Capacity-Key Tradeoff in Covert Communication

Abdelaziz Bounhar[⋆], Mireille Sarkiss[§], and Michèle Wigger[†]
⋆MBZUAI, Paris, France. Email: abdelaziz.bounhar@mbzuai.ac.ae
[§]SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France
Email: mireille.sarkiss@telecom-sudparis.eu
[†]LTCI, Télécom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France. Email: michele.wigger@telecom-paris.fr

***Abstract*—This paper explores the tradeoff between covert communication capacity and secret key requirements over discrete memoryless channels (DMCs). We focus on settings where under a covertness constraint both communication and key rates are measured as the number of bits per square root of the blocklength. While previous work has identified the maximum covert communication rates and the corresponding minimum key rates needed to achieve them, our study characterizes the minimum key rates necessary for all of desired covert communication rates. In equivalent terms, we determine, for any given key rate, the set of achievable covert rates. This relationship defines what we call the covert capacity-key tradeoff.**

Our analysis reveals several new insights. In scenarios where only small key rates are available and the adversary has a stronger channel than the intended receiver, binary signaling is optimal—regardless of the specific channel characteristics or input alphabets. In these cases, the covert capacity increases linearly with the available key rate. In other cases and for larger key rates, the covert capacity-key tradeoff grows sublinearly.

We also extend our findings to multi-access channels (MACs) with binary inputs.

***Index Terms*—Covert communication, key rates, covert capacity-key tradeoff, discrete memoryless networks.**

I. Introduction

This paper investigates the problem of *covert* communication, where the goal is to ensure that an external observer—referred to as the warden—cannot determine whether communication is taking place. Specifically, we focus on characterizing the maximum number of data bits that an encoder can reliably transmit over a channel without being detected by the warden.

The fundamental limits of covert communication were first established for Gaussian channels in [1], where covertness was formalized via a constraint on the Kullback—Leibler (KL) divergence between the distributions of the wardens observations in the presence and absence of communication. This divergence is required to vanish as the communication blocklength increases.

A key result in [1] is the *square-root law*, which states that the number of bits that can be covertly transmitted scales with the square root of the number of channel uses, and not linearly in the number of channel uses for traditional non-covert communication. The square-root law has since been shown to hold across a variety of channel models, including general Gaussian and discrete memoryless channels (DMCs) [2]–[5].

Subsequent work has extended these results to more complex scenarios [6]–[12] and network scenarios [13]–[16].

Prior studies [3], [16] have characterized the key rate needed to attain the *covert capacity*, i.e., the maximum covert rate under unlimited key resources. However, a more refined question remains open: given a fixed key budget, what covert rates are achievable? This tradeoff, termed the *covert capacity—key tradeoff*, was first explored for binary-input DMCs in [15], where it was shown to grow linearly with key rate before saturating.

In this paper, we extend this analysis to DMCs with arbitrary input alphabets. We find that the covert capacity—key tradeoff depends critically on the relative strength of the warden's and legitimate receiver's (Rx) channels across input symbols, measured by the KL divergence between their respective output distributions. When the warden is uniformly stronger over all inputs, binary signaling is optimal at low key rates, irrespective of the input alphabets of the DMC, and the covert capacity-key tradeoff grows linearly in this regime. When the warden is uniformly weaker than the legitimate Rx, covert capacity is achievable without any key. For mixed cases, the tradeoff is sublinear and binary signaling suboptimal.

We also characterize the covert capacity—key tradeoff region for a two-user multiple-access channel with binary inputs. Our results show that the interaction between the two users' covert rates depends heavily on the relative strength of the warden with respect to the legitimate Rx, revealing regimes of both independence and tradeoff.

Notation: In this paper, we follow standard information theory notations. We use calligraphic fonts for sets (e.g. $\mathcal{S}$) and note by $|\mathcal{S}|$ the cardinality of a set $\mathcal{S}$. Random variables are denoted by upper case letters (e.g., X), while their realizations are denoted by lowercase letters (e.g. x). We write X^n and x^n for the tuples $(X_1, \ldots, X_n)$ and $(x_1, \ldots, x_n)$, respectively, for any positive integer $n > 0$. For a distribution P on $\mathcal{X}$, we note its product distribution on $\mathcal{X}^n$ by $P^{\otimes n}(x^n) = \prod_{i=1}^n P(x_i)$. We also denote by $\text{Supp}(P)$ the support of a distribution P, i.e. $\text{Supp}(P) = \{x\colon P(x) \neq 0\}$. For two distributions P and Q on $\mathcal{X}$, we let $\mathbb{D}(P\|Q) = \sum_{x\in\mathcal{X}} P(x) \log\left(\frac{P(x)}{Q(x)}\right)$ denote the Kullback-Leibler (KL) divergence. We shall use Landau notation and We abbreviate *probability mass function* by *pmf* and *independent and identically distributed* by *i.i.d.*.

II. The Single-User Setup

Consider the setup illustrated in Figure 1. The transmitter (Tx) wishes to send a message W to the legitimate Rx while

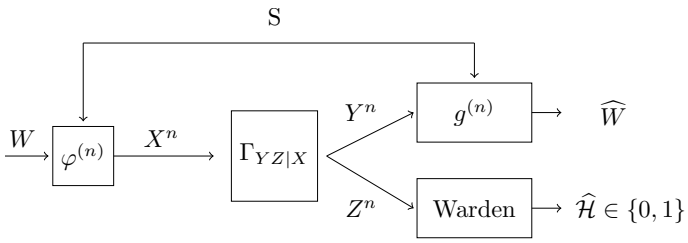

Fig. 1: Point-to-point covert communication over a Discrete Memoryless Channel of law $\Gamma_{YZ|X}$.

avoiding detection by the warden which attempts to detect the presence of communication. Communication takes place over a block of n channel uses. The Tx produces channel inputs in a finite alphabet $\mathcal{X}$ and the legitimate Rx and the warden observe channel outputs within finite alphabets $\mathcal{Y}$ and $\mathcal{Z}$. These outputs are produced by a DMC, that means, if the Tx produces the n channel inputs $X^n = x^n$ then for any $i \in \{1, \dots, n\}$ the i-th output symbols Y_i and Z_i observed at the legitimate Rx and the warden are generated from the i-th input x_i according to the conditional laws $\Gamma_{Y|X}(\cdot|x_i)$ and $\Gamma_{Z|X}(\cdot|x_i)$, repectively.

The Tx encodes message W using some encoding function $\varphi^{(n)}$ defined on appropriate domains, along with a secret-key S. Subsequently, it sends the resulting codeword

$$X^n = \varphi^{(n)}(W, S) \tag{1}$$

over the channel. For readability, we will write $x^n(w, s)$ instead of $\varphi^{(n)}(w, s)$. Let the message W and the secret-key S be represented by two sequences of m and p i.i.d. Bernoulli-$1/2$ bits, where these numbers m and p will depend on the blocklength n. The secret-key S is exclusively known to the Tx and the Rx but not to the warden.

The legitimate Rx estimates the message as:

$$\widehat{W} = g^{(n)}(Y^n, S) \tag{2}$$

using an appropriate decoding function.

To ensure reliability of communication, we seek for systems (encoding and decoding functions) where

$$\lim_{n\to\infty} \Pr[\widehat{W} \neq W] = 0. \tag{3}$$

At the same time we impose that the output distribution implied at the warden

$$\widehat{Q}^n(z^n) \triangleq \frac{1}{2^m 2^p} \sum_{(w,s)} \Gamma_{Z|X}^{\otimes n}(z^n|x^n(w,s)). \tag{4}$$

be almost indistinguishable from the warden's output distribution when the all-zero sequence is transmitted (which stands for absence of communication), i.e., from

$$\Gamma_{Z|X}^{\otimes n}(z^n|0^n). \tag{5}$$

Notice that we impose that the 0-symbol be part of the input alphabet $\mathcal{X}$ and then require that the KL divergence

$$\delta_n \triangleq \mathbb{D}\left(\widehat{Q}^n(\cdot) \middle\| \Gamma_{Z|X}^{\otimes n}(\cdot|0^n)\right) \tag{6}$$

vanishes in the regime of large blocklengths $n \to \infty$.

A. Achievable Covert Rates and Covert Capacity-Key Tradeoff

Our goal is to determine the largest data rate in function of the available key rate. In covert-communication tradition [2]–[4] rates are obtained by scaling the number of message or secret key bits with the square-root of the number of channel uses and the divergence-δ_n. This yields the following definition.

Definition 1: For any given $k \geq 0$, define the *covert capacity-key tradeoff* $r^\star(k)$ as the largest rate r such that there exists a sequence of tuples (m, p) and encoding/decoding functions $\{(\varphi^{(n)}, g^{(n)})\}_n$ satisfying

$$\lim_{n\to\infty} \Pr[\widehat{W} \neq W] = 0 \tag{7a}$$

and

$$r \leq \liminf_{n\to\infty} \frac{m}{\sqrt{n\delta_n}}, \tag{8a}$$

$$k \geq \limsup_{n\to\infty} \frac{p}{\sqrt{n\delta_n}}. \tag{8b}$$

To ensure that the problem is non-degenerate, we assume:

$$\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x)\Gamma_{Z|X}(\cdot|x) \neq \Gamma_{Z|X}(\cdot|0), \quad \forall \psi(\cdot), \tag{9a}$$

$$\text{Supp}\left(\Gamma_{Z|X}(\cdot|x)\right) \subseteq \text{Supp}\left(\Gamma_{Z|X}(\cdot|0)\right), \quad \forall x \in \mathcal{X}, \tag{9b}$$

$$\text{Supp}\left(\Gamma_{Y|X}(\cdot|x)\right) \subseteq \text{Supp}\left(\Gamma_{Y|X}(\cdot|0)\right), \quad \forall x \in \mathcal{X} \tag{9c}$$

where in the above, $\psi(\cdot)$ indicates any pmf over $\mathcal{X}\backslash\{0\}$.

Under these assumptions, we define the covert capacity as:

$$C_{\text{covert}} := \sup_{k\geq 0} r^\star(k). \tag{10}$$

III. Single-User Results

Given a pmf $\psi(\cdot)$ over $\mathcal{X}\backslash\{0\}$, we use the abbreviations

$$\chi^2(\psi) \triangleq \sum_{z\in\mathcal{Z}} \frac{\left(\sum_{x\in\mathcal{X}} \psi(x)\Gamma_{Z|X}(z|x) - \Gamma_{Z|X}(z|0)\right)^2}{\Gamma_{Z|X}(z|0)}. \tag{11}$$

Moreover, for given $x \in \mathcal{X}\backslash\{0\}$, we define:

$$\mathbb{D}_Y(x) \triangleq \mathbb{D}\left(\Gamma_{Y|X}(\cdot|x)\|\Gamma_{Y|X}(\cdot|0)\right), \tag{12}$$

$$\mathbb{D}_Z(x) \triangleq \mathbb{D}\left(\Gamma_{Z|X}(\cdot|x)\|\Gamma_{Z|X}(\cdot|0)\right). \tag{13}$$

Finally, we define a function $k \mapsto f(\psi, k)$ for each choice of the pmf $\psi(\cdot)$. The definition of $k \mapsto f(\psi, k)$ depends on whether the difference

$$\Delta_\psi \triangleq \sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x) \cdot (\mathbb{D}_Z(x) - \mathbb{D}_Y(x)) \tag{14}$$

is positive or not. For pmfs ψ for which $\Delta_\psi > 0$, we define

$$f(\psi, k) = \min\left\{ \frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x) \cdot \mathbb{D}_Y(x)}{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x) \cdot (\mathbb{D}_Z(x) - \mathbb{D}_Y(x))} \cdot k, \right.$$
$$\left. \sqrt{2}\frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x) \cdot \mathbb{D}_Y(x)}{\sqrt{\chi^2(\psi)}} \right\}, \tag{15}$$

and for pmfs ψ for which $\Delta_\psi \leq 0$, we define

$$f(\psi, k) = C_\psi \triangleq \sqrt{2}\frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x)\cdot \mathbb{D}_Y(x)}{\sqrt{\chi^2(\psi)}}, \quad (16)$$

So, if $\Delta_\psi \leq 0$ the function is constant equal to C_ψ. If $\Delta_\psi > 0$, the function is a straight line from the origin $(k = 0, r = 0)$ to the point

$$k = k_\psi \triangleq \sqrt{2}\frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x)\cdot (\mathbb{D}_Z(x) - \mathbb{D}_Y(x))}{\sqrt{\chi^2(\psi)}} \quad (17)$$

$$r = C_\psi \triangleq \sqrt{2}\frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x)\cdot \mathbb{D}_Y(x)}{\sqrt{\chi^2(\psi)}}, \quad (18)$$

and is constant equal to $f(\psi, k) = C_\psi$ for all key rates $k \geq k_\psi$.

We now state the covert capacity-key tradeoff of a DMC:

Theorem 1: We have

$$r^\star(k) = \max_\psi f(\psi, k), \quad (19)$$

where the maximum is over all pmfs $\psi(\cdot)$ on $\mathcal{X}\backslash\{0\}$, and

$$C_{\text{covert}} := \sup_\psi C_\psi. \quad (20)$$

Proof: Omitted due to lack of space. ∎

A. Simplification of the Covert Capacity-Key Tradeoff $r^\star(k)$

To further analyze $r^\star(k)$, we distinguish three cases depending on whether Δ_ψ is positive for all pmfs ψ, is negative for all ψ, or positive for some ψs and negative for others.

Consider an example where Δ_ψ is always positive:

Example 1 (Adversary Stronger than Legitimate Rx): Consider a channel with input alphabet $\mathcal{X} = \{0, 1, 2\}$, output alphabets $\mathcal{Z} = \mathcal{Y} = \{0, 1, 2, 3, 4\}$, and transition pmfs:

$$\Gamma_{Y|X} = \begin{bmatrix} 0.23,\ 0.20,\ 0.25,\ 0.05,\ 0.27 \\ 0.35,\ 0.22,\ 0.10,\ 0.05,\ 0.28 \\ 0.27,\ 0.24,\ 0.24,\ 0.07,\ 0.18 \end{bmatrix} \quad (21a)$$

and

$$\Gamma_{Z|X} = \begin{bmatrix} 0.22,\ 0.32,\ 0.15,\ 0.12,\ 0.19 \\ 0.36,\ 0.03,\ 0.39,\ 0.21,\ 0.01 \\ 0.31,\ 0.20,\ 0.07,\ 0.22,\ 0.20 \end{bmatrix} \quad (21b)$$

where the rows pertain to the three input symbols $x = 0, 1, 2$ and the five columns to the five output symbols $0, 1, 2, 3, 4$. For this channel the difference $\Delta_\psi > 0$ for all pmfs ψ, since for all $x \in \mathcal{X}\backslash\{0\}$ we have $\mathbb{D}_Z(x) - \mathbb{D}_Y(x) > 0$. Figure 2 illustrates the function $k \mapsto f(\psi, k)$ for different choices of the input pmf $\psi(\cdot)$. The line with $+$ markers corresponds to the degenerate choice $\psi(1) = 1$ and the line with $*$-markers to the degenerate choice $\psi(2) = 1$.

We observe that in this example the extreme lines $f(\psi, k)$ with smallest and largest slopes correspond to the degenerate pmfs that put probability 1 on a single non-zero input. This is true in general as can be seen by the following lemma and noting that the slope of $f(\psi, k)$ is given by

$$S_\psi \triangleq \frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x)\cdot \mathbb{D}_Y(x)}{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x)\cdot (\mathbb{D}_Z(x) - \mathbb{D}_Y(x))} > 0. \quad (22)$$

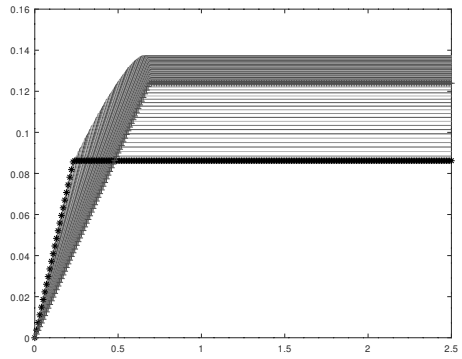

Fig. 2: Figure illustrates the functions $k \mapsto f(\psi, k)$ for different pmfs ψ. The capacity-secret-key tradeoff $r^\star(k)$ corresponds to the upper convex hull of all these curves.

Lemma 1: If

$$\mathbb{D}_Z(x) - \mathbb{D}_Y(x) > 0, \qquad \forall x \in \mathcal{X}\backslash\{0\}, \quad (23)$$

then $S_\psi > 0$ for all input pmfs $\psi(\cdot)$ and is largest (smallest) for a degenerate pmf $\psi(\cdot)$ that puts probability mass 1 on one of the non-zero inputs.

Proof: Omitted due to lack of space. ∎

Using above lemma, we can simplify Theorem 1 for the class of channels satisfying (23).

Corollary 2 (Adversary Stronger than Legitimate Rx): If (23) holds, then

$$r^\star(k) = \begin{cases} S_{\max}\cdot k & \text{if } k \in [0,\ k_{\text{lin}}], \\ \max_{\psi\in\mathcal{L}(k)} C_\psi & \text{if } k \in (k_{\text{lin}},\ k_{\text{sat}}), \\ C_{\text{covert}} & \text{if } k \in [k_{\text{sat}}, \infty), \end{cases} \quad (24)$$

where

$$S_{\max} \triangleq \max_{x\in\mathcal{X}\backslash\{0\}} \frac{\mathbb{D}_Y(x)}{\mathbb{D}_Z(x) - \mathbb{D}_Y(x)}; \quad (25)$$

and

$$k_{\text{lin}} \triangleq \sqrt{2}\frac{\mathbb{D}_Z(x_{\text{best}}) - \mathbb{D}_Y(x_{\text{best}})}{\sqrt{\chi^2(\delta_{x_{\text{best}}})}} \quad (26)$$

$$k_{\text{sat}} \triangleq \sqrt{2}\frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi^*(x)\cdot \mathbb{D}_Y(x)}{\sqrt{\chi^2(\psi^*)}}, \quad (27)$$

where x_{best} is the maximizer in (25); $\delta_{x_{\text{best}}}$ indicates the degenerate pmf with probability 1 at x_{best}; and ψ^* the maximizer in (20). Moreover,

$$\mathcal{L}(k) \triangleq \left\{\psi:\ \ k \geq \sqrt{2}\frac{\sum_{x\in\mathcal{X}\backslash\{0\}} \psi(x)\cdot \mathbb{D}_Y(x)}{\sqrt{\chi^2(\psi)}}\right\}. \quad (28)$$

Remark 1: From Lemma 1 and above Corollary 2, we can directly deduce that for channels satisfying (23), binary signaling on inputs 0 and x_{best} is optimal for small key rates $k \leq k_{\text{lin}}$, where recall that x_{best} is the maximizing input in (25). Moreover, in the regime of small key rates the covert capacity-key tradeoff is linear. Specifically, increasing the key-rate by 1 will increase the largest achievable rate by $S_{\max}$.

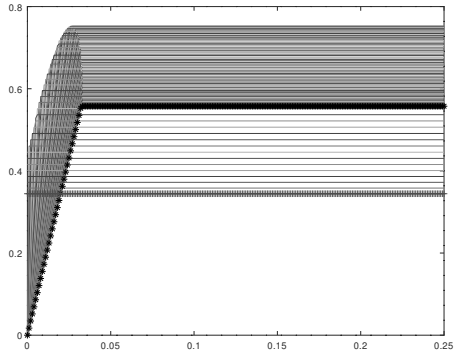

Fig. 3: Figure illustrates the functions $k \mapsto f(\psi, k)$ for different pmfs ψ. The covert capacity-key tradeoff $r^\star(k)$ corresponds to the upper convex hull of all these curves.

For channels where (23) holds, the covert capacity-key tradeoff starts at the origin: $r^\star(0) = 0$. For all other channels, i.e., when

$$\mathbb{D}_Z(x) - \mathbb{D}_Y(x) \leq 0, \qquad \text{for some } x \in \mathcal{X}\backslash\{0\}, \tag{29}$$

we have $r^\star(0) > 0$, see the following corollary directly obtained from Theorem 1 and our discussion above.

Corollary 3 (Adversary Sometimes Weaker than Legitimate Rx): If (29) holds, then

$$r^\star(k) = \max_{\psi \in \mathcal{L}(k)} \sqrt{2} \frac{\sum_{x \in \mathcal{X}\backslash\{0\}} \psi(x) \cdot \mathbb{D}_Y(x)}{\sqrt{\chi^2(\psi)}}. \tag{30}$$

In particular,

$$r^\star(k) = C_{\text{covert}}, \qquad k \geq k_{\text{sat}}. \tag{31}$$

The following example illustrates the covert capacity-key tradeoff for a channel satisfying Condition (29).

Example 2 (Adversary Sometimes Weaker than Legitimate Rx): Consider a channel with ternary inputs $\mathcal{X} = \{0, 1, 2\}$, quinary outputs $\mathcal{Y} = \mathcal{Z} = \{0, 1, 2, 3, 4\}$, and transition pmfs:

$$\Gamma_{Y|X} = \begin{bmatrix} 0.24 & 0.10 & 0.22 & 0.22 & 0.22 \\ 0.20 & 0.14 & 0.26 & 0.328 & 0.072 \\ 0.06 & 0.19 & 0.2 & 0.05 & 0.50 \end{bmatrix} \tag{32a}$$

and

$$\Gamma_{Z|X} = \begin{bmatrix} 0.32 & 0.22 & 0.23 & 0.13 & 0.10 \\ 0.47 & 0.25 & 0.10 & 0.14 & 0.04 \\ 0.38 & 0.01 & 0.15 & 0.12 & 0.34 \end{bmatrix} \tag{32b}$$

For this channel, Condition (29) holds for $x = 1$ but not for $x = 2$. Figure 3 illustrates the function $k \mapsto f(\psi, k)$ for above channel and for different choices of the input pmf $\psi(\cdot)$. The lines with $+$ and $*$ markers correspond to the degenerate choices $\psi(1) = 1$ and $\psi(2) = 1$.

IV. The Multiple-Access Channel

We turn our focus to a two-user multi-access channel (MAC) with binary inputs, see Figure 4, where the two Txs produce binary input alphabets $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$, the legitimate Rx and warden observe outputs in the finite alphabets $\mathcal{Y}$ and $\mathcal{Z}$.

A. Model

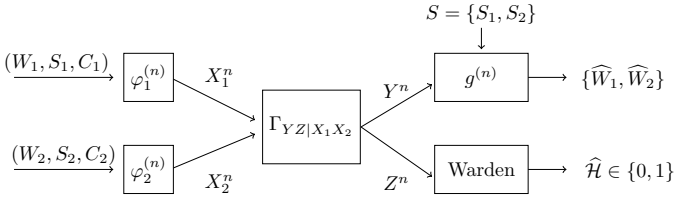

Fig. 4: Covert communication over a two-user MAC.

This covert MAC has previously been studied in [14]. The main difference here is that each Tx $j \in \{1, 2\}$ has access to additional *local* randomness described by C_i which consists of g_j i.i.d. Bernoulli-$1/2$ bits. Each Tx $j \in \{1, 2\}$ thus produces its channel inputs as

$$X_j^n = \varphi_j^{(n)}(W_j, S_j, C_j), \tag{33}$$

where W_j and S_j are independent i.i.d. Bernoulli-$1/2$ bit-strings of lengths m_j and p_j (growing with n), respectively, and S_j is know to Tx j and the Rx, while W_j and C_j only to Tx j.

As for the single-user setup, the legitimate Rx and the warden observe outputs generated by discrete memoryless channels $\Gamma_{Y|X_1X_2}(\cdot|\cdot,\cdot)$ and $\Gamma_{Z|X_1X_2}(\cdot|\cdot,\cdot)$ based on the input sequences produced at the two Txs. That means, if Tx 1 sends inputs $X_1^n = x_1^n$ and Tx 2 sends $X_2^n = x_2^n$, then for each $i \in \{1, \ldots, n\}$, the legitimate Rx observes output symbol Y_i following the conditional law $\Gamma_{Y|X_1X_2}(\cdot|x_{1,i}, x_{2,i})$ and the adversary observes output Z_i following the conditional law $\Gamma_{Z|X_1X_2}(\cdot|x_{1,i}, x_{2,i})$.

After observing outputs Y^n, the legitimate Rx decodes both messages W_1 and W_2 as:

$$(\widehat{W}_1, \widehat{W}_2) = g^{(n)}(Y^n, S_1, S_2). \tag{34}$$

using an appropriate decoding function $g^{(n)}$.

Covertness imposes that the warden's output distribution

$$\widehat{Q}^n(z^n) \triangleq \frac{\sum\limits_{\substack{(w_1,w_2,s_1,\\ s_2,c_1,c_2)}} \Gamma_{Z|X_1X_2}^{\otimes n}(z^n|x_1^n(w_1,s_1,c_1), x_2^n(w_2,s_2,c_2))}{2^{m_1+m_2} 2^{p_1+p_2} 2^{g_1+g_2}}, \tag{35}$$

be almost indistinguishable from the warden's output distribution when the all-zero sequence is transmitted by both Txs:

$$\Gamma_{Z|X_1X_2}^{\otimes n}(\cdot|0^n, 0^n). \tag{36}$$

We thus require a vanishing KL-divergence

$$\delta_n \triangleq \mathbb{D}\left(\widehat{Q}^n \middle\| \Gamma_{Z|X_1X_2}^{\otimes n}(\cdot|0^n, 0^n)\right). \tag{37}$$

Definition 2: For given $k_1, k_2 \geq 0$, define the *covert capacity-key tradeoff region* $\mathcal{R}^\star(k_1, k_2)$ as the set of all pairs (r_1, r_2) for which there exists a sequence (in n) of tuples $(m_1, m_2, p_1, p_2, g_1, g_2)$ and encoding/decoding functions $(\varphi_1^{(n)}, \varphi_2^{(n)}, g^{(n)})$ satisfying $\lim_{n\to\infty} \Pr[(\widehat{W}_1, \widehat{W}_2) \neq (W_1, W_2)] = 0$, $\lim_{n\to\infty} \delta_n = 0$, and

$$r_j \leq \liminf_{n\to\infty} \frac{m_j}{\sqrt{n\delta_n}}, \quad j \in \{1, 2\}, \tag{38}$$

$$k_j \geq \limsup_{n\to\infty} \frac{p_j}{\sqrt{n\delta_n}}, \quad j \in \{1, 2\}. \tag{39}$$

Define

$$\tilde{\mathcal{X}} \triangleq ((\mathcal{X}_1 \backslash \{0\}) \times \{0\}) \cup (\{0\} \times (\mathcal{X}_1 \backslash \{0\})). \tag{40}$$

To avoid that the problem be trivial or impossible, we restrict to DMMACs where for all $(x_1, x_2) \in \tilde{\mathcal{X}}$:

$$\sum_{\psi} \psi(x_1, x_2) \Gamma_{Z|X_1X_2}(\cdot|x_1, x_2) \neq \Gamma_{Z|X_1X_2}(\cdot|0, 0), \ \forall \psi, \tag{41a}$$

$$\text{Supp}(\Gamma_{Y|X_1X_2}(\cdot|x_1, x_2)) \subseteq \text{Supp}(\Gamma_{Y|X_1X_2}(\cdot|0, 0)) \quad \forall x_1 \in \mathcal{X}, x_2 \in \mathcal{X}_2, \tag{41b}$$

$$\text{Supp}(\Gamma_{Z|X_1X_2}(\cdot|x_1, x_2)) \subseteq \text{Supp}(\Gamma_{Z|X_1X_2}(\cdot|0, 0)), \quad \forall x_1 \in \mathcal{X}, x_2 \in \mathcal{X}_2, \tag{41c}$$

where here ψ denotes any pmf over $(\mathcal{X}_1 \times \mathcal{X}_2)\backslash\{(0, 0)\}$.

B. Results

For any input pair (x_1, x_2), define

$$\mathbb{D}_Y(x_1, x_2) = \mathbb{D}\left(\Gamma_{Y|X_1X_2}(\cdot|x_1, x_2) \,\|\, \Gamma_{Y|X_1X_2}(\cdot|0, 0)\right), \tag{42}$$

$$\mathbb{D}_Z(x_1, x_2) = \mathbb{D}\left(\Gamma_{Z|X_1X_2}(\cdot|x_1, x_2) \,\|\, \Gamma_{Z|X_1X_2}(\cdot|0, 0)\right). \tag{43}$$

Theorem 4: The covert capacity-key tradeoff region is:

$$\mathcal{C}(k_1, k_2) = \bigcup_{\rho\in[0,1]} \{0 \leq r_1 \leq r_1(\rho)\} \times \{0 \leq r_2 \leq r_2(\rho)\}$$

where

$$\rho_1(\rho) \triangleq \min\left\{\frac{\rho\mathbb{D}_Y(1, 0)}{\max\{\rho(\mathbb{D}_Z(1, 0) - \mathbb{D}_Y(1, 0)), \, 0\}} k_1, \sqrt{2}\frac{\rho\mathbb{D}_Y(1, 0)}{\sqrt{\chi^2(\rho)}}\right\} \tag{44}$$

$$\rho_2(\rho) \triangleq \min\left\{\frac{(1-\rho)\mathbb{D}_Y(0, 1)}{\max\{(1-\rho)(\mathbb{D}_Z(0, 1) - \mathbb{D}_Y(0, 1)), \, 0\}} k_2, \sqrt{2}\frac{(1-\rho)\mathbb{D}_Y(0, 1)}{\sqrt{\chi^2(\rho)}}\right\}, \tag{45}$$

and

$$\chi^2(\rho) \triangleq \sum_{z\in\mathcal{Z}} \frac{\left(\Gamma_\rho(z) - \Gamma_{Z|X_1X_2}(z|0, 0)\right)^2}{\Gamma_{Z|X_1X_2}(z|0, 0)}, \tag{46}$$

for $\Gamma_\rho(z) := \rho\Gamma_{Z|X_1X_2}(z|1, 0) + (1-\rho)\Gamma_{Z|X_1X_2}(z|0, 1)$.

Proof: Omitted due to lack of space. ∎

Example 3 (Binary Inputs with a Strong Adversary): Consider a DMMAC with binary input alphabets $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$, output alphabet $\mathcal{Y} = \mathcal{Z} = \{0, 1, 2, 3\}$, and channel transition pmfs:

$$\Gamma_{Y|X_1X_2} = \begin{bmatrix} 0.25 & 0.16 & 0.08 & 0.51 \\ 0.45 & 0.10 & 0.35 & 0.10 \\ 0.04 & 0.18 & 0.20 & 0.58 \\ 0.15 & 0.18 & 0.32 & 0.35 \end{bmatrix} \tag{47a}$$

and

$$\Gamma_{Z|X_1X_2} = \begin{bmatrix} 0.24 & 0.39 & 0.32 & 0.05 \\ 0.49 & 0.03 & 0.41 & 0.07 \\ 0.22 & 0.39 & 0.23 & 0.16 \\ 0.25 & 0.10 & 0.35 & 0.30 \end{bmatrix} \tag{47b}$$

where the four columns correspond to the four output symbols $y = 0, 1, 2, 3$ and the four rows correspond to the four possible input pairs (x_1, x_2) in increasing alphabetical ordering, i.e., $(0, 0), (0, 1), (1, 0), (1, 1)$. Figure 5 illustrates

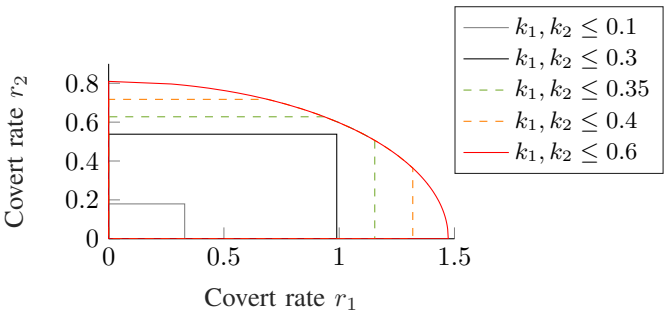

Fig. 5: $\mathcal{R}^*(k_1, k_2)$ for the channel in (47) when symmetric constraints are imposed on the key rates k_1 and k_2.

the covert capacity-key tradeoff regions for this channel at different secret-key rates. We observe that for small key rates, the tradeoff regions are square regions and there is no tradeoff between the largest covert rates r_1 and r_2 that are simultaneously achievable at the two users. For larger key rates a tradeoff arises between the two covert rates

V. Summary and Conclusion

We determined the covert capacity-key tradeoff of DMCs and binary DMMACs. For DMCs we found that the covert rate grows linearly with small key rates—but only when the adversary is uniformly stronger (in the KL-divergence sense) than the legitimate Rx. In that case, simple binary signaling (between zero and one nonzero input) is optimal. If the adversary isnt uniformly stronger or the key rate is higher, the covert rate grows sublinearly and binary signaling may no longer be optimal.

Similarly, for the DMMAC, when both users have small key budgets and the warden is uniformly stronger than the legitimate Rx, each user can independently attain its optimal covert rate. But with higher key rates or uneven adversary strength, a tradeoff emerges, and the region grows sublinearly.

Acknowledgments

The authors acknowledge funding from the ERC under Grant Agreement 101125691 and from the ANR Project CLECI.

REFERENCES

[1] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on awgn channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1921–1930, 2013.

[2] B. A. Bash, D. Goeckel, D. Towsley, and S. Guha, "Hiding information in noise: fundamental limits of covert wireless communication," *IEEE Communications Magazine*, vol. 53, no. 12, pp. 26–31, 2015.

[3] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334–2354, 2016.

[4] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493–3503, 2016.

[5] C. Bouette, L. Luzzi, and L. Wang, "Covert communication over additive-noise channels," *IEEE Transactions on Information Theory*, vol. 71, no. 3, pp. 2157–2169, 2025.

[6] P. H. Che, M. Bakshi, C. Chan, and S. Jaggi, "Reliable deniable communication with channel uncertainty," in *2014 IEEE Information Theory Workshop (ITW 2014)*, pp. 30–34, 2014.

[7] S.-H. Lee, L. Wang, A. Khisti, and G. W. Wornell, "Covert communication with channel-state information at the transmitter," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2310–2319, 2018.

[8] H. ZivariFard, M. Bloch, and A. Nosratinia, "Keyless covert communication in the presence of non-causal channel state information," in *2019 IEEE Information Theory Workshop (ITW)*, pp. 1–5, 2019.

[9] H. ZivariFard, M. R. Bloch, and A. Nosratinia, "Keyless covert communication via channel state information," *CoRR*, vol. abs/2003.03308, 2020.

[10] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Trans. Wirel. Comm.*, vol. 16, no. 9, pp. 6193–6206, 2017.

[11] O. Shmuel, A. Cohen, and O. Gurewitz, "Multi-antenna jamming in covert communication," *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4644–4658, 2021.

[12] H. ZivariFard, M. R. Bloch, and A. Nosratinia, "Covert communication via non-causal cribbing from a cooperative jammer," in *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 202–207, 2021.

[13] K. S. Kumar Arumugam and M. R. Bloch, "Embedding covert information in broadcast communications," *IEEE Trans. Inf. Forens. and Sec.*, vol. 14, no. 10, pp. 2787–2801, 2019.

[14] K. S. K. Arumugam and M. R. Bloch, "Covert communication over a k -user multiple-access channel," *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7020–7044, 2019.

[15] A. Bounhar, M. Sarkiss, and M. Wigger, "Whispering secrets in a crowd: Leveraging non-covert users for covert communications," 2024.

[16] K.-H. Cho and S.-H. Lee, "Treating interference as noise is optimal for covert communication over interference channels," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 322–332, 2021.

[17] K.-H. Cho and S.-H. Lee, "Treating interference as noise is optimal for covert communication over interference channels," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 322–332, 2021.