variable with $n+2$ degrees of freedom and noncentrality parameter $\frac{\|\mathbf{x}\|^2}{\sigma_1^2}$ and

$$M_i(y) = \frac{1}{\sigma_i^2}\left(\frac{\mathsf{R}}{y}\mathsf{h}_{\frac{n}{2}}\left(\frac{\mathsf{R}}{\sigma_i^2}y\right) - 1\right), \qquad i \in \{1,2\} \tag{A2}$$

$$\widetilde{M}_2(y) = \mathbb{E}[M_2(\|y + \mathbf{W}\|)], \tag{A3}$$

where $\mathbf{W} \sim \mathcal{N}(\mathbf{0}_{n+2}, (\sigma_2^2 - \sigma_1^2)\mathbf{I}_{n+2})$.

Proof. We start with the secrecy density expressed in spherical coordinates. A quick way to obtain the information densities in this coordinate system is to note that:

$$\begin{aligned}
&I(\mathbf{X}; \mathbf{Y}_i)\\
&= h(\mathbf{Y}_i) - h(\mathbf{N}_i) && \text{(A4)}\\
&= h(\|\mathbf{Y}_i\|) + (n-1)\mathbb{E}[\log\|\mathbf{Y}_i\|] + h_\lambda\left(\frac{\mathbf{Y}_i}{\|\mathbf{Y}_i\|}\right) - h(\mathbf{N}_i) && \text{(A5)}\\
&= h(\|\mathbf{Y}_i\|^2) + \Big(\frac{n}{2} - 1\Big)\mathbb{E}[\log\|\mathbf{Y}_i\|^2] + \log\frac{\pi^{\frac{n}{2}}}{\Gamma\big(\frac{n}{2}\big)} - \frac{n}{2}\log(2\pi e\sigma_i^2) && \text{(A6)}\\
&= h\left(\sigma_i^2\left\|\frac{\mathbf{X}}{\sigma_i} + \widetilde{\mathbf{N}}_i\right\|^2\right) + \Big(\frac{n}{2} - 1\Big)\mathbb{E}\left[\log\left(\sigma_i^2\left\|\frac{\mathbf{X}}{\sigma_i} + \widetilde{\mathbf{N}}_i\right\|^2\right)\right] + \log\frac{\pi^{\frac{n}{2}}}{\Gamma\big(\frac{n}{2}\big)} - \frac{n}{2}\log(2\pi e\sigma_i^2) && \text{(A7)}\\
&= h\left(\left\|\frac{\mathbf{X}}{\sigma_i} + \widetilde{\mathbf{N}}_i\right\|^2\right) + \Big(\frac{n}{2} - 1\Big)\mathbb{E}\left[\log\left\|\frac{\mathbf{X}}{\sigma_i} + \widetilde{\mathbf{N}}_i\right\|^2\right] - \log\Big((2e)^{\frac{n}{2}}\Gamma\Big(\frac{n}{2}\Big)\Big), && \text{(A8)}
\end{aligned}$$

where (A5) holds by [47], Lemma 6.17, and by independence between $\|\mathbf{Y}_i\|$ and $\frac{\mathbf{Y}_i}{\|\mathbf{Y}_i\|}$; the term $h_\lambda(\cdot)$ is a differential entropy-like quantity for random vectors on the n-dimensional unit sphere ([47], Lemma 6.16); (A6) holds because $\frac{\mathbf{Y}_i}{\|\mathbf{Y}_i\|}$ is uniform on the unit sphere and thanks to [47], Lemma 6.15; the term $\Gamma(z)$ is the gamma function; and in (A7) we have $\widetilde{\mathbf{N}}_i \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$. It is now required to write the secrecy density as follows:

$$\widetilde{\Xi}(\|\mathbf{x}\|; P_{\|\mathbf{X}\|}) = i_1(\|\mathbf{x}\|; P_\mathbf{X}) - i_2(\|\mathbf{x}\|; P_\mathbf{X}) \tag{A9}$$

where

$$i_j(\|\mathbf{x}\|; P_\mathbf{X}) = -\int_0^\infty f_{\chi^2_n(\frac{\|\mathbf{x}\|^2}{\sigma_j^2})}(y)\log\frac{\int_0^\mathsf{R} f_{\chi^2_n(\frac{t^2}{\sigma_j^2})}(y)\,dP_{\|\mathbf{X}\|}(t)}{y^{\frac{n}{2}-1}}dy - \log\Big((2e)^{\frac{n}{2}}\Gamma\Big(\frac{n}{2}\Big)\Big), \tag{A10}$$

for $j \in \{1,2\}$. The term $f_{\chi^2_n(\lambda)}(y)$ is the noncentral chi-square pdf with n degrees of freedom and noncentrality parameter λ.

Given two values ρ_1, ρ_2 with $\rho_1 > \rho_2$, write

$$i_j(\rho_1; P_\mathbf{X}) - i_j(\rho_2; P_\mathbf{X}) = \int_0^\infty \left(f_{\chi^2_n(\frac{\rho_1^2}{\sigma_j^2})}(y) - f_{\chi^2_n(\frac{\rho_2^2}{\sigma_j^2})}(y)\right)\log\frac{y^{\frac{n}{2}-1}}{f_{\|\frac{\mathbf{X}}{\sigma_j}\|^2}(y; P_\mathbf{X})}dy \tag{A11}$$

$$= \int_0^\infty \left(F_{\chi^2_n(\frac{\rho_2^2}{\sigma_j^2})}(y) - F_{\chi^2_n(\frac{\rho_1^2}{\sigma_j^2})}(y) \right) \frac{d}{dy} \log \frac{y^{\frac{n}{2}-1}}{f_{\|\frac{\mathbf{Y}}{\sigma_j}\|^2}(y; P_{\mathbf{X}})} dy \tag{A12}$$

where we have integrated by parts and where $F_{\chi^2_n(\lambda)}(y)$ is the cumulative distribution function of $\chi^2_n(\lambda)$. Now notice that

$$\int_0^\infty \left(F_{\chi^2_n(\frac{\rho_2^2}{\sigma_j^2})}(y) - F_{\chi^2_n(\frac{\rho_1^2}{\sigma_j^2})}(y) \right) dy = \frac{\rho_1^2 - \rho_2^2}{\sigma_j^2}. \tag{A13}$$

Since $\chi^2_n(\frac{\rho_1^2}{\sigma_j^2})$ statistically dominates $\chi^2_n(\frac{\rho_2^2}{\sigma_j^2})$, the integrand function in (A13) is always positive. We can introduce an auxiliary output random variable Q_j, for $j \in \{1,2\}$, with pdf

$$f_{Q_j}(y; \rho_1, \rho_2) = \frac{\sigma_j^2}{\rho_1^2 - \rho_2^2} \left(F_{\chi^2_n(\frac{\rho_2^2}{\sigma_j^2})}(y) - F_{\chi^2_n(\frac{\rho_1^2}{\sigma_j^2})}(y) \right), \tag{A14}$$

for $y > 0$, to rewrite (A12) as follows:

$$i_j(\rho_1; P_{\mathbf{X}}) - i_j(\rho_2; P_{\mathbf{X}}) = -\frac{\rho_1^2 - \rho_2^2}{\sigma_j^2} \int_0^\infty f_{Q_j}(y; \rho_1, \rho_2) \frac{d}{dy} \log \frac{f_{\|\frac{\mathbf{Y}}{\sigma_j}\|^2}(y; P_{\mathbf{X}})}{y^{\frac{n}{2}-1}} dy. \tag{A15}$$

We evaluate the derivative in (A15) as:

$$\frac{d}{dy} \log \frac{f_{\|\frac{\mathbf{Y}}{\sigma_j}\|^2}(y; P_{\mathbf{X}})}{y^{\frac{n}{2}-1}}$$

$$= \frac{y^{\frac{n}{2}-1}}{f_{\|\frac{\mathbf{Y}}{\sigma_j}\|^2}(y; P_{\mathbf{X}})} \int_0^{\mathsf{R}} \frac{d}{dy} \frac{f_{\chi^2_n(\frac{t^2}{\sigma_j^2})}(y)}{y^{\frac{n}{2}-1}} dP_{\|\mathbf{X}\|}(t) \tag{A16}$$

$$= \frac{y^{\frac{n}{2}-1}}{f_{\|\frac{\mathbf{Y}}{\sigma_j}\|^2}(y; P_{\mathbf{X}})} \int_0^{\mathsf{R}} \left(\frac{f_{\chi^2_{n-2}(\frac{t^2}{\sigma_j^2})}(y)}{2y^{\frac{n}{2}-1}} - \left(\frac{1}{2} + \frac{\frac{n}{2}-1}{y} \right) \frac{f_{\chi^2_n(\frac{t^2}{\sigma_j^2})}(y)}{y^{\frac{n}{2}-1}} \right) dP_{\|\mathbf{X}\|}(t) \tag{A17}$$

$$= \mathbb{E}\left[\frac{1}{2} \frac{f_{\chi^2_{n-2}(\frac{\|\mathbf{X}\|^2}{\sigma_j^2})}(\frac{\|\mathbf{Y}\|^2}{\sigma_j^2})}{f_{\chi^2_n(\frac{\|\mathbf{X}\|^2}{\sigma_j^2})}(\frac{\|\mathbf{Y}\|^2}{\sigma_j^2})} - \left(\frac{1}{2} + \frac{\frac{n}{2}-1}{\frac{\|\mathbf{Y}\|^2}{\sigma_j^2}} \right) | \frac{\|\mathbf{Y}\|^2}{\sigma_j^2} = y \right] \tag{A18}$$

$$= \mathbb{E}\left[\frac{1}{2} \frac{\|\mathbf{X}\|}{\|\mathbf{Y}\|} \frac{\mathsf{I}_{\frac{n}{2}-2}(\frac{\|\mathbf{X}\|\|\mathbf{Y}\|}{\sigma_j^2})}{\mathsf{I}_{\frac{n}{2}-1}(\frac{\|\mathbf{X}\|\|\mathbf{Y}\|}{\sigma_j^2})} - \left(\frac{1}{2} + \frac{\frac{n}{2}-1}{\frac{\|\mathbf{Y}\|^2}{\sigma_j^2}} \right) | \frac{\|\mathbf{Y}\|^2}{\sigma_j^2} = y \right] \tag{A19}$$

$$= \mathbb{E}\left[\frac{1}{2} \frac{\|\mathbf{X}\|}{\|\mathbf{Y}\|} \mathsf{h}_{\frac{n}{2}} \left(\frac{\|\mathbf{X}\|\|\mathbf{Y}\|}{\sigma_j^2} \right) - \frac{1}{2} | \frac{\|\mathbf{Y}\|^2}{\sigma_j^2} = y \right] \tag{A20}$$

where, in (A16), we used

$$f_{\|\frac{\mathbf{Y}}{\sigma_j}\|^2}(y; P_{\mathbf{X}}) = \int_0^{\mathsf{R}} f_{\chi^2_n(\frac{t^2}{\sigma_j^2})}(y) dP_{\|\mathbf{X}\|}(t); \tag{A21}$$

in (A17), we used the relationship

$$\frac{d}{dy}f_{\chi^2_n(\rho^2)}(y) = \frac{1}{2}f_{\chi^2_{n-2}(\rho^2)}(y) - \frac{1}{2}f_{\chi^2_n(\rho^2)}(y); \tag{A22}$$

and (A20) follows from the recurrence relationship

$$\mathsf{I}_{\nu-1}(z) - \mathsf{I}_{\nu+1}(z) = \frac{2\nu}{z}\mathsf{I}_\nu(z). \tag{A23}$$

Putting together (A15) and (A20), we find

$$i_j(\rho_1; P_{\mathbf{X}}) - i_j(\rho_2; P_{\mathbf{X}}) = -\frac{\rho_1^2 - \rho_2^2}{2\sigma_j^2}\mathbb{E}\left[\mathbb{E}\left[\frac{\|\mathbf{X}\|}{\|\mathbf{Y}\|}\mathsf{h}_{\frac{n}{2}}\left(\frac{\|\mathbf{X}\|\|\mathbf{Y}\|}{\sigma_j^2}\right) - 1|\frac{\|\mathbf{Y}\|^2}{\sigma_j^2} = Q_j\right]\right]. \tag{A24}$$

We are now in the position to compute the derivative of the information density as

$$i_j'(\rho; P_{\mathbf{X}}) = \lim_{h\to 0}\frac{i_j(\rho+h; P_{\mathbf{X}}) - i_j(\rho; P_{\mathbf{X}})}{h} \tag{A25}$$

$$= -\frac{\rho}{\sigma_j^2}\,\mathbb{E}\left[\mathbb{E}\left[\frac{\|\mathbf{X}\|}{\|\mathbf{Y}\|}\mathsf{h}_{\frac{n}{2}}\left(\frac{\|\mathbf{X}\|\|\mathbf{Y}\|}{\sigma_j^2}\right) - 1|\frac{\|\mathbf{Y}\|^2}{\sigma_j^2} = Q'\right]\right], \tag{A26}$$

where $Q' \sim \chi^2_{n+2}(\frac{\rho^2}{\sigma_j^2})$ thanks to Lemma A2.

The final result is obtained by letting

$$\widetilde{\Xi}'(\|\mathbf{x}\|; P_{\|\mathbf{x}\|}) = i_1'(\|\mathbf{x}\|; P_{\mathbf{X}}) - i_2'(\|\mathbf{x}\|; P_{\mathbf{X}}) \tag{A27}$$

and by specializing the result to the input $P_{\mathbf{X}_\mathsf{R}}$. □

Lemma A2. *Consider the pdf $f_{Q_j}(y; \rho_1, \rho_2)$ defined in (A14). For any $\rho \geq 0$ we have*

$$\lim_{h\to 0} f_{Q_j}(y; \rho+h, \rho) = f_{\chi^2_{n+2}(\frac{\rho^2}{\sigma_j^2})}(y), \qquad y > 0. \tag{A28}$$

Proof. Thanks to the definition (A14), we have

$$\lim_{h\to 0} f_{Q_j}(y; \rho+h, \rho)$$

$$= \lim_{h\to 0}\frac{\sigma_j^2}{h(2\rho+h)}\left(F_{\chi^2_n(\frac{\rho^2}{\sigma_j^2})}(y) - F_{\chi^2_n(\frac{(\rho+h)^2}{\sigma_j^2})}(y)\right) \tag{A29}$$

$$= \lim_{h\to 0}\frac{\sigma_j^2}{h(2\rho+h)}\int_0^y\left(f_{\chi^2_n(\frac{\rho^2}{\sigma_j^2})}(t) - f_{\chi^2_n(\frac{(\rho+h)^2}{\sigma_j^2})}(t)\right)dt \tag{A30}$$

$$= \frac{\sigma_j^2}{2\rho}\int_0^y\sum_{i=0}^{\infty}\lim_{h\to 0}\frac{1}{h}\left(\frac{\mathrm{e}^{-\frac{\rho^2}{2\sigma_j^2}}\left(\frac{\rho^2}{2\sigma_j^2}\right)^i}{i!} - \frac{\mathrm{e}^{-\frac{(\rho+h)^2}{2\sigma_j^2}}\left(\frac{(\rho+h)^2}{2\sigma_j^2}\right)^i}{i!}\right)f_{\chi^2_{n+2i}}(t)dt \tag{A31}$$

$$= \frac{\sigma_j^2}{2\rho}\int_0^y\sum_{i=0}^{\infty}\frac{d}{d\rho}\left(\frac{\mathrm{e}^{-\frac{\rho^2}{2\sigma_j^2}}\left(\frac{\rho^2}{2\sigma_j^2}\right)^i}{i!}\right)f_{\chi^2_{n+2i}}(t)dt \tag{A32}$$

$$= \frac{1}{2}\int_0^y \sum_{i=0}^{\infty}\left(-\frac{\mathrm{e}^{-\frac{\rho^2}{2\sigma_j^2}}\left(\frac{\rho^2}{2\sigma_j^2}\right)^i}{i!}+\frac{\mathrm{e}^{-\frac{\rho^2}{2\sigma_j^2}}\left(\frac{\rho^2}{2\sigma_j^2}\right)^{i-1}}{(i-1)!}1(i\geq 1)\right) f_{\chi^2_{n+2i}}(t)dt \tag{A33}$$

$$= \frac{1}{2}\int_0^y \left(-f_{\chi^2_n(\frac{\rho^2}{\sigma_j^2})}(t)+f_{\chi^2_{n+2}(\frac{\rho^2}{\sigma_j^2})}(t)\right)dt \tag{A34}$$

$$= \int_0^y \frac{d}{dt} f_{\chi^2_{n+2}(\frac{\rho^2}{\sigma_j^2})}(t)dt \tag{A35}$$

$$= f_{\chi^2_{n+2}(\frac{\rho^2}{\sigma_j^2})}(y), \tag{A36}$$

where $1(\cdot)$ is the indicator function; in (A31) we used the Poisson-weighted mixture representation of the noncentral chi-square pdf, and in (A35), we used (A22). □

Lemma A3. *There exists some* $L = L(\sigma_1, \sigma_2, \mathsf{R}) < \infty$ *such that*

$$\mathrm{N}\left(\mathbb{R}, g(\cdot)+\log\left(\frac{\sigma_2}{\sigma_1}\right)-C_s\right) = \mathrm{N}\left([-L,L], g(\cdot)+\log\left(\frac{\sigma_2}{\sigma_1}\right)-C_s\right) < \infty. \tag{A37}$$

Furthermore, L can be upper-bounded as follows:

$$L \leq \mathsf{R}d_1 + d_2 \tag{A38}$$

where

$$d_1 = \frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}, \tag{A39}$$

$$d_2 = \sqrt{\frac{\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}+2C_s}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}}} \leq \sqrt{\frac{\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}+2C_G}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}}}, \tag{A40}$$

with

$$C_G(\sigma_1^2, \sigma_2^2, \mathsf{R}^2, 1) = \frac{1}{2}\log\frac{1+\mathsf{R}^2/\sigma_1^2}{1+\mathsf{R}^2/\sigma_2^2}. \tag{A41}$$

Proof. First, note that $C_s \leq C_G$ thanks to (69). Second, for $|y| \geq \mathsf{R}$, we can lower-bound the function g as follows:

$$g(y) = \mathbb{E}\left[\log f_{Y_2^\star}(y+N)\right] - \log f_{Y_1^\star}(y) \tag{A42}$$

$$= \mathbb{E}[\log \mathbb{E}[\phi_{\sigma_2}(y+N-X^\star)|N]] - \log \mathbb{E}[\phi_{\sigma_1}(y-X^\star)] \tag{A43}$$

$$\geq \mathbb{E}[\log \phi_{\sigma_2}(y+N-X^\star)] - \log \mathbb{E}[\phi_{\sigma_1}(y-X^\star)] \tag{A44}$$

$$\geq \log\frac{\sigma_1}{\sigma_2} - \mathbb{E}\left[\frac{(y+N-X^\star)^2}{2\sigma_2^2}\right] + \frac{(|y|-\mathsf{R})^2}{2\sigma_1^2} \tag{A45}$$

$$= \log\frac{\sigma_1}{\sigma_2} - \mathbb{E}\left[\frac{(y-X^\star)^2}{2\sigma_2^2}\right] - \frac{\sigma_2^2-\sigma_1^2}{2\sigma_2^2} + \frac{(|y|-\mathsf{R})^2}{2\sigma_1^2} \tag{A46}$$

$$\geq \log\frac{\sigma_1}{\sigma_2} - \frac{(|y|+\mathsf{R})^2}{2\sigma_2^2} - \frac{\sigma_2^2-\sigma_1^2}{2\sigma_2^2} + \frac{(|y|-\mathsf{R})^2}{2\sigma_1^2}, \tag{A47}$$

where (A44) follows from applying Jensen's inequality and the law of iterated expectation to the first term; (A45) follows from

$$\mathbb{E}[\phi_{\sigma_1}(y - X^\star)] \le \phi_{\sigma_1}(|y| - \mathsf{R}), \qquad |y| \ge \mathsf{R}; \tag{A48}$$

and (A47) follows from $(y - X^\star)^2 \le (|y| + \mathsf{R})^2$ for all $|y| \ge \mathsf{R} \ge |X^\star|$. The RHS of

$$g(y) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s \ge -\frac{(|y| + \mathsf{R})^2}{2\sigma_2^2} - \frac{\sigma_2^2 - \sigma_1^2}{2\sigma_2^2} + \frac{(|y| - \mathsf{R})^2}{2\sigma_1^2} - C_s \tag{A49}$$

is strictly positive when

$$|y| > \frac{\mathsf{R}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) + \sqrt{\frac{4\mathsf{R}^2}{\sigma_1^2\sigma_2^2} + \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)\left(\frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2} + 2C_s\right)}}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}}. \tag{A50}$$

By using the bound $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$, we arrive at

$$|y| \ge \mathsf{R}\frac{\sigma_2 + \sigma_1}{\sigma_2 - \sigma_1} + \sqrt{\frac{\frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2} + 2C_s}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}}}. \tag{A51}$$

This concludes the proof for the bound on L. □

Lemma A4. *Let $\check{h} : \mathbb{C} \to \mathbb{C}$ denote the complex extension of the function h in (123). Then, for $\mathsf{B} \ge \mathsf{R}$, we have that*

$$\max_{|z| \le \mathsf{B}} |\check{h}(z)| \le \frac{1}{\sqrt{2\pi\sigma_1^2}} \mathrm{e}^{\frac{\mathsf{B}^2}{2\sigma_1^2}} \left(a_1\mathsf{B}^2 + a_2\mathsf{B} + a_3\right) \tag{A52}$$

where

$$a_1 = \frac{3\sigma_1^2}{\sigma_2^2\sqrt{\sigma_2^2 - \sigma_1^2}}, \tag{A53}$$

$$a_2 = \frac{\sqrt{2}\sigma_1^2}{\sqrt{\sigma_2^2\sqrt{\sigma_2^2 - \sigma_1^2}}} + 2, \tag{A54}$$

$$a_3 = \frac{\sigma_1^2}{\sqrt{\sigma_2^2 - \sigma_1^2}} \left(\sqrt{|\log(2\pi\sigma_2^2)|^2 + \frac{24(\sigma_2^2 - \sigma_1^2)^2}{\sigma_2^4} + \pi^2}\right). \tag{A55}$$

Proof. Let us denote $z = z_R + iz_I$, where z_R and z_I are real numbers and $i = \sqrt{-1}$ is the imaginary unit. Then, by triangular inequality, we have:

$$|\check{h}(z)| = \left|\frac{\sigma_1^2 f_{Y_1}(z)\mathbb{E}\left[N \log f_{Y_2}(z+N)\right]}{\sigma_2^2 - \sigma_1^2} - \mathbb{E}[X^\star \phi_{\sigma_1}(z - X^\star)] + z f_{Y_1}(z)\right| \tag{A56}$$

$$\le |f_{Y_1}(z)| \left(\frac{\sigma_1^2}{\sigma_2^2 - \sigma_1^2}\mathbb{E}\left[|N| \cdot |\log f_{Y_2}(z+N)|\right] + |z|\right) + \mathbb{E}\left[|X^\star| \cdot |\phi_{\sigma_1}(z - X^\star)|\right]. \tag{A57}$$

Next, let us upper-bound each contribution of (A57). For $|z| \le \mathsf{B}$, we have

$$\left|\log f_{Y_2}(z+n)\right|^2$$

$$= \left|\log\left|f_{Y_2}(z+n)\right| + i\arg(f_{Y_2}(z+n))\right|^2 \tag{A58}$$

$$= \log^2\left|f_{Y_2}(z+n)\right| + \arg^2(f_{Y_2}(z+n)) \tag{A59}$$

$$= \log^2|\mathbb{E}[\phi_{\sigma_2}(z+n-X^\star)]| + \arg^2(\mathbb{E}[\phi_{\sigma_2}(z+n-X^\star)]) \tag{A60}$$

$$\leq \log^2\left(\frac{1}{\sqrt{2\pi\sigma_2^2}}\mathbb{E}\left[\exp\left(-\frac{(z_R+n-X^\star)^2-z_I^2}{2\sigma_2^2}\right)\right]\right) + \arg^2\left(\sum_x \alpha_x \exp(i\theta_x)\right) \tag{A61}$$

$$\leq \left(\frac{z_I^2}{2\sigma_2^2} - \frac{1}{2}\log(2\pi\sigma_2^2) + \log\mathbb{E}\left[\mathrm{e}^{-\frac{(z_R+n-X^\star)^2}{2\sigma_2^2}}\right]\right)^2 + \pi^2 \tag{A62}$$

$$\leq 2\left(\frac{z_I^2}{2\sigma_2^2} - \frac{1}{2}\log(2\pi\sigma_2^2)\right)^2 + 2\log^2\mathbb{E}\left[\mathrm{e}^{-\frac{(z_R+n-X^\star)^2}{2\sigma_2^2}}\right] + \pi^2 \tag{A63}$$

$$\leq 2\left(\frac{z_I^2}{2\sigma_2^2} - \frac{1}{2}\log(2\pi\sigma_2^2)\right)^2 + 2\frac{\mathbb{E}^2\left[(z_R+n-X^\star)^2\right]}{4\sigma_2^4} + \pi^2 \tag{A64}$$

$$\leq 2\left(\frac{z_I^2}{2\sigma_2^2} - \frac{1}{2}\log(2\pi\sigma_2^2)\right)^2 + 2\frac{\left((z_R+n)^2+\mathsf{R}^2\right)^2}{4\sigma_2^4} + \pi^2 \tag{A65}$$

$$\leq \frac{2\mathsf{B}^2}{\sigma_2^2} + |\log(2\pi\sigma_2^2)|^2 + \frac{8(\mathsf{B}^4+n^4)+\mathsf{R}^4}{\sigma_2^4} + \pi^2, \tag{A66}$$

where step (A61) holds by triangular inequality; step (A62) holds by noticing that

$$-\pi < \arg\left(\sum_{x\in\mathsf{supp}(P_{X^\star})} \alpha_x \exp(i\theta_x)\right) \leq \pi, \tag{A67}$$

where $\{\alpha_x\}$ and $\{\theta_x\}$ are real numbers that depend on x; (A63) follows from using the bound $(a+b)^2 \leq 2(a^2+b^2)$; (A64) holds because $x \mapsto \log^2(x)$ is a decreasing function for $x < 1$ and because $\mathbb{E}\left[\mathrm{e}^{-\frac{(z_R+n-X^\star)^2}{2\sigma_2^2}}\right] \geq \mathrm{e}^{-\frac{\mathbb{E}\left[(z_R+n-X^\star)^2\right]}{2\sigma_2^2}}$, which follows from Jensen's inequality; (A65) follows from $\mathbb{E}[X^\star] = 0$ and $\mathbb{E}[(X^\star)^2] \leq \mathsf{R}^2$; and (A66) follows from the bound $|a+b|^k \leq 2^{k-1}(|a|^k+|b|^k)$ for $k \geq 1$. Furthermore, given that $|z_R| \leq \mathsf{B}$ and $|z_I| \leq \mathsf{B}$, we arrive at the bound

$$\left((z_R+n)^2+\mathsf{R}^2\right)^2 \leq 2\left(8(\mathsf{B}^4+n^4)+\mathsf{R}^4\right). \tag{A68}$$

Consequently,

$$\frac{\mathbb{E}\left[|N|\cdot|\log f_{Y_2}(z+N)|\right]}{\sqrt{\sigma_2^2-\sigma_1^2}} \leq \frac{\sqrt{\mathbb{E}\left[|N|^2\right]\mathbb{E}\left[|\log f_{Y_2}(z+N)|^2\right]}}{\sqrt{\sigma_2^2-\sigma_1^2}} \tag{A69}$$

$$\leq \sqrt{\frac{2\mathsf{B}^2}{\sigma_2^2} + |\log(2\pi\sigma_2^2)|^2 + \frac{8(\mathsf{B}^4+\mathbb{E}[N^4])+\mathsf{R}^4}{\sigma_2^4} + \pi^2} \tag{A70}$$

$$= \sqrt{\frac{2\mathsf{B}^2}{\sigma_2^2} + |\log(2\pi\sigma_2^2)|^2 + \frac{8\mathsf{B}^4+24(\sigma_2^2-\sigma_1^2)^2+\mathsf{R}^4}{\sigma_2^4} + \pi^2}, \tag{A71}$$

where (A69) follows from Cauchy–Schwarz inequality; (A70) follows from $\mathbb{E}[N^4] = 3(\sigma_2^2 - \sigma_1^2)^2$. Moreover, we have

$$|f_{Y_1}(z)| \leq \mathbb{E}[|\phi_{\sigma_1}(z - X^\star)|] \tag{A72}$$

$$= \frac{1}{\sqrt{2\pi\sigma_1^2}} \mathbb{E}\left[\exp\left(-\frac{(z_R - X^\star)^2 - z_I^2}{2\sigma_1^2}\right)\right] \tag{A73}$$

$$\leq \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{\mathsf{B}^2}{2\sigma_1^2}\right), \tag{A74}$$

and finally

$$\mathbb{E}\left[|X^\star| \cdot |\phi_{\sigma_1}(z - X^\star)|\right] \leq \mathsf{R}\, \mathbb{E}\left[|\phi_{\sigma_1}(z - X^\star)|\right] \tag{A75}$$

$$\leq \mathsf{R} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{\mathsf{B}^2}{2\sigma_1^2}\right). \tag{A76}$$

Putting all contributions together, we get

$$|\check{h}(z)|\sqrt{2\pi\sigma_1^2} e^{-\frac{\mathsf{B}^2}{2\sigma_1^2}} \leq \frac{\sigma_1^2 \sqrt{\frac{2\mathsf{B}^2}{\sigma_2^2} + |\log(2\pi\sigma_2^2)|^2 + \frac{8\mathsf{B}^4 + 24(\sigma_2^2 - \sigma_1^2)^2 + \mathsf{R}^4}{\sigma_2^4} + \pi^2}}{\sqrt{\sigma_2^2 - \sigma_1^2}} + \mathsf{B} + \mathsf{R} \tag{A77}$$

$$\leq a_1 \mathsf{B}^2 + a_2 \mathsf{B} + a_3, \tag{A78}$$

where, in the last step, we have used that $\sqrt{\sum_i x_i} \leq \sum_i \sqrt{x_i}$ and the fact that $\mathsf{R} \leq \mathsf{B}$. □

Lemma A5. *Let $\check{h} : \mathbb{C} \to \mathbb{C}$ denote the complex extension of the function h in (123). Then, for*

$$\mathsf{B} \geq \mathsf{R} \frac{\sigma_2^2 + \sigma_1^2}{\sigma_2^2 - \sigma_1^2}, \tag{A79}$$

we have that

$$\max_{|z| \leq \mathsf{B}} |\check{h}(z)| \geq (c_1 \mathsf{B} - c_2 \mathsf{R}) \frac{\exp\left(-\frac{(\mathsf{B}+\mathsf{R})^2}{2\sigma_1^2}\right)}{\sqrt{2\pi\sigma_1^2}} > 0, \tag{A80}$$

where $c_1 = 1 - \frac{\sigma_1^2}{\sigma_2^2}$ and $c_2 = 1 + \frac{\sigma_1^2}{\sigma_2^2}$.

Proof. First, note that

$$\frac{\mathbb{E}_N[\mathbb{E}[X^\star | Y_2 = \mathsf{B} + N]]}{\sigma_2^2} - \frac{\mathbb{E}[X^\star | Y_1 = \mathsf{B}]}{\sigma_1^2} \geq -\frac{\mathsf{R}}{\sigma_2^2} - \frac{\mathsf{R}}{\sigma_1^2}. \tag{A81}$$

Second, note that the condition in (A79) implies that

$$0 \leq \mathsf{B}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) - \frac{\mathsf{R}}{\sigma_2^2} - \frac{\mathsf{R}}{\sigma_1^2}. \tag{A82}$$

Therefore, by using (111) together with (A81) and (A82), we arrive at

$$\max_{|z| \leq \mathsf{B}} |\check{h}(z)| \geq \left|\check{h}(\mathsf{B})\right| \tag{A83}$$

$$= \left| \frac{\mathbb{E}[\mathbb{E}[X^\star|Y_2 = \mathsf{B}+N]] - \mathsf{B}}{\sigma_2^2} - \frac{\mathbb{E}[X^\star|Y_1 = \mathsf{B}] - \mathsf{B}}{\sigma_1^2} \right| \sigma_1^2 f_{Y_1}(\mathsf{B}) \tag{A84}$$

$$\geq \left(\mathsf{B}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} \right) - \frac{\mathsf{R}}{\sigma_2^2} - \frac{\mathsf{R}}{\sigma_1^2} \right) \sigma_1^2 f_{Y_1}(\mathsf{B}) \tag{A85}$$

$$\geq \left(\mathsf{B}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} \right) - \frac{\mathsf{R}}{\sigma_2^2} - \frac{\mathsf{R}}{\sigma_1^2} \right) \frac{\sigma_1^2}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mathsf{B}+\mathsf{R})^2}{2\sigma_1^2} \right), \tag{A86}$$

where in last bound we have used Jensen's inequality to arrive at

$$f_{Y_1}(\mathsf{B}) = \mathbb{E}[\phi_{\sigma_1}(\mathsf{B} - X^\star)] \tag{A87}$$

$$= \frac{1}{\sqrt{2\pi\sigma_1^2}} \mathbb{E}\left[\exp\left(-\frac{(\mathsf{B}-X^\star)^2}{2\sigma_1^2} \right) \right] \tag{A88}$$

$$\geq \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mathsf{B}+\mathsf{R})^2}{2\sigma_1^2} \right). \tag{A89}$$

This concludes the proof. □

Appendix C. Proof of Theorem 7

To study the large n behavior, we need the following bounds on the function h_ν [39,40]: for $\nu > \frac{1}{2}$

$$\mathsf{h}_\nu(x) = \frac{x}{\frac{2\nu-1}{2} + \sqrt{\frac{(2\nu-1)^2}{4} + x^2}} \cdot g_\nu(x), \tag{A90}$$

where

$$1 \geq g_\nu(x) \geq \frac{\frac{2\nu-1}{2} + \sqrt{\frac{(2\nu-1)^2}{4} + x^2}}{\nu + \sqrt{\nu^2 + x^2}}. \tag{A91}$$

Moreover, let

$$U_n = \|\mathsf{R} + \sqrt{s}\mathbf{Z}\| \tag{A92}$$

with $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}_n, \sigma^2 \mathbf{I}_n)$. Consequently,

$$\lim_{n\to\infty} \mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2 \left(\frac{\|\mathsf{R} + \sqrt{s}\mathbf{Z}\|\mathsf{R}}{s} \right) \right]$$

$$= \mathbb{E}\left[\lim_{n\to\infty} \mathsf{h}_{\frac{n}{2}}^2 \left(\frac{\|\mathsf{R} + \sqrt{s}\mathbf{Z}\|\mathsf{R}}{s} \right) \right] \tag{A93}$$

$$= \mathbb{E}\left[\lim_{n\to\infty} \frac{U_n^2 \frac{\mathsf{R}^2}{s^2}}{\left(\frac{n-1}{2} + \sqrt{\frac{(n-1)^2}{4} + U_n^2 \frac{\mathsf{R}^2}{s^2}} \right)^2} \cdot g_{\frac{n}{2}}^2 \left(U_n \frac{\mathsf{R}}{s} \right) \right] \tag{A94}$$

$$= \mathbb{E}\left[\lim_{n\to\infty} \frac{\frac{1}{n} U_n^2 \frac{\mathsf{R}^2}{s^2}}{n \cdot \left(\frac{1}{2} + \sqrt{\frac{1}{4} + \left(\frac{1}{n} U_n \frac{\mathsf{R}}{s} \right)^2} \right)^2} \cdot g_{\frac{n}{2}}^2 \left(U_n \frac{\mathsf{R}}{s} \right) \right] \tag{A95}$$

$$= 0, \tag{A96}$$

where (A93) follows from the dominated convergence theorem, since $|h_\nu| \leq 1$; (A94) follows from using (A90); (A96) follows from using the strong law of large numbers to note that

$$\lim_{n\to\infty} \frac{1}{n} U_n^2 = \lim_{n\to\infty} \frac{\|\mathsf{R} + \sqrt{s}\mathbf{Z}\|^2}{n} = s. \tag{A97}$$

Now, combining the capacity expression in (58) and (A96), we have that

$$\lim_{n\to\infty} C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}, n) = \frac{1}{2}\int_{\sigma_1^2}^{\sigma_2^2} \frac{\mathsf{R}^2}{s^2} \mathrm{d}s = \mathsf{R}^2 \left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_2^2} \right). \tag{A98}$$

Appendix D. Proof of Theorem 8

Let $\mathsf{R}_n = c\sqrt{n}$

$$\lim_{n\to\infty} \frac{C_s(\sigma_1^2, \sigma_2^2, \mathsf{R}_n, n)}{n} = \frac{c^2}{2} \int_{\sigma_1^2}^{\sigma_2^2} \frac{1 - \lim_{n\to\infty} \mathbb{E}\left[\mathsf{h}_{\frac{n}{2}}^2 \left(\frac{\|\mathsf{R}_n + \sqrt{s}\mathbf{Z}\| \mathsf{R}_n}{s} \right) \right]}{s^2} \mathrm{d}s \tag{A99}$$

$$= \frac{c^2}{2} \int_{\sigma_1^2}^{\sigma_2^2} \frac{1 - \frac{c^2(c^2+s)}{\left(\frac{s}{2} + \sqrt{\frac{s^2}{4} + c^2(c^2+s)} \right)^2}}{s^2} \mathrm{d}s \tag{A100}$$

$$= \frac{1}{2} \log \left(\frac{\sigma_2^2(c^2 + \sigma_1^2)}{\sigma_1^2(c^2 + \sigma_2^2)} \right), \tag{A101}$$

where (A100) follows from the limit established in (98). This concludes the proof.

Appendix E. Partial Derivatives for the Gradient Ascent Algorithm

The partial derivatives of the secrecy information, with respect to any mass point $\rho_l \in \mathsf{supp}(P_{\|\mathbf{X}\|})$, are defined as

$$\frac{\partial}{\partial \rho_l} I_s(\|\mathbf{X}\|; P_{\|\mathbf{X}\|}) = \sum_{k=1}^{K} p_i \cdot \frac{\partial}{\partial \rho_l} \widetilde{\Xi}\left(\rho_k \, ; \hat{P}_{\|\mathbf{X}\|} \right), \quad l = 1, \ldots, K. \tag{A102}$$

By (A9), we have that $\widetilde{\Xi}\left(\|\mathbf{x}\| \, ; \hat{P}_{\|\mathbf{X}\|} \right) = i_1(\|\mathbf{x}\|; P_{\mathbf{X}}) - i_2(\|\mathbf{x}\|; P_{\mathbf{X}})$, where $i_j(\|\mathbf{x}\|; P_{\mathbf{X}})$, for $j = 1, 2$, is defined in (A10). Therefore, to compute (A102), we define the following derivatives

$$\frac{\partial}{\partial \rho_l} i_j(\rho_k; P_{\mathbf{X}}) = \int_0^\infty \frac{\partial}{\partial \rho_l} \left(f_{\chi_n^2(\rho_k^2/\sigma_j^2)}(y) \log \frac{y^{\frac{n}{2}-1}}{\sum_{m=1}^{K} p_m f_{\chi_n^2(\rho_m^2/\sigma_j^2)}(y)} \right) dy, \tag{A103}$$

where $f_{\chi_n^2(\rho_k^2/\sigma_j^2)}(y)$ is the noncentral chi-square pdf with noncentrality parameter ρ_k^2/σ_j^2 and n degrees of freedom. Notice that the derivative of $f_{\chi_n^2(\rho_k^2/\sigma_j^2)}(y)$ with respect to ρ_l is different from zero only when $k = l$ and is given by

$$\frac{\partial}{\partial \rho_l} f_{\chi_n^2(\rho_l^2/\sigma_j^2)}(y) = \frac{\rho_l}{\sigma_j^2} \left(f_{\chi_{n+2}^2(\rho_l^2/\sigma_j^2)}(y) - f_{\chi_n^2(\rho_l^2/\sigma_j^2)}(y) \right). \tag{A104}$$

Moreover, given the probability p_l associated with ρ_l, we have that

$$\frac{\partial}{\partial \rho_l} \log \frac{y^{\frac{n}{2}-1}}{\sum_{k=1}^{K} p_k f_{\chi_n^2(\rho_k^2/\sigma_j^2)}(y)} = -p_l \frac{\frac{\partial}{\partial \rho_l} f_{\chi_n^2(\rho_l^2/\sigma_j^2)}(y)}{\sum_{k=1}^{K} p_k f_{\chi_n^2(\rho_k^2/\sigma_j^2)}(y)}. \tag{A105}$$

Finally, by combining everything together, we find

$$\frac{\partial}{\partial \rho_l} I_s(\|\mathbf{X}\|; P_{\|\mathbf{X}\|}) =$$
$$p_l \int_0^\infty \frac{\rho_l}{\sigma_1^2} \left(f_{\chi^2_{n+2}(\rho_l^2/\sigma_1^2)}(y) - f_{\chi^2_n(\rho_l^2/\sigma_1^2)}(y) \right) \left[\log \frac{y^{\frac{n}{2}-1}}{\sum_{k=1}^K p_k f_{\chi^2_n(\rho_k^2/\sigma_1^2)}(y)} - 1 \right] dy$$
$$- p_l \int_0^\infty \frac{\rho_l}{\sigma_2^2} \left(f_{\chi^2_{n+2}(\rho_l^2/\sigma_2^2)}(y) - f_{\chi^2_n(\rho_l^2/\sigma_2^2)}(y) \right) \left[\log \frac{y^{\frac{n}{2}-1}}{\sum_{k=1}^K p_k f_{\chi^2_n(\rho_k^2/\sigma_2^2)}(y)} - 1 \right] dy. \quad \text{(A106)}$$

References

1. Favano, A.; Barletta, L.; Dytso, A. Simulated Data. Available online: https://github.com/ucando83/WiretapCapacity (accessed on 26 April 2023).
2. Wyner, A.D. The wire-tap channel. *Bell Syst. Tech. J.* **1975**, *54*, 1355–1387. [CrossRef]
3. Leung-Yan-Cheong, S.; Hellman, M. The Gaussian wire-tap channel. *IEEE Trans. Inf. Theory* **1978**, *24*, 451–456. [CrossRef]
4. Bloch, M.; Barros, J. *Physical-Layer Security: From Information Theory to Security Engineering*; Cambridge University Press: Cambridge, UK, 2011.
5. Oggier, F.; Hassibi, B. A Perspective on the MIMO Wiretap Channel. *Proc. IEEE* **2015**, *103*, 1874–1882. [CrossRef]
6. Liang, Y.; Poor, H.V.; Shamai (Shitz), S. Information theoretic security. *Found. Trends Commun. Inf. Theory* **2009**, *5*, 355–580. [CrossRef]
7. Poor, H.V.; Schaefer, R.F. Wireless physical layer security. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 19–26. [CrossRef] [PubMed]
8. Mukherjee, A.; Fakoorian, S.A.A.; Huang, J.; Swindlehurst, A.L. Principles of physical layer security in multiuser wireless networks: A survey. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 1550–1573. [CrossRef]
9. Gopala, P.K.; Lai, L.; El Gamal, H. On the secrecy capacity of fading channels. *IEEE Trans. Inf. Theory* **2008**, *54*, 4687–4698. [CrossRef]
10. Bloch, M.; Barros, J.; Rodrigues, M.R.; McLaughlin, S.W. Wireless information-theoretic security. *IEEE Trans. Inf. Theory* **2008**, *54*, 2515–2534. [CrossRef]
11. Khisti, A.; Tchamkerten, A.; Wornell, G.W. Secure broadcasting over fading channels. *IEEE Trans. Inf. Theory* **2008**, *54*, 2453–2469. [CrossRef]
12. Liang, Y.; Poor, H.V.; Shamai, S. Secure communication over fading channels. *IEEE Trans. Inf. Theory* **2008**, *54*, 2470–2492. [CrossRef]
13. Shafiee, S.; Liu, N.; Ulukus, S. Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel. *IEEE Trans. Inf. Theory* **2009**, *55*, 4033–4039. [CrossRef]
14. Khisti, A.; Wornell, G.W. Secure transmission with multiple antennas–Part II: The MIMOME wiretap channel. *IEEE Trans. Inf. Theory* **2010**, *56*, 5515–5532. [CrossRef]
15. Oggier, F.; Hassibi, B. The secrecy capacity of the MIMO wiretap channel. *IEEE Trans. Inf. Theory* **2011**, *57*, 4961–4972. [CrossRef]
16. Guo, D.; Shamai, S.; Verdú, S. Mutual information and minimum mean-square error in Gaussian channels. *IEEE Trans. Inf. Theory* **2005**, *51*, 1261–1282. [CrossRef]
17. Bustin, R.; Liu, R.; Poor, H.V.; Shamai, S. An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel. *Eurasip J. Wirel. Commun. Netw.* **2009**, *2009*, 1–8. [CrossRef]
18. Liu, T.; Shamai, S. A note on the secrecy capacity of the multiple-antenna wiretap channel. *IEEE Trans. Inf. Theory* **2009**, *55*, 2547–2553. [CrossRef]
19. Loyka, S.; Charalambous, C.D. An algorithm for global maximization of secrecy rates in Gaussian MIMO wiretap channels. *IEEE Trans. Commun.* **2015**, *63*, 2288–2299. [CrossRef]
20. Loyka, S.; Charalambous, C.D. Optimal signaling for secure communications over Gaussian MIMO wiretap channels. *IEEE Trans. Inf. Theory* **2016**, *62*, 7207–7215. [CrossRef]
21. Ozel, O.; Ekrem, E.; Ulukus, S. Gaussian wiretap channel with amplitude and variance constraints. *IEEE Trans. Inf. Theory* **2015**, *61*, 5553–5563. [CrossRef]
22. Soltani, M.; Rezki, Z. Optical wiretap channel with input-dependent Gaussian noise under peak-and average-intensity constraints. *IEEE Trans. Inf. Theory* **2018**, *64*, 6878–6893. [CrossRef]
23. Soltani, M.; Rezki, Z. The Degraded Discrete-Time Poisson Wiretap Channel. *arXiv* **2021**, arXiv:2101.03650.
24. Nam, S.H.; Lee, S.H. Secrecy Capacity of a Gaussian Wiretap Channel with One-bit ADCs is Always Positive. In Proceedings of the 2019 IEEE Information Theory Workshop (ITW), Visby, Sweden, 25–28 August 2019; pp. 1–5. [CrossRef]
25. Dytso, A.; Egan, M.; Perlaza, S.M.; Poor, H.V.; Shitz, S.S. Optimal Inputs for Some Classes of Degraded Wiretap Channels. In Proceedings of the 2018 IEEE Information Theory Workshop (ITW), Guangzhou, China, 25–29 November 2018; pp. 1–5. [CrossRef]
26. Karlin, S. Pólya type distributions, II. *Ann. Math. Stat.* **1957**, *28*, 281–308. [CrossRef]

27. Dytso, A.; Al, M.; Poor, H.V.; Shamai Shitz, S. On the Capacity of the Peak Power Constrained Vector Gaussian Channel: An Estimation Theoretic Perspective. *IEEE Trans. Inf. Theory* **2019**, *65*, 3907–3921. [CrossRef]
28. Favano, A.; Ferrari, M.; Magarini, M.; Barletta, L. The Capacity of the Amplitude-Constrained Vector Gaussian Channel. In Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT), Melbourne, Australia, 12–20 July 2021; pp. 426–431. [CrossRef]
29. Berry, J.C. Minimax estimation of a bounded normal mean vector. *J. Multivar. Anal.* **1990**, *35*, 130–139. [CrossRef]
30. Dytso, A.; Yagli, S.; Poor, H.V.; Shamai (Shitz), S. The Capacity Achieving Distribution for the Amplitude Constrained Additive Gaussian Channel: An Upper Bound on the Number of Mass Points. *IEEE Trans. Inf. Theory* **2020**, *66*, 2006–2022. [CrossRef]
31. Cover, T.; Thomas, J. *Elements of Information Theory*, 2nd, ed.; Wiley: Hoboken, NJ, USA, 2006.
32. Han, T.S.; Endo, H.; Sasaki, M. Reliability and Secrecy Functions of the Wiretap Channel Under Cost Constraint. *IEEE Trans. Inf. Theory* **2014**, *60*, 6819–6843. [CrossRef]
33. Barletta, L.; Dytso, A. Amplitude-Constrained Gaussian Wiretap Channel: Computation of the Optimal Input Distribution. In Proceedings of the 2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Athens, Greece, 5–8 September 2022; pp. 106–111. [CrossRef]
34. Rose, K. A mapping approach to rate-distortion computation and analysis. *IEEE Trans. Inf. Theory* **1994**, *40*, 1939–1952. [CrossRef]
35. Blahut, R. Computation of channel capacity and rate-distortion functions. *IEEE Trans. Inf. Theory* **1972**, *18*, 460–473. [CrossRef]
36. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
37. Yasui, K.; Suko, T.; Matsushima, T. An algorithm for computing the secrecy capacity of broadcast channels with confidential messages. In Proceedings of the 2007 IEEE International Symposium on Information Theory (ISIT), Nice, France, 24–29 June 2007; pp. 936–940. [CrossRef]
38. Verdú, S. Mismatched estimation and relative entropy. *IEEE Trans. Inf. Theory* **2010**, *56*, 3712–3720. [CrossRef]
39. Segura, J. Bounds for ratios of modified Bessel functions and associated Turán-type inequalities. *J. Math. Anal. Appl.* **2011**, *374*, 516–528. [CrossRef]
40. Baricz, Á. Bounds for Turánians of modified Bessel functions. *Expo. Math.* **2015**, *33*, 223–251. [CrossRef]
41. Tijdeman, R. On the number of zeros of general exponential polynomials. In *Proceedings of the Indagationes Mathematicae*; North-Holland: Amsterdam, The Netherlands, 1971; Volume 74, pp. 1–7.
42. Esposito, R. On a relation between detection and estimation in decision theory. *Inf. Control* **1968**, *12*, 116–120. [CrossRef]
43. Dytso, A.; Poor, H.V.; Shitz, S.S. A general derivative identity for the conditional mean estimator in Gaussian noise and some applications. In Proceedings of the 2020 IEEE International Symposium on Information Theory (ISIT), Los Angeles, CA, USA, 21–26 June 2020; pp. 1183–1188. [CrossRef]
44. Barletta, L.; Dytso, A. Scalar Gaussian Wiretap Channel: Bounds on the Support Size of the Secrecy-Capacity-Achieving Distribution. In Proceedings of the 2021 IEEE Information Theory Workshop (ITW), Kanazawa, Japan, 17–21 October 2021; pp. 1–6. [CrossRef]
45. Favano, A.; Barletta, L.; Dytso, A. On the Capacity Achieving Input of Amplitude Constrained Vector Gaussian Wiretap Channel. In Proceedings of the 2022 IEEE International Symposium on Information Theory (ISIT), Espoo, Finland, 26 June–1 July 2022; pp. 850–855. [CrossRef]
46. Favano, A. The Capacity of Amplitude-Constrained Vector Gaussian Channels. Ph.D. Dissertation, Politecnico di Milano, Milan, Italy, 2022.
47. Lapidoth, A.; Moser, S.M. Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels. *IEEE Trans. Inf. Theory* **2003**, *49*, 2426–2467. [CrossRef]