Scalar Gaussian Wiretap Channel: Bounds on the Support Size of the Secrecy-Capacity-Achieving Distribution

Luca Barletta*, Alex Dytso**
* Politecnico di Milano, Milano, 20133, Italy. Email: luca.barletta@polimi.it
** New Jersey Institute of Technology, Newark, NJ 07102, USA. Email: alex.dytso@njit.edu

***Abstract*—This work studies the secrecy-capacity of a scalar-Gaussian wiretap channel with an amplitude constraint on the input. It is known that for this channel, the secrecy-capacity-achieving distribution is discrete with finitely many points. This work improves such result by showing an upper bound of the order $\frac{\mathsf{A}}{\sigma_1^2}$ where A is the amplitude constraint and σ_1^2 is the variance of the Gaussian noise over the legitimate channel.**

I. Introduction

Consider the Gaussian wiretap channel with outputs

$$Y_1 = X + N_1, \tag{1}$$
$$Y_2 = X + N_2, \tag{2}$$

where $N_1 \sim \mathcal{N}(0, \sigma_1^2)$ and $N_2 \sim \mathcal{N}(0, \sigma_2^2)$, and with (X, N_1, N_2) independent of each other. The output Y_1 is observed by the legitimate receiver whereas the output Y_2 is observed by the malicious receiver. In this work, we assume that the input X is limited by a peak-power constraint or amplitude constraint given by $|X| \leq \mathsf{A}$. For this setting, the secrecy-capacity is given by

$$C_s(\sigma_1, \sigma_2, \mathsf{A}) = \max_{P_X : |X| \leq \mathsf{A}} I(X; Y_1) - I(X; Y_2) \tag{3}$$
$$= \max_{P_X : |X| \leq \mathsf{A}} I(X; Y_1 | Y_2). \tag{4}$$

We are interested in studying the input distribution $P_{X^\star}$ that maximizes (4). It can be shown that for $\sigma_1^2 \geq \sigma_2^2$ the secrecy-capacity is equal to zero. Therefore, in the remaining, we assume that $\sigma_1^2 < \sigma_2^2$.

Literature Review

The wiretap channel was introduced by Wyner in [1], who also established the secrecy-capacity of the degraded wiretap channel. The wiretap channel plays a central role in network information theory; the interested reader is referred to [2]–[5] and reference therein for an in-detail treatment of the topic.

The secrecy-capacity of a Gaussian wiretap channel with an average power constraint was shown by Leung and Hellman in [6] where the secrecy-capacity-achieving input distribution was shown to be Gaussian. The secrecy-capacity of Gaussian wiretap channel with an amplitude and power constraint was considered by Ozel et al. in [7] where the authors showed that the secrecy-capacity-achieving input distribution is discrete with finitely many points. The work of [7] was extended to noise-dependent channels by Soltani and Rezki in [8]. For further studies of the properties of secrecy-capacity-achieving input distribution for a class of degraded wiretap channels, the interested reader is referred to [9]–[11].

The classical approach for demonstrating that the secrecy-capacity-achieving distributions are discrete relies on an analytic argument introduced to information theory by Smith in [12]. The drawback of this technique is that it does not provide any bounds on the support size of the secrecy-capacity-achieving distribution and only asserts that the support is countable. In this work, instead of following the approach of [12], we follow the approach introduced in [13], which relies on the *variation diminishing* property [14].

This work has two goals. The first goal is to sharpen the results of [7] by establishing a firm upper bound on the number of points in the support of the secrecy-capacity-achieving distribution. The second goal is to study the necessary techniques required to extend the method introduced in [13] to network information theory problems. The wiretap channel serves as an ideal first test candidate in this research program.

Outline and Contributions

Section II presents our main results, which includes two new upper bounds on the cardinality of the support of the optimal input distribution. Section III is dedicated to the proofs. Section IV concludes the paper with a discussion on interesting future directions. Some of the proofs can be found in the extended version of this paper [15].

Notation

Throughout the paper, the deterministic scalar quantities are denoted by lower-case letters and random variables are denoted by uppercase letters.

We denote the distribution of a random variable X by P_X. The support set of P_X is denoted and defined as

$$\mathsf{supp}(P_X) = \{x : \text{ for every open set } \mathcal{D} \ni x \text{ we have that } P_X(\mathcal{D}) > 0\}. \tag{5}$$

The relative entropy between distributions P and Q will be denoted by $\mathsf{D}(P\|Q)$. The pdf of a Gaussian random variable with zero mean and variance σ^2 is denoted by $\phi_\sigma(\cdot)$

Finally, the number of zeros 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 $\mathcal{D}$.

II. Main Result

In this section, we state our main results. We first present the following ancillary lemma the first part of which was shown in [7].

Lemma 1. *$P_{X^\star}$ maximizes* (4) *if and only if*

$$\Xi(x) = C_s(\sigma_1, \sigma_2, \mathsf{A}),\ x \in \mathsf{supp}(P_{X^\star}), \tag{6}$$

$$\Xi(x) \le C_s(\sigma_1, \sigma_2, \mathsf{A}),\ x \in [-\mathsf{A}, \mathsf{A}], \tag{7}$$

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

$$\Xi(x) = \mathsf{D}(f_{Y_1|X}(\cdot|x)\|f_{Y_1^\star}) - \mathsf{D}(f_{Y_2|X}(\cdot|x)\|f_{Y_2^\star}) \tag{8}$$

$$= \mathbb{E}\left[g(Y_1)|X = x\right] + \log\left(\frac{\sigma_2}{\sigma_1}\right), \tag{9}$$

and where

$$g(y) = \mathbb{E}\left[\log \frac{f_{Y_2^\star}(y+N)}{f_{Y_1^\star}(y)}\right],\ y \in \mathbb{R}, \tag{10}$$

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

Proof: The first part of Lemma 1 was shown in [7]. The proof of (9) goes as follows:

$$\mathsf{D}(f_{Y_1|X}(\cdot|x)\|f_{Y_1^\star}) - \mathsf{D}(f_{Y_2|X}(\cdot|x)\|f_{Y_2^\star}) - \log\left(\frac{\sigma_2}{\sigma_1}\right) \tag{11}$$

$$= \int_{-\infty}^{\infty} \log \frac{1}{f_{Y_1^\star}(y)} \phi_{\sigma_1}(y-x)\mathrm{d}y - \int_{-\infty}^{\infty} \log \frac{1}{f_{Y_2^\star}(y)} \mathbb{E}[\phi_{\sigma_1}(y-x-N)]\mathrm{d}y \tag{12}$$

$$= \int_{-\infty}^{\infty} \log \frac{1}{f_{Y_1^\star}(y)} \phi_{\sigma_1}(y-x)\mathrm{d}y - \int_{-\infty}^{\infty} \mathbb{E}\left[\log \frac{1}{f_{Y_2^\star}(y+N)}\right] \phi_{\sigma_1}(y-x)\mathrm{d}y \tag{13}$$

$$= \int_{-\infty}^{\infty} \mathbb{E}\left[\log \frac{f_{Y_2^\star}(y+N)}{f_{Y_1^\star}(y)}\right] \phi_{\sigma_1}(y-x)\mathrm{d}y \tag{14}$$

$$= \int_{-\infty}^{\infty} g(y)\phi_{\sigma_1}(y-x)\mathrm{d}y, \tag{15}$$

where in (12) we have introduced $N \sim \mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$; and in (13) we applied the change of variable $y \mapsto y + N$. This concludes the proof. ∎

The main result of this paper is summarized in the following theorem.

Theorem 1. *For $\mathsf{A} > 0$*

$$|\mathsf{supp}(P_{X^\star})| \le \mathrm{N}\left([-R, R], g(\cdot) + \kappa_1\right) < \infty \tag{16}$$

where

$$\kappa_1 = \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s, \tag{17}$$

$$R = \mathsf{A}\frac{\sigma_2 + \sigma_1}{\sigma_2 - \sigma_1} + \sqrt{\frac{\frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2} + 2C_s}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}}}. \tag{18}$$

Moreover,

$$\mathrm{N}\left([-R, R], g(\cdot) + \kappa_1\right) \le \rho\frac{\mathsf{A}^2}{\sigma_1^2} + O(\log(\mathsf{A})), \tag{19}$$

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

III. Proofs of the Main Results

A. Proof of the bound in (16)

The function $g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s$ will play an important role in our proof in this section. We start with the following lemma, which characterizes the region on which the zeros of the function $g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s$ concentrate.

Lemma 2. *Let*

$$\overline{C}_s = \frac{1}{2}\log\left(\frac{1 + \frac{\mathsf{A}^2}{\sigma_1^2}}{1 + \frac{\mathsf{A}^2}{\sigma_2^2}}\right). \tag{20}$$

Then,

$$C_s \le \overline{C}_s. \tag{21}$$

Moreover, there exists some $R = R(\sigma_1, \sigma_2, \mathsf{A}) < \infty$ such that

$$\mathrm{N}\left(\mathbb{R}, g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) = \mathrm{N}\left([-R, R], g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) < \infty. \tag{22}$$

Furthermore, R can be upper-bounded as follows:

$$R \le \mathsf{A}d_1 + d_2 \tag{23}$$

where

$$d_1 = \frac{\sigma_2 + \sigma_1}{\sigma_2 - \sigma_1}, \tag{24}$$

$$d_2 = \sqrt{\frac{\frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2} + 2C_s}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}}} \le \sqrt{\frac{\frac{\sigma_2^2 - \sigma_1^2}{\sigma_2^2} + 2\overline{C}_s}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}}}. \tag{25}$$

Proof. First, note that

$$C_s = \max_{P_X\colon |X| \le \mathsf{A}} I(X; Y_1|Y_2) \tag{26}$$

$$= \max_{P_X\colon |X| \le \mathsf{A},\, \mathbb{E}[X^2] \le \mathsf{A}^2} I(X; Y_1|Y_2) \tag{27}$$

$$\le \max_{P_X\colon \mathbb{E}[X^2] \le \mathsf{A}^2} I(X; Y_1|Y_2). \tag{28}$$

The last expression is the secrecy-capacity of a Gaussian wiretap channel with an average power constraint, which is given in (20).

Second, for $|y| \ge \mathsf{A}$, we can lower-bound the function g as follows:

$$g(y) = \mathbb{E}\left[\log f_{Y_2^\star}(y+N)\right] - \log f_{Y_1^\star}(y) \tag{29}$$

$$= \mathbb{E}\left[\log \mathbb{E}[\phi_{\sigma_2}(y + N - X^\star)|N]\right] - \log \mathbb{E}[\phi_{\sigma_1}(y - X^\star)] \tag{30}$$

$$\geq \mathbb{E}\left[\log \phi_{\sigma_2}(y+N-X^\star)\right] - \log \mathbb{E}[\phi_{\sigma_1}(y-X^\star)] \tag{31}$$

$$\geq \log\frac{\sigma_1}{\sigma_2} - \mathbb{E}\left[\frac{(y+N-X^\star)^2}{2\sigma_2^2}\right] + \frac{(|y|-\mathsf{A})^2}{2\sigma_1^2} \tag{32}$$

$$= \log\frac{\sigma_1}{\sigma_2} - \mathbb{E}\left[\frac{(y-X^\star)^2}{2\sigma_2^2}\right] - \frac{\sigma_2^2-\sigma_1^2}{2\sigma_2^2} + \frac{(|y|-\mathsf{A})^2}{2\sigma_1^2} \tag{33}$$

$$\geq \log\frac{\sigma_1}{\sigma_2} - \frac{(|y|+\mathsf{A})^2}{2\sigma_2^2} - \frac{\sigma_2^2-\sigma_1^2}{2\sigma_2^2} + \frac{(|y|-\mathsf{A})^2}{2\sigma_1^2}, \tag{34}$$

where (31) follows by applying Jensen's inequality to the first term; (32) follows by

$$\mathbb{E}[\phi_{\sigma_1}(y-X^\star)] \leq \phi_{\sigma_1}(|y|-\mathsf{A}), \qquad |y| \geq \mathsf{A}; \tag{35}$$

and (34) follows by $(y-X^\star)^2 \leq (|y|+\mathsf{A})^2$ for all $|y| \geq \mathsf{A} \geq |X^\star|$. The function

$$g(y) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s$$
$$\geq -\frac{(|y|+\mathsf{A})^2}{2\sigma_2^2} - \frac{\sigma_2^2-\sigma_1^2}{2\sigma_2^2} + \frac{(|y|-\mathsf{A})^2}{2\sigma_1^2} - C_s \tag{36}$$

is strictly positive when

$$|y| > \frac{\mathsf{A}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right) + \sqrt{\frac{4\mathsf{A}^2}{\sigma_1^2\sigma_2^2} + \left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}\right)\left(\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}+2C_s\right)}}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}}. \tag{37}$$

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

$$|y| \geq \mathsf{A}\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1} + \sqrt{\frac{\frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2}+2C_s}{\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}}}. \tag{38}$$

This concludes the proof for the bound on R. □

To show the bound on the number of points, we need to first present a number of ancillary results. We start with the following definition.

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

$$\mathscr{S}(\xi) = \sup_{m\in\mathbb{N}}\left\{\sup_{y_1<\cdots<y_m\subseteq\Omega} \mathscr{N}\{\xi(y_i)\}_{i=1}^m\right\}, \tag{39}$$

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

The following theorem, shown in [14], will be a key step in the proof of the upper bound on the number of mass points.

Theorem 2 (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.*[1] *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*

$$\Xi(x) = \int \xi(y)p(x,y)\mathrm{d}\mu(y). \tag{40}$$

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 $\Xi(x)$, which is the output of integral transformation, is less than the number of sign changes of the function $\xi(y)$, which is the input to the integral transformation. The fact that the Gaussian pdf is a strictly totally positive kernel was shown in [14].

We are now in the position to show the upper bound in (16):

$$|\mathsf{supp}(P_{X^\star})|$$
$$\leq \mathrm{N}\left([-\mathsf{A},\mathsf{A}], \Xi(x) - C_s(\sigma_1,\sigma_2,\mathsf{A})\right) \tag{41}$$

$$= \mathrm{N}\left([-\mathsf{A},\mathsf{A}], \mathbb{E}\left[g(Y_1) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s \middle| X = x\right]\right) \tag{42}$$

$$\leq \mathscr{S}\left(g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \tag{43}$$

$$\leq \mathrm{N}\left(\mathbb{R}, g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \tag{44}$$

$$= \mathrm{N}\left([-R,R], g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \tag{45}$$

$$< \infty, \tag{46}$$

where (41) follows by using the following inclusion which is a consequence of Lemma 1

$$\mathsf{supp}(P_{X^\star}) \subseteq \{x \in [-\mathsf{A},\mathsf{A}] : \Xi(x) - C_s = 0\}; \tag{47}$$

(42) follows by using (9); (43) follows by applying Theorem 2 and the fact that Gaussian pdf is a strictly totally positive kernel; (45) is proved in Lemma 2; and (46) follows since $g(\cdot)$ is an analytic function in $(-R, R)$.

B. Counting the number of zeros: Proof of the bound in (19)

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 [16])**.** *Let R, s, t be positive numbers such that $s > 1$. For the complex valued function $f \neq 0$ which is analytic on $|z| < (st+s+t)R$, its number of zeros $\mathrm{N}(\mathcal{D}_R, f)$ within the disk $\mathcal{D}_R = \{z\colon |z| \leq R\}$ satisfies*

$$\mathrm{N}(\mathcal{D}_R, f)$$
$$\leq \frac{1}{\log s}\left(\log \max_{|z|\leq(st+s+t)R}|f(z)| - \log \max_{|z|\leq tR}|f(z)|\right). \tag{48}$$

Furthermore, the following loosened version of the bound in (16) will be useful.

Lemma 4.

$$|\mathsf{supp}(P_{X^\star})| \leq \mathrm{N}\left([-R,R], h(\cdot)\right) + 1 \tag{49}$$

[1] 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 $\det\left([f(x_i,y_j)]_{i,j=1}^m\right) > 0$ for all $1 \leq m \leq n$, and for all $x_1 < \cdots < x_m \in \mathbb{I}_1$, and $y_1 < \cdots < y_m \in \mathbb{I}_2$. If f is totally positive kernel of order n for all $n \in \mathbb{N}$, then f is a strictly totally positive kernel.

where

$$\frac{h(y)}{\sigma_1^2 f_{Y_1}(y)} = \frac{\mathbb{E}_N\left[\mathbb{E}[X^\star|Y_2 = y+N]\right] - y}{\sigma_2^2} - \frac{\mathbb{E}[X^\star|Y_1 = y] - y}{\sigma_1^2} \quad (50)$$

$$= \frac{\mathbb{E}\left[N \log f_{Y_2}(y+N)\right]}{\sigma_2^2 - \sigma_1^2} - \frac{\mathbb{E}[X^\star|Y_1 = y] - y}{\sigma_1^2}, \quad (51)$$

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

Proof: Starting from (45), we can write

$$|\mathsf{supp}(P_{X^\star})| \leq \mathrm{N}\left([-R, R], g(\cdot) + \log\left(\frac{\sigma_2}{\sigma_1}\right) - C_s\right) \quad (52)$$

$$\leq \mathrm{N}\left([-R, R], g'(\cdot)\right) + 1 \quad (53)$$

$$= \mathrm{N}\left([-R, R], \sigma_1^2 f_{Y_1}(\cdot) g'(\cdot)\right) + 1 \quad (54)$$

where in step (53) we have applied Rolle's theorem, and in step (54) 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:

$$g'(y) = \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_2}(y+N)\right] - \frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_1}(y) \quad (55)$$

$$= \frac{\mathbb{E}_N\left[\mathbb{E}[X^\star|Y_2 = y+N]\right] - y}{\sigma_2^2} - \frac{\mathbb{E}[X^\star|Y_1 = y] - y}{\sigma_1^2}, \quad (56)$$

where in the last step we have used the well-known Tweedy's formula (see for example [17], [18]):

$$\mathbb{E}[X^\star|Y_i = y] = y + \sigma_i^2 \frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_i}(y). \quad (57)$$

An alternative expression for the first term in the RHS of (55) is as follows:

$$\mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_2}(y+N)\right]$$

$$= \int_{-\infty}^{\infty} f_N(n) \frac{\mathrm{d}}{\mathrm{d}y} \log f_{Y_2}(y+n) \mathrm{d}n \quad (58)$$

$$= -\int_{-\infty}^{\infty} \left(\frac{\mathrm{d}}{\mathrm{d}n} f_N(n)\right) \cdot \log f_{Y_2}(y+n) \mathrm{d}n \quad (59)$$

$$= \int_{-\infty}^{\infty} \frac{n}{\sigma_2^2 - \sigma_1^2} f_N(n) \cdot \log f_{Y_2}(y+n) \mathrm{d}n \quad (60)$$

$$= \frac{1}{\sigma_2^2 - \sigma_1^2} \mathbb{E}\left[N \log f_{Y_2}(y+N)\right], \quad (61)$$

The proof is concluded by letting

$$h(y) \triangleq \sigma_1^2 f_{Y_1}(y) g'(y). \quad (62)$$

■

With the goal of getting an explicit bound on the number of zeros, through the application of Tijdeman's number of zeros Lemma, the following lemmas propose upper and lower bounds to the maximum module of the complex analytic extension of h over the disk $\mathcal{D}_R = \{z : |z| \leq R\}$.

Lemma 5. *Let* $\breve{h} : \mathbb{C} \to \mathbb{C}$ *denote the complex extension of the function* h *in* (62)*. Then, for* $\mathsf{B} \geq \mathsf{A}$*, we have that*

$$\max_{|z| \leq \mathsf{B}} |\breve{h}(z)| \leq \frac{1}{\sqrt{2\pi\sigma_1^2}} \mathrm{e}^{\frac{\mathsf{B}^2}{2\sigma_1^2}} \left(a_1 \mathsf{B}^2 + a_2 \mathsf{B} + a_3\right) \quad (63)$$

where

$$a_1 = \frac{3\sigma_1^2}{\sigma_2^2 \sqrt{\sigma_2^2 - \sigma_1^2}}, \quad (64)$$

$$a_2 = \frac{\sqrt{2}\sigma_1^2}{\sqrt{\sigma_2^2}\sqrt{\sigma_2^2 - \sigma_1^2}} + 2, \quad (65)$$

$$a_3 = \frac{\sigma_1^2}{\sqrt{\sigma_2^2 - \sigma_1^2}} \left(\sqrt{|\log(2\pi\sigma_2^2)|^2 + \frac{24(\sigma_2^2 - \sigma_1^2)^2}{\sigma_2^4} + \pi^2}\right). \quad (66)$$

Lemma 6. *Let* $\breve{h} : \mathbb{C} \to \mathbb{C}$ *denote the complex extension of the function* h *in* (62)*. Then, for*

$$\mathsf{B} \geq \mathsf{A} \frac{\sigma_2^2 + \sigma_1^2}{\sigma_2^2 - \sigma_1^2}, \quad (67)$$

we have that

$$\max_{|z| \leq \mathsf{B}} |\breve{h}(z)| \geq (c_1 \mathsf{B} - c_2 \mathsf{A}) \frac{\exp\left(-\frac{(\mathsf{B}+\mathsf{A})^2}{2\sigma_1^2}\right)}{\sqrt{2\pi\sigma_1^2}} > 0, \quad (68)$$

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

$$\frac{\mathbb{E}_N\left[\mathbb{E}[X^\star|Y_2 = \mathsf{B}+N]\right]}{\sigma_2^2} - \frac{\mathbb{E}[X^\star|Y_1 = \mathsf{B}]}{\sigma_1^2} \geq -\frac{\mathsf{A}}{\sigma_2^2} - \frac{\mathsf{A}}{\sigma_1^2}. \quad (69)$$

Second, note that the condition in (67) implies that

$$0 \leq \mathsf{B}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) - \frac{\mathsf{A}}{\sigma_2^2} - \frac{\mathsf{A}}{\sigma_1^2}. \quad (70)$$

Therefore, by using (50) together with (69) and (70), we arrive at

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

$$= \left|\frac{\mathbb{E}[\mathbb{E}[X^\star|Y_2 = \mathsf{B}+N]] - \mathsf{B}}{\sigma_2^2} - \frac{\mathbb{E}[X^\star|Y_1 = \mathsf{B}] - \mathsf{B}}{\sigma_1^2}\right| \sigma_1^2 f_{Y_1}(\mathsf{B}) \quad (71)$$

$$\geq \left(\mathsf{B}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) - \frac{\mathsf{A}}{\sigma_2^2} - \frac{\mathsf{A}}{\sigma_1^2}\right) \sigma_1^2 f_{Y_1}(\mathsf{B}) \quad (72)$$

$$\geq \left(\mathsf{B}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) - \frac{\mathsf{A}}{\sigma_2^2} - \frac{\mathsf{A}}{\sigma_1^2}\right) \frac{\sigma_1^2}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mathsf{B}+\mathsf{A})^2}{2\sigma_1^2}\right), \quad (73)$$

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

$$f_{Y_1}(\mathsf{B}) = \mathbb{E}\left[\phi_{\sigma_1}(\mathsf{B} - X^\star)\right] \quad (74)$$

$$= \frac{1}{\sqrt{2\pi\sigma_1^2}} \mathbb{E}\left[\exp\left(-\frac{(\mathsf{B} - X^\star)^2}{2\sigma_1^2}\right)\right] \quad (75)$$

$$\geq \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mathsf{B}+\mathsf{A})^2}{2\sigma_1^2}\right). \tag{76}$$

This concludes the proof. □

With Lemma 5 and Lemma 6 at our disposal we are now ready to use Tijdeman's Number of Zeros Lemma to provide an upper bound on the number of mass points:

$$\mathrm{N}\left([-R,R],h(\cdot)\right) \tag{77}$$

$$\leq \mathrm{N}\left(\mathcal{D}_R,\breve{h}(\cdot)\right) \tag{78}$$

$$\leq \min_{s>1,\,t>0}\left\{\frac{\log\frac{\max_{|z|\leq(st+s+t)R}|\breve{h}(z)|}{\max_{|z|\leq tR}|\breve{h}(z)|}}{\log s}\right\} \tag{79}$$

$$\leq \log\frac{\frac{\mathrm{e}^{\frac{(2\mathrm{e}+1)^2R^2}{2\sigma_1^2}}}{\sqrt{2\pi\sigma_1^2}}\left(a_1(2\mathrm{e}+1)^2R^2+a_2(2\mathrm{e}+1)R+a_3\right)}{(c_1R-c_2\mathsf{A})\frac{\exp\left(-\frac{(R+\mathsf{A})^2}{2\sigma_1^2}\right)}{\sqrt{2\pi\sigma_1^2}}} \tag{80}$$

$$= \frac{(2\mathrm{e}+1)^2R^2}{2\sigma_1^2}+\frac{(R+\mathsf{A})^2}{2\sigma_1^2} + \log\frac{a_1(2\mathrm{e}+1)^2R^2+a_2(2\mathrm{e}+1)R+a_3}{c_1R-c_2\mathsf{A}} \tag{81}$$

$$= \frac{(2\mathrm{e}+1)^2(d_1\mathsf{A}+d_2)^2}{2\sigma_1^2}+\frac{((d_1+1)\mathsf{A}+d_2)^2}{2\sigma_1^2} + \log\frac{a_1(2\mathrm{e}+1)^2(d_1\mathsf{A}+d_2)^2+a_2(2\mathrm{e}+1)(d_1\mathsf{A}+d_2)+a_3}{(c_1d_1-c_2)\mathsf{A}+c_1d_2} \tag{82}$$

$$\leq b_1\frac{\mathsf{A}^2}{\sigma_1^2}+b_2+\log\frac{b_3\mathsf{A}^2+b_4\mathsf{A}+b_5}{b_6\mathsf{A}+b_7} \tag{83}$$

$$\leq b_1\frac{\mathsf{A}^2}{\sigma_1^2}+O(\log(\mathsf{A})), \tag{84}$$

where (78) follows since extending to larger domain can only increase the number of zeros; (79) follows by the Tijdeman's Number of Zeros Lemma; (80) follows by choosing $s=\mathrm{e}$ and $t=1$ and using bounds in Lemma 5 and Lemma 6; (82) follows by using the value of R in (23); (83) using the bound $(a+b)^2\leq 2(a^2+b^2)$ and defining coefficients $b_i, i\in[1:7]$ which are combination of c_i's, a_i's and d_i's (the exact values of b_i's are given in the extended version in [15]); and (84) follows from the fact that the b_1, b_3, b_4 and b_6 coefficients do not depend A and the fact that the coefficients b_2, b_5 and b_4, while do depend on A through C_s, do not grow with A. The fact that C_s does not grow with A follows from the bound in (20).

IV. Conclusion

This works has focused on deriving upper bounds on the number of mass points of secrecy-capacity-achieving distribution.

The upper bounds in Theorem 1 are generalizations of the upper bounds on the number of points presented in [13] in the context of a point-to-point additive white Gaussian noise (AWGN) channel with an amplitude constraint. Indeed, if we let $\sigma_2\to\infty$, while keeping σ_1 and A fixed, then the wiretap channel reduces to the AWGN point-to-point channel.

An interesting future direction would be to find a matching implicit lower bound in (16). In [13] such a matching lower bound was found and shown to be tight with a multiplicative factor of two from the upper bound. These results effectively show that the oscillation theorem (see Theorem 2) is a strong enough tool for producing upper bounds on the cardinality of secrecy-capacity-achieving distributions for point-to-point channels. A matching lower bound in the case of the wiretap channel would demonstrate that oscillation theorem can also play an important role in network information theory problems. In [13], the key tool to finding the lower bound was the observation that a linear combination of $n+1$ distinct Gaussians with distinct variances can have at most $2n$ zeros. In the wiretap channel, due to a more complicated structure of the function g in (10), it is not immediately clear how such an argument can be applied.

It will also be interesting to augment an explicit upper bound on the number of points in (19) with a lower bound on the number of points. A possible line of attack consists of the following steps:

$$C_s(\sigma_1,\sigma_2,\mathsf{A}) = I(X^\star;Y_1)-I(X^\star;Y_2) \tag{85}$$

$$\leq H(X^\star)-I(X^\star;Y_2) \tag{86}$$

$$\leq \log(|\mathsf{supp}(P_{X^\star})|)-I(X^\star;Y_2), \tag{87}$$

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

$$C_s(\sigma_1,\sigma_2,\mathsf{A})\geq\frac{1}{2}\log\left(1+\frac{\frac{2\mathsf{A}^2}{\pi\mathrm{e}\sigma_1^2}}{1+\frac{\mathsf{A}^2}{\sigma_2^2}}\right). \tag{88}$$

Combing bounds in (87) and (88) we arrive at the following lower bound on the number of points:

$$|\mathsf{supp}(P_{X^\star})|\geq\sqrt{1+\frac{\frac{2\mathsf{A}^2}{\pi\mathrm{e}\sigma_1^2}}{1+\frac{\mathsf{A}^2}{\sigma_2^2}}\mathrm{e}^{I(X^\star;Y_2)}}. \tag{89}$$

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 A while the upper bound increases with A. A possible way of establishing a lower bound that is increasing in A is by showing that $I(X^\star;Y_2)\approx\frac{1}{2}\log\left(1+\frac{\mathsf{A}^2}{\sigma_2^2}\right)$. 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.

References

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] M. Bloch and J. Barros, *Physical-Layer Security:From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[3] F. Oggier and B. Hassibi, "A perspective on the MIMO wiretap channel," *Proc. of IEEE*, vol. 103, no. 10, pp. 1874–1882, 2015.

[4] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4–5, pp. 355–580, 2009.

[5] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," *Proc. the Natl. Acad. Sci. U.S.A.*, vol. 114, no. 1, pp. 19–26, 2017.

[6] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, 1978.

[7] O. Ozel, E. Ekrem, and S. Ulukus, "Gaussian wiretap channel with amplitude and variance constraints," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5553–5563, 2015.

[8] M. Soltani and Z. Rezki, "Optical wiretap channel with input-dependent Gaussian noise under peak-and average-intensity constraints," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6878–6893, 2018.

[9] A. Dytso, M. Egan, S. M. Perlaza, H. V. Poor, and S. S. Shitz, "Optimal inputs for some classes of degraded wiretap channels," in *Proc. IEEE Inf. Theory Workshop*. IEEE, 2018, pp. 1–5.

[10] M. Soltani and Z. Rezki, "The degraded discrete-time Poisson wiretap channel," *arXiv preprint arXiv:2101.03650*, 2021.

[11] S.-H. Nam and S.-H. Lee, "Secrecy capacity of a Gaussian wiretap channel with one-bit ADCs is always positive," in *Proc. IEEE Inf. Theory Workshop*. IEEE, 2019, pp. 1–5.

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

[13] A. Dytso, S. Yagli, H. V. Poor, and S. Shamai (Shitz), "The capacity achieving distribution for the amplitude constrained additive Gaussian channel: An upper bound on the number of mass points," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 2006–2022, 2020.

[14] S. Karlin, "Pólya type distributions, ii," *The Ann. Math. Stat.*, vol. 28, no. 2, pp. 281–308, 1957.

[15] L. Barletta and A. Dytso, "Scalar Gaussian wiretap channel: Properties of the support size of the secrecy-capacity-achieving distribution," 2021. [Online]. Available: https://arxiv.org/abs/2109.01566

[16] R. Tijdeman, "On the number of zeros of general exponential polynomials," in *Indagationes Mathematicae (Proceedings)*, vol. 74. North-Holland, 1971, pp. 1–7.

[17] R. Esposito, "On a relation between detection and estimation in decision theory," *Inf. Control*, vol. 12, no. 2, pp. 116–120, February 1968.

[18] A. Dytso, H. V. Poor, and S. Shamai (Shitz), "A general derivative identity for the conditional mean estimator in Gaussian noise and some applications," 2021. [Online]. Available: https://arxiv.org/abs/2104.01883