Amplitude-Constrained Gaussian Wiretap Channel: Computation of the Optimal Input 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 paper studies a scalar Gaussian wiretap channel where instead of an average input power constraint, we consider a peak amplitude constraint on the input. The goal is to obtain insights into the secrecy-capacity and the structure of the secrecy-capacity-achieving distribution. Capitalizing on the recent theoretical progress on the structure of the secrecy-capacity-achieving distribution, this paper develops a numerical procedure, based on the gradient ascent algorithm and a version of the Blahut-Arimoto algorithm, for computing the secrecy-capacity and the secrecy-capacity-achieving input and output distributions.**

I. Introduction

Consider the scalar 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. The block diagram for the Gaussian wiretap channel is shown in Fig. 1.

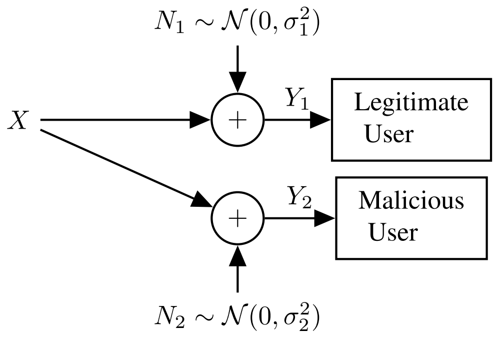

$N_1 \sim \mathcal{N}(0, \sigma_1^2)$
X
Y_1
Legitimate User
Y_2
Malicious User
$N_2 \sim \mathcal{N}(0, \sigma_2^2)$

Fig. 1: The Gaussian wiretap channel.

In this work, we assume that the input X is limited by a peak-power constraint or, equivalently, by a peak amplitude constraint given by $|X| \le \mathsf{A}$. For this setting, the secrecy-capacity is given by

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

$$= \max_{P_X : |X| \le \mathsf{A}} I(X; Y_1 | Y_2), \tag{4}$$

where P_X denotes the distribution of the input random variable X, and where $I(X; Y_i)$ is the mutual information between X and Y_i, $i \in \{1, 2\}$. The step in (4) is due to the degradedness of the channel. We are interested in studying the input distribution $P_{X^\star}$ that maximizes (4). It can be shown that for $\sigma_1^2 \ge \sigma_2^2$ the secrecy-capacity is equal to zero. Therefore, in the remaining, we assume that $\sigma_1^2 < \sigma_2^2$.

The $P_{X^\star}$ and $C_s(\sigma_1, \sigma_2, \mathsf{A})$, besides a few cases, are in general unknown [1]. The expression of the capacity is an important benchmark in communications theory, and knowing the secrecy-capacity-achieving distribution is important as it is useful for the code or modulation design. Furthermore, the availability of numerical examples of $P_{X^\star}$ may also guide theoretical work by pointing out possible properties of $P_{X^\star}$ that we might want to prove. Therefore, it is both of theoretical and practical need to produce numerical example of $P_{X^\star}$ and $C_s(\sigma_1, \sigma_2, \mathsf{A})$. In this work, we take a numerical approach to computing $P_{X^\star}$ and $C_s(\sigma_1, \sigma_2, \mathsf{A})$ and provide numerical examples of these quantities for various parameter regimes of $(\mathsf{A}, \sigma_1, \sigma_2)$. Data from these simulations (e.g., calculated $P_X^\star$ and capacity) are made available at [2].

A. 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 cardinality of $\mathsf{supp}(P_X)$ will be denoted by $|\mathsf{supp}(P_X)|$. The relative entropy between distributions P and Q will be denoted by $\mathsf{D}(P \| Q)$.

Let g_1 and g_2 be nonnegative functions, then $g_1(x) = O(g_2(x))$ means that there exists a constant $c > 0$ and x_0 such that $\frac{g_1(x)}{g_2(x)} \le c$ for all $x > x_0$;

B. Literature Review

The wiretap channel was introduced by Wyner in [3], 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 [4]–[7] 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 [8] where the capacity-achieving input distribution was shown to be Gaussian. The secrecy-capacity of the Gaussian wiretap channel with an amplitude and power constraint was considered by Ozel *et al.* in [1] where the authors showed that the capacity-achieving input distribution is discrete with finitely many support points. Recently, the result of [1] was sharpened in [9] in several ways. First, an explicit upper bound on the number of support points of $P_{X^\star}$ of the following from has been shown:

$$|\mathsf{supp}(P_{X^\star})| \le \rho \frac{\mathsf{A}^2}{\sigma_1^2} + O(\log(\mathsf{A})), \tag{6}$$

where $\rho = (2\mathrm{e}+1)^2 \left(\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}\right)^2 + \left(\frac{\sigma_2+\sigma_1}{\sigma_2-\sigma_1}+1\right)^2$. Second, the following two lower bounds on the cardinality of the support of the following form have been shown:

$$\begin{aligned} |\mathsf{supp}(P_{X^\star})| &\ge \sqrt{1 + \frac{\frac{2\mathsf{A}^2}{\pi \mathrm{e} \sigma_1^2}}{1+\frac{\mathsf{A}^2}{\sigma_2^2}} \mathrm{e}^{I(X^\star;Y_2)}} && (7) \\ &\ge \sqrt{1 + \frac{\frac{2\mathsf{A}^2}{\pi \mathrm{e} \sigma_1^2}}{1+\frac{\mathsf{A}^2}{\sigma_2^2}}}, && (8) \end{aligned}$$

where the second bound follows from the trivial bound $\mathrm{e}^{I(X^\star;Y_2)} \ge 1$.

In this work, specifically capitalizing on the upper bound in (6), we propose an algorithm that produces extensive numerical examples for $P_{X^\star}$. A firm upper bound on the size of the support in [9] allows us to move the optimization from the space of probability distributions to the $\mathbb{R}^{2n}$ space, where n is the number of support points of $P_{X^\star}$ (i.e., $n = |\mathsf{supp}(P_{X^\star})|$). Working in $\mathbb{R}^{2n}$ allows us to employ finite-dimensional methods such as the gradient ascent together with a variant of the Blahut-Arimoto algorithm [10] for the wiretap channel [11]. The firm bound n also guarantees that our algorithm will converge to at least a local maximum. To guarantee our solution is close to the global maximum, we also check it against the Karush-Kuhn-Tucker (KKT) conditions. We note that there are several other numerical recipes for generating an optimal input distribution, and the interested reader is referred to [10], [12], [13]. However, most of these approaches ultimately optimize over the space of distributions, which is an infinite-dimensional space.

To the best of our knowledge, there are no existing prior works that focus on generating numerical examples of secrecy-capacity-achieving distribution for the Gaussian wiretap channel in the regime where $P_{X^\star}$ is not available analytically. However, there is extensive prior work on generating examples of the capacity-achieving distribution for the point-to-point channel, which is a special case of the wiretap channel when $\sigma_2 = \infty$; The interested reader is referred to [14]–[16] for examples of such implementations.

C. Outline and Contributions

The outline and the contribution of the paper are as follows. Section II will present the following: the KKT conditions needed for the optimality of $P_{X^\star}$; the algorithm that will produce numerical examples of $P_{X^\star}$ and other quantities related to $P_{X^\star}$. Section III will present our extensive numerical simulations and the observations surrounding these simulations.

II. Proposed Algorithm

In this section, we first present the KKT equations, which provide sufficient and necessary conditions for the optimality of $P_{X^\star}$. Second, we present the algorithm which will be used to generate numerical examples of $P_{X^\star}$. The proposed algorithm is the combination of a version of the Blahut-Arimoto algorithm and the gradient ascent algorithm. The key step of the algorithm evaluates the KKT conditions with the candidate solution, which ensures that the solution is close to the true solution.

A. KKT conditions

The starting point for the design of our algorithm are the following KKT conditions shown in [1].

Lemma 1. *The secrecy-capacity-achieving input distribution $P_{X^\star}$ and induced secrecy-capacity-achieving output distributions $P_{Y_1^\star}$ and $P_{Y_2^\star}$ satisfy the following*

$$\begin{aligned} \Xi(x; P_{X^\star}) &= C_s(\sigma_1, \sigma_2, \mathsf{A}), \qquad x \in \mathsf{supp}(P_{X^\star}), && (9a) \\ \Xi(x; P_{X^\star}) &\le C_s(\sigma_1, \sigma_2, \mathsf{A}), \qquad x \in [-\mathsf{A}, \mathsf{A}], && (9b) \end{aligned}$$

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

$$\Xi(x; P_{X^\star}) = \mathsf{D}(P_{Y_1|X}(\cdot|x) \| P_{Y_1^\star}) - \mathsf{D}(P_{Y_2|X}(\cdot|x) \| P_{Y_2^\star}). \tag{10}$$

In the numerical computation of $P_{X^\star}$ we find more convenient to check the negated version of (9), where a tolerance parameter ε is introduced which trades off accuracy with computational burden. Specifically, P_X is not an optimal input pmf if any of the following conditions is satisfied:

$$\begin{aligned} &|\Xi(x; P_X) - \Xi(\mathsf{A}; P_X)| > \varepsilon, \text{ for some } x \in \mathsf{supp}(P_X) && (11a) \\ &\Xi(\mathsf{A}; P_X) + \varepsilon < \Xi(x; P_X), \text{ for some } x \in [-\mathsf{A}, \mathsf{A}]. && (11b) \end{aligned}$$

Note that in (11) in place of the secrecy-capacity $C_s(\sigma_1, \sigma_2, \mathsf{A})$, which is unknown, we used the value of $\Xi(\mathsf{A}; P_X)$, thanks to the fact that $\mathsf{A} \in \mathsf{supp}(P_{X^\star})$ for any $(\sigma_1, \sigma_2, \mathsf{A})$. Condition (11a) is derived by negating (9a): there exists an $x \in \mathsf{supp}(P_X)$ such that $\Xi(x; P_X)$ is ε-away from the estimated secrecy-capacity $\Xi(\mathsf{A}; P_X)$. Condition (11b) is the negated version of (9b): there exists an $x \in [-\mathsf{A}, \mathsf{A}]$ such that $\Xi(x; P_X)$ is at least ε-larger than the estimated secrecy-capacity $\Xi(\mathsf{A}; P_X)$. With some abuse of notation, we refer to (11) as to the ε-KKT conditions.

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 by suitably changing the KKT conditions as reported in [1, Eq. (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.

B. Numerical Algorithm

The algorithm that gives a numerical lower bound to secrecy-capacity and that estimates the optimal input pmf $P_{X^\star}$ is given in Algorithm 1. The algorithm takes as input: the value A of the amplitude constraint; the values (σ_1, σ_2) of the standard deviations of the additive noise; the vector $\mathbf{x}$ of the support points of the initial tentative P_X; the vector $\mathbf{p}$ of the probabilities associated with the points in $\mathbf{x}$; and a value ε which is related to the accuracy of the capacity estimation. Using the bound in [9, Eq. (83)], we can set the dimensionally of $\mathbf{p}$ and $\mathbf{x}$ to be less than

$$b_1 \frac{\mathsf{A}^2}{\sigma_1^2} + b_2 + \log \frac{b_3\mathsf{A} + b_4\mathsf{A} + b_5}{b_6\mathsf{A} + b_7}, \tag{12}$$

where the constants $\{b_i\}_{i=1}^7$ depend on $(\sigma_1, \sigma_2, C_s)$. We can easily prove that if $x \in \mathsf{supp}(P_{X^\star})$, then $-x \in \mathsf{supp}(P_{X^\star})$ and $P_{X^\star}(x) = P_{X^\star}(-x)$; We can exploit the symmetry by storing in $\mathbf{x}$ only the nonnegative valued support points.

The optimization procedure is iterative and divided into two parts: (i) N_{BA} iterations of the Blahut-Arimoto algorithm are used to let the tentative pmf converge to stable values, and (ii) N_{GA} iterations of a gradient ascent algorithm are used to modify the positions of the support points of the tentative input distribution. The objective function of the gradient ascent is the secrecy-information $I(X;Y_1) - I(X;Y_2)$. We use a backtracking line search version of the gradient ascent [17], in order to ensure that the sequence of secrecy-information values obtained along the iterations is non-decreasing, which ensures convergence to a local maximum.

There are several ways to perform the initialization of Algorithm 1. For example, the initial P_X can be chosen to be uniformly distributed and uniformly spaced with the number of points given by the upper bound in (6). Alternatively, if we already have the solution for some amplitude A_1, and seek to find the optimal input distribution for amplitude A_2 that is 'close' to A_1, then the solution for A_1 can be used as the initial condition for A_2. To generate results in this work, we use the latter procedure.

The Blahut-Arimoto algorithm is a well-known computation method in information theory. Thus, we do not go over the details of its implementation, and the interested reader can refer to classical text such as [18] or its implementation for the wiretap channel in [11].

After the two-step optimization, if there is a set of support points that are driven too close to each other by the gradient ascent, then such points are clustered together in a new support point which assumes a probability equal to the sum of the probabilities of the clustered points, and the cluster center is taken to be the average of the clustered points. Note that the clustering operation is equivalent to having a minimum distance constraint among support points. We have observed that this constraint in general helps numerical stability. We have chosen a minimum distance of 10^{-2}, which seems to be an inactive capacity constraint for all values of $\mathsf{A} > 10^{-2}$ and of (σ_1, σ_2). We caution, however, this constraint might be active around the transition points (i.e., when one point splits into two), but becomes inactive once again after we move away by 10^{-2} from the transition point.

Algorithm 1 Capacity and input PMF estimation

```
procedure MAIN($\mathsf{A}, \sigma_1, \sigma_2, \mathbf{x}, \mathbf{p}, \varepsilon$)
    $N_{\mathsf{BA}} \leftarrow 100$ / Number of Blahut-Arimoto iterations
    $N_{\mathsf{GA}} \leftarrow 20$ / Number of gradient ascent iterations
    repeat
        $k \leftarrow 0$
        while $k < 100$ do
            $k \leftarrow k + 1$
            $\mathbf{p} \leftarrow$ BLAHUT-ARIMOTO($\mathbf{x}, \mathbf{p}, N_{\mathsf{BA}}$)
            $\mathbf{x} \leftarrow$ GRADIENT-ASCENT($\mathbf{x}, \mathbf{p}, N_{\mathsf{GA}}$)
        end while
        $(\mathbf{x}, \mathbf{p}) \leftarrow$ CLUSTER($\mathbf{x}, \mathbf{p}$)
        valid $\leftarrow$ KKT-VALIDATION($\mathbf{x}, \mathbf{p}, \varepsilon$)
        if valid = False then
            $(\mathbf{x}, \mathbf{p}) \leftarrow$ UPDATE($\mathbf{x}, \mathbf{p}$)
        end if
    until valid = True
    $C_s(\sigma_1, \sigma_2, \mathsf{A}) \leftarrow \Xi(\mathbf{x}, \mathbf{p})$
    return $\mathbf{x}, \mathbf{p}, C_s(\sigma_1, \sigma_2, \mathsf{A})$
end procedure
```

The optimality of the resulting distribution P_X is tested with the ε-KKT conditions. If any of the conditions (11) is satisfied, then the function UPDATE modifies P_X according to the following rules. Let

$$\widehat{x} = \arg \max_{x \in [-\mathsf{A}, \mathsf{A}]} \Xi(x; P_X) \tag{13}$$

be the candidate novel support point (or points, due to the even symmetry of $\Xi(\cdot; P_X)$), and

$$\mathcal{S} = \{x \in \mathsf{supp}(P_X) : \text{(11a) is true}\} \tag{14}$$

be the set of support points for which $\Xi(\cdot; P_X)$ falls outside the horizontal ε-strip that contains $\Xi(\mathsf{A}; P_X)$. If both conditions in (11) are verified and there exist $x_1, x_2 \in \mathcal{S}$ such that $|x_1 - x_2| < \delta = 0.1$ and $\widehat{x} \in [x_1, x_2]$, then $\widehat{x}$ replaces both x_1 and x_2 in the support of P_X with $P_X(\widehat{x}) = P_X(x_1) + P_X(x_2)$. Otherwise, if only (11b) is satisfied, then $\widehat{x}$ is added to the support of P_X and the probabilities are set to $P_X(x) = |\mathsf{supp}(P_X)|^{-1}$ for all $x \in \mathsf{supp}(P_X)$.

The main algorithm repeats the two-step optimization procedure until P_X is validated, which becomes the proposed $P_{X^\star}$.

III. Simulations and Observations

In this section, we use Algorithm 1 to provide examples of the $P_{X^\star}$. For several values of the parameters $(\sigma_1, \sigma_2, \mathsf{A})$, we will plot the following: $C_s(\sigma_1, \sigma_2, \mathsf{A})$; the location of the support points of $P_{X^\star}$; the cardinality of $\mathsf{supp}(P_{X^\star})$; the output pdfs $P_{Y_1^\star}$ and $P_{Y_2^\star}$. The simulation data results are publicly available [2].

We consider two regimes. In the first regime we examine a scenario where the eavesdropper experiences significantly large noise than the legitimate user and specifically we consider $\sigma_1^2 = 1$ and $\sigma_2^2 = 10$. In the second regime, we examine a scenario where the noise at both users is comparable, and specifically, we consider $\sigma_1^2 = 1$ and $\sigma_2^2 = 1.5$.

Fig. 2 presents the structure of the support of the secrecy-capacity-achieving distribution vs. A and the values of $C_s(\sigma_1, \sigma_2, \mathsf{A})$ vs. A. We observe the following:

- In Fig. 2a, we consider discrete values for A and for each value of A we plot the corresponding $\mathsf{supp}(P_{X^\star})$, where thanks to symmetry of $P_{X^\star}$ we only plot the nonnegative points in $\mathsf{supp}(P_{X^\star})$. Different colors refers to different values of σ_2. In addition, in Fig. 2b we normalize the values of the support and plot the nonnegative values of $\frac{\mathsf{supp}(P_{X^\star})}{\mathsf{A}}$ vs. A. First, we see that, as A increases, the new points appear only at zero while the nonzero points travel away from the origin. Moreover, the points are created in only the following two ways: either a point at zero simply emerges, or an existing point at zero splits into two points. This behavior is analogous to the behavior of the support of the capacity-achieving distribution for the point-to-point channel [19].

  Second, note that the support points of $P_{X^\star}$ for $\sigma_2^2 = 10$ lag behind (*i.e.*, for larger values of A) the support points of $P_{X^\star}$ for $\sigma_2^2 = 1.5$. This could be explained by noting that when the eavesdropper faces less additive noise, the transmitter has to create more equivocation in order to keep the leakage information rate at the eavesdropper equal to zero: This could be done only by adding a new support point in $P_{X^\star}$.
- In Fig. 2c we plot the cardinality of $\mathsf{supp}(P_{X^\star})$ vs. A. First, we observe that the cardinality increases linearly in A. Should this behavior continue for larger values of A, the upper bound in (6) will be loose. Note a similar issue exists for a point-to-point channel with a peak amplitude constraint where the cardinality of the support of the capacity-achieving distribution is lower-bounded by a term in the order of A while the upper bound is in the order of A^2 [19]. Therefore, since the point-to-point channel is a special case of the wiretap channel, it is not surprising that similar behavior occurs here too.

  Second, Fig. 2c suggests that the bound in (7) might be loose. However, we conjecture that this bound, while it may not be attaining the correct constants, is tight asymptotically as A increases.
- In Fig. 2d we plot the difference between the nonnegative adjacent points in $\mathsf{supp}(P_{X^\star})$ starting from the largest and the second largest. An interesting observation here is that the distance between two adjacent points eventually stops fluctuating and concentrates on a single value. In other words, the space between adjacent points eventually does not change. However, it is important to note that the spacing is not uniform. Along this direction, an interesting future direction would be to consider the asymptotic behavior of $\frac{X^\star}{\mathsf{A}}$ and find the limiting distribution as $\mathsf{A} \to \infty$.
- In Fig. 2e we plot the secrecy-capacity vs. A. It is interesting to note that the secrecy-capacity is not concave for small values A. There are two reasons for this. First, note that C_s is given as the difference of mutual informations. Therefore, even if each mutual information is concave in A, this does not imply that C_s is concave in A as the difference of concave functions is not concave. Second, note that even the single mutual information may not be concave in A: See Fig. 2f where the mutual information is plotted vs. A. Note that because of the I-MMSE relationship [20], the mutual information is a concave function of the signal-to-noise ratio and is, therefore, concave in A^2. However, this does not imply that the mutual information is concave in A. In fact, as we see from Fig. 2f, for small values of A, the mutual information can be convex in A.

In addition to studying the optimal input distribution, it is also interesting to observe the behavior of the output distributions $P_{Y_1^\star}$ at the legitimate user and $P_{Y_2^\star}$ at the eavesdropper, and to compare them. Fig. 3 provides examples of $P_{Y_1^\star}$, $P_{Y_2^\star}$, and $P_{X^\star}$ for $\sigma_1^2 = 1$ and $\sigma_2^2 = 10$, and Fig. 4 provides examples of $P_{Y_1^\star}$, $P_{Y_2^\star}$, and $P_{X^\star}$ for $\sigma_1^2 = 1$ and $\sigma_2^2 = 1.5$.

The eavesdropper's pdf $P_{Y_2^\star}$ and the legitimate user's pdf $P_{Y_1^\star}$ have clear differences. In particular, $P_{Y_2^\star}$ has just one inflection point for positive abscissa values, and very much resembles the Gaussian pdf. To aid this comparison of $P_{Y_2^\star}$ to the Gaussian distribution, we also plot P_{Y_G} where Y_G is the Gaussian random variable with variance $\mathbb{E}[(X^\star)^2] + \sigma_2^2$ (*i.e.*, Y_G and $Y_2^\star$ have the same variance). In contrast, $P_{Y_1^\star}$ has many peaks and inflection points, which correspond to the locations of the support points of $P_{X^\star}$. Because of the difference in structure of the two pdfs, estimating $X^\star$ is *easier* from the sample $Y_1^\star$ than from the sample $Y_2^\star$. Of course, this is expected, as we desire the eavesdropper to gain little information from the observation $Y_2^\star$. The proximity to Gaussian can also be seen through the mutual information. In particular, Fig. 2f plots $I(X^\star; Y_2^\star)$ and compares it to the mutual information attained by the Gaussian input with same variance as $X^\star$.

As one possible future direction, it would be interesting to show that $P_{Y_2^\star}$ is indeed close to Gaussian in some distance between distributions (e.g., relative entropy). It would also be interesting to study how the smoothnesses of $P_{Y_1^\star}$ and $P_{Y_2^\star}$ differ. This can be accomplished by studying the maximum number of inflection points of each pdf.

The interested reader is also referred to our recent extension of this work to the vector wiretap channel [21].

References

[1] 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.

[2] L. Barletta and A. Dytso. (2021) Simulated data. [Online]. Available: https://github.com/ucando83/WiretapCapacity

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

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

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

[6] 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.

[7] 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.

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

[9] L. Barletta and A. Dytso, “Scalar Gaussian wiretap channel: Bounds on the support size of the secrecy-capacity-achieving distribution,” in *2021 IEEE Information Theory Workshop (ITW)*, 2021.

[10] R. Blahut, “Computation of channel capacity and rate-distortion functions,” *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460–473, 1972.

[11] K. Yasui, T. Suko, and T. Matsushima, “An algorithm for computing the secrecy capacity of broadcast channels with confidential messages,” in *2007 IEEE International Symposium on Information Theory*. IEEE, 2007, pp. 936–940.

[12] C.-I. Chang and L. D. Davisson, “On calculating the capacity of an infinite-input finite (infinite)-output channel,” *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 1004–1010, 1988.

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

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

[15] A. Favano, M. Ferrari, M. Magarini, and L. Barletta, “The capacity of the amplitude-constrained vector Gaussian channel,” in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 426–431.

[16] D. Xiao, L. Wang, D. Song, and R. D. Wesel, “Finite-support capacity-approaching distributions for AWGN channels,” in *2020 IEEE Information Theory Workshop (ITW)*. IEEE, 2021, pp. 1–5.

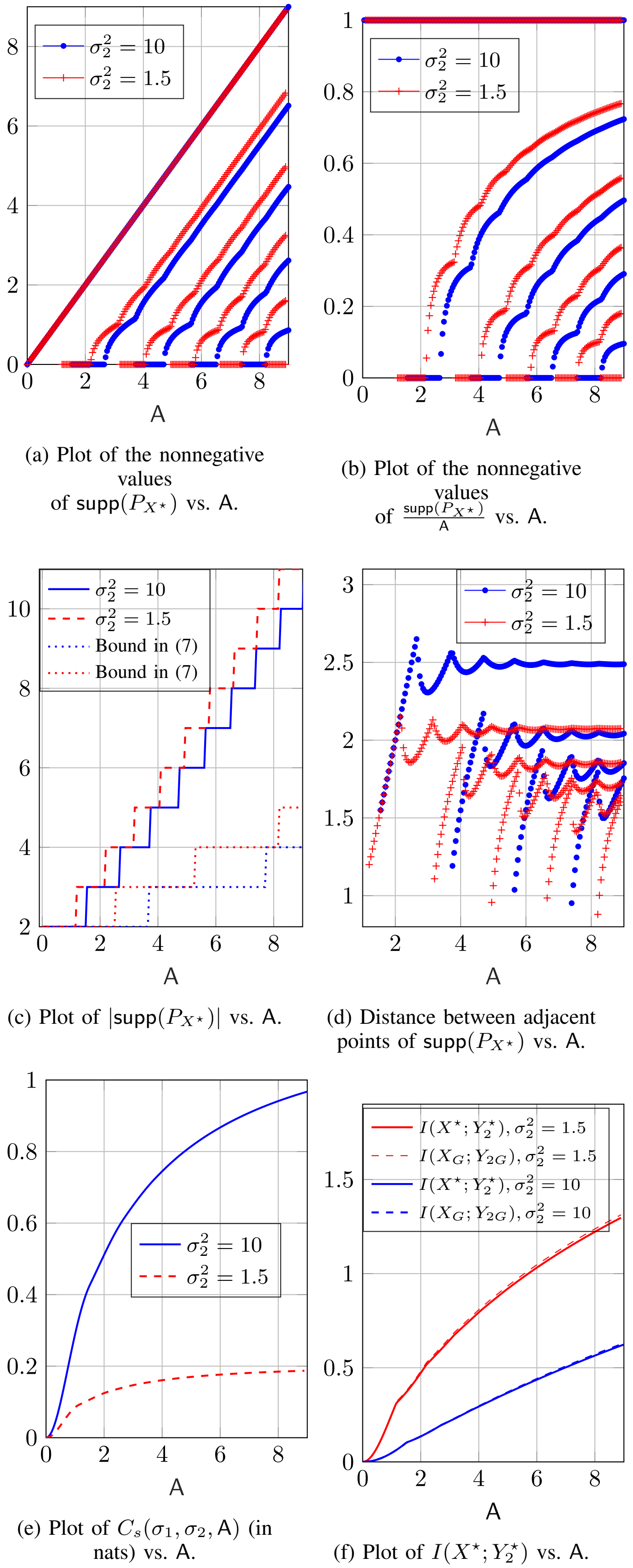
[17] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[18] T. Cover and J. Thomas, *Elements of Information Theory: Second Edition*. Wiley, 2006.

[19] 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.

[20] D. Guo, S. Shamai, and S. Verdú, “Mutual information and minimum mean-square error in Gaussian channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[21] A. Favano, L. Barletta, and A. Dytso, “On the capacity achieving input of amplitude constrained vector Gaussian wiretap channel,” in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022. [Online]. Available: https://arxiv.org/pdf/2202.00586.pdf



(a) Plot of the nonnegative values of $\mathsf{supp}(P_{X^\star})$ vs. A.

(b) Plot of the nonnegative values of $\frac{\mathsf{supp}(P_{X^\star})}{\mathsf{A}}$ vs. A.

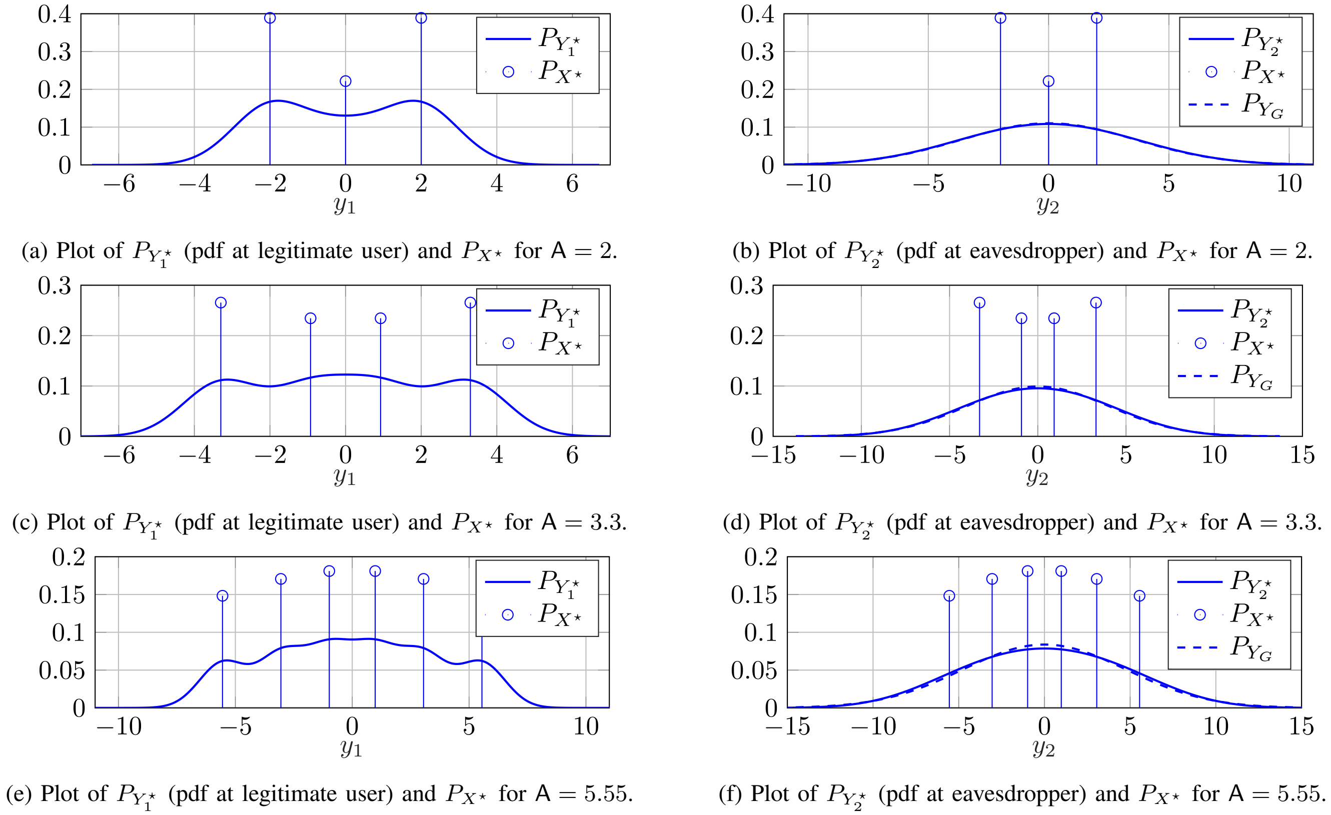
(c) Plot of $|\mathsf{supp}(P_{X^\star})|$ vs. A.

(d) Distance between adjacent points of $\mathsf{supp}(P_{X^\star})$ vs. A.

(e) Plot of $C_s(\sigma_1, \sigma_2, \mathsf{A})$ (in nats) vs. A.

(f) Plot of $I(X^\star; Y_2^\star)$ vs. A.

Fig. 2: Simulations results for the case of $\sigma_1^2 = 1$ and $\sigma_2^2 = \{1.5, 10\}$.



(a) Plot of $P_{Y_1^\star}$ (pdf at legitimate user) and $P_{X^\star}$ for $\mathsf{A} = 2$.

(b) Plot of $P_{Y_2^\star}$ (pdf at eavesdropper) and $P_{X^\star}$ for $\mathsf{A} = 2$.

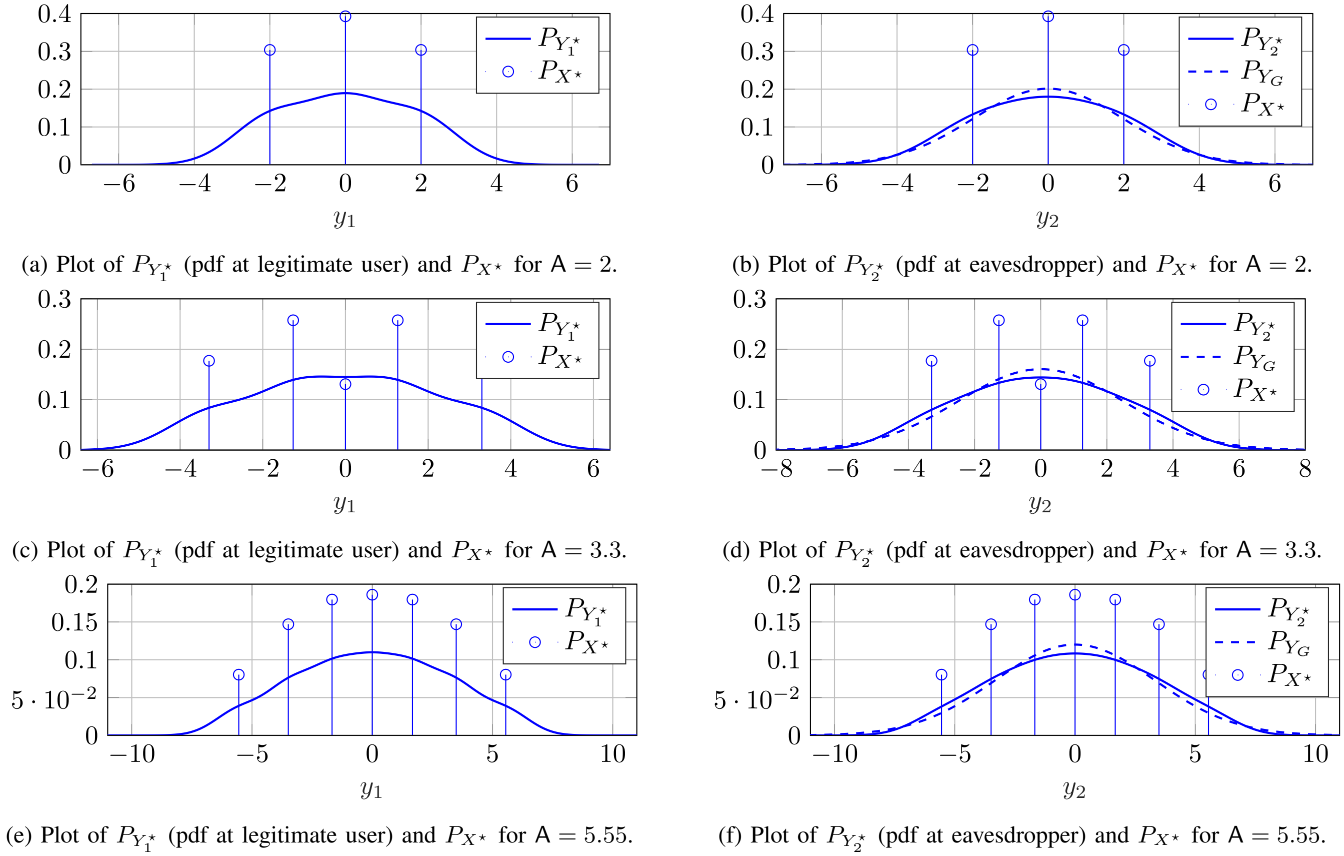
(c) Plot of $P_{Y_1^\star}$ (pdf at legitimate user) and $P_{X^\star}$ for $\mathsf{A} = 3.3$.

(d) Plot of $P_{Y_2^\star}$ (pdf at eavesdropper) and $P_{X^\star}$ for $\mathsf{A} = 3.3$.

(e) Plot of $P_{Y_1^\star}$ (pdf at legitimate user) and $P_{X^\star}$ for $\mathsf{A} = 5.55$.

(f) Plot of $P_{Y_2^\star}$ (pdf at eavesdropper) and $P_{X^\star}$ for $\mathsf{A} = 5.55$.

Fig. 3: Examples of input and output distributions for the case of $\sigma_1^2 = 1$ and $\sigma_2^2 = 10$.



(a) Plot of $P_{Y_1^\star}$ (pdf at legitimate user) and $P_{X^\star}$ for $\mathsf{A} = 2$.

(b) Plot of $P_{Y_2^\star}$ (pdf at eavesdropper) and $P_{X^\star}$ for $\mathsf{A} = 2$.

(c) Plot of $P_{Y_1^\star}$ (pdf at legitimate user) and $P_{X^\star}$ for $\mathsf{A} = 3.3$.

(d) Plot of $P_{Y_2^\star}$ (pdf at eavesdropper) and $P_{X^\star}$ for $\mathsf{A} = 3.3$.

(e) Plot of $P_{Y_1^\star}$ (pdf at legitimate user) and $P_{X^\star}$ for $\mathsf{A} = 5.55$.

(f) Plot of $P_{Y_2^\star}$ (pdf at eavesdropper) and $P_{X^\star}$ for $\mathsf{A} = 5.55$.

Fig. 4: Examples of input and output distributions for the case of $\sigma_1^2 = 1$ and $\sigma_2^2 = 1.5$.