Covert Multi-Access Communication with a Non-Covert User

Abdelaziz Bounhar*, Mireille Sarkiss§, Michèle Wigger*

*LTCI, Télécom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France

{abdelaziz.bounhar, michele.wigger}@telecom-paris.fr

§SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France

{mireille.sarkiss}@telecom-sudparis.eu

Abstract—In this paper, we characterize the fundamental limits of a communication system with three users (i.e., three transmitters) and a single receiver where communication from two covert users must remain undetectable to an external warden. Our results show a tradeoff between the highest rates that are simultaneously achievable for the three users. They further show that the presence of a non-covert user in the system can enhance the capacities of the covert users under stringent secret-key constraints. To derive our fundamental limits, we provide an information-theoretic converse proof and present a coding scheme that achieves the performance of our converse result. Our coding scheme is based on multiplexing different code phases, which seems to be essential to exhaust the entire tradeoff region between the rates at the covert and the two non-covert users. This property is reminiscent of the setup with multiple non-covert users, where multiplexing is also required to exhaust the entire rate-region.

Index Terms—Physical Layer Security, Covert Communication, Undetectable Communication, IoT

Fig. 1: MAC setup with 2 covert users and a non-covert user in the presence of an external warden.

I. INTRODUCTION

Guaranteeing privacy and security of Internet of Things (IoT) and sensor networks is a major challenge for future wireless systems [1]. In many IoT applications, devices are resource-constrained and transmit sporadically a small number of bits while remaining silent for most of the time. Such transmissions can be secured through the paradigm of covert communication, a physical layer security technique, where users convey information without being detected by external wardens. It was shown in [2] that it is possible to communicate covertly as long as the number of communicated bits scales like $\mathcal{O}(\sqrt{n})$, for n the number of channel uses, which is compliant with IoT scenarios. Covert-rates were first characterized according to this so-called *square-root law* over AWGN channels in [2]. Several subsequent works [3]–[5], made this *square-root law* become the de-facto standard limit of covert communication for most scenarios and channels. Extensions to Broadcast Channels (BCs) and Multiple Access Channels (MACs) were proposed in [6]–[9].

This paper generalizes our previous work [9] to a Discrete Memoryless Multiple Access Channel (DM-MAC) with two covert users communicating with a legitimate receiver without being detected by an external warden, while a third non-covert user is not subject to any such covertness constraint. We establish the fundamental limits of all achievable tuples of non-covert rate, covert rates, and secret-key rates.
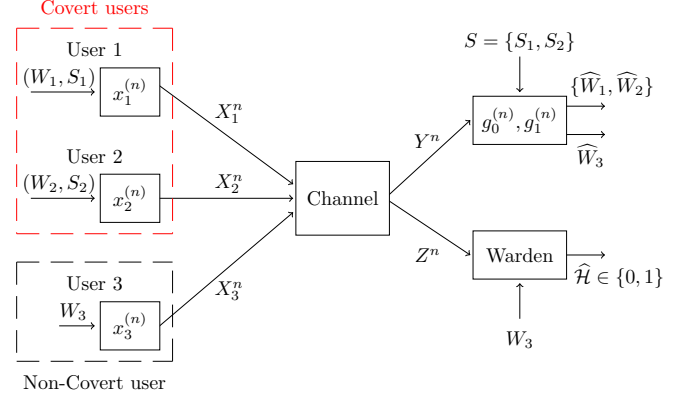
Our results show a fundamental tradeoff between the rates achievable by all users. They also confirm our previous conclusions in [9] on the fundamental role of multiplexing different code strategies to exhaust the fundamental tradeoff between the covert and non-covert rates, and on the benefits of the non-covert user to improve the covert users' capacity under a secret-key rate constraint.

II. NOTATION

We follow standard notations in [9]–[11]. In particular, we denote a random variable by X and its realization by 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 $P^{\otimes n}(x^n) = \prod_{i=1}^n P(x_i)$. For two distributions P and Q on $\mathcal{X}$, $\mathbb{D}(P\|Q) = \sum_{x \in \mathcal{X}} P(x) \log(\frac{P(x)}{Q(x)})$ denotes the Kullback-Leibler divergence between P and Q.

III. PROBLEM STATEMENT

Consider the three-user single-receiver setup in Figure 1 where an external warden should not be able to detect communication from Users 1 and 2. Communication from User 3 has no detectability constraints, and we can even allow the warden to know its transmitted message.[1] We model our setup

---

[1] Providing the warden with the message of User 3 makes the warden only stronger. The rates that are achievable under such a strong warden remain also achievable under weaker assumptions on the warden.

using two hypotheses $\mathcal{H} = 0$ and $\mathcal{H} = 1$, where under $\mathcal{H} = 0$ only User 3 is transmitting while under $\mathcal{H} = 1$ all three users transmit, and the warden wishes to guess the true hypothesis. Details are as follows. For simplicity of illustration, we assume that Users 1 and 2 produce inputs in the binary alphabets $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$ and we consider that 0 is the "off-symbol", i.e. the symbol transmitted by Users 1 and 2 under $\mathcal{H} = 0$. User 3's input alphabet $\mathcal{X}_3$ is finite but arbitrary otherwise. The legitimate receiver and the warden observe channel outputs in the finite alphabets $\mathcal{Y}$ and $\mathcal{Z}$. Define the message and key sets

$$\mathcal{M}_\ell \triangleq \{1, \ldots, \mathsf{M}_\ell\}, \quad \forall \ell \in \{1, 2, 3\}, \tag{1}$$
$$\mathcal{K}_\ell \triangleq \{1, \ldots, \mathsf{K}_\ell\}, \quad \forall \ell \in \{1, 2\}, \tag{2}$$

for given numbers M_1, M_2, M_3, K_1, and K_2 and let the messages W_1, W_2, W_3 and the keys S_1 and S_2 be independent of each other and uniform over $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{K}_1$ and $\mathcal{K}_2$, respectively. For each $\ell \in \{1, 2\}$, the key S_ℓ is known to User ℓ and to the legitimate receiver, while message W_ℓ is known to User ℓ only. In contrast W_3 is known to User 3 and is given to the warden.

Under $\mathcal{H} = 0$: Users 1 and 2 send the all-zero sequences

$$X_\ell{}^n = 0^n, \quad \forall \ell \in \{1, 2\}, \tag{3}$$

whereas User 3 applies an encoding function $x_3^{(n)} : \mathcal{M}_3 \to \mathcal{X}_3^n$ to its message W_3 and sends the resulting codeword

$$X_3^n = x_3^{(n)}(W_3) \tag{4}$$

over the channel.

Under $\mathcal{H} = 1$: For each $\ell \in \{1, 2\}$, User ℓ applies an encoding function $x_\ell{}^{(n)} : \mathcal{M}_\ell \times \mathcal{K}_\ell \to \mathcal{X}_\ell{}^n$ to its message W_ℓ and to the secret key S_ℓ and sends the resulting codeword

$$X_\ell{}^n = x_\ell{}^{(n)}(W_\ell, S_\ell), \quad \forall \ell \in \{1, 2\}, \tag{5}$$

over the channel. User 3, unaware of whether $\mathcal{H} = 0$ or $\mathcal{H} = 1$, constructs its channel inputs as in (4).

The legitimate receiver, which knows the hypothesis $\mathcal{H}$, decodes the desired messages W_3 (under $\mathcal{H} = 0$) or (W_1, W_2, W_3) (under $\mathcal{H} = 1$) based on its knowledge of the secret-keys (S_1, S_2) and its observed outputs $Y^n = (Y_1, \ldots, Y_n)$ which are generated by a discrete memoryless channel $\Gamma_{Y|X_1 X_2 X_3}$ from the input sequences X_1^n, X_2^n, X_3^n. That means, if $X_1^n = x_1^n$, $X_2^n = x_2^n$, and $X_3^n = x_3^n$ then the i-th output symbol Y_i is generated from the i-th inputs $x_{1,i}, x_{2,i}, x_{3,i}$ according to the conditional channel law $\Gamma_{Y|X_1 X_2 X_3}(\cdot|x_{1,i}, x_{2,i}, x_{3,i})$ for any $i \in \{1, \ldots, n\}$.

Under $\mathcal{H} = 0$, the decoder uses a decoding function $g_0^{(n)} : \mathcal{Y}^n \to \mathcal{M}_3$ to produce the single guess

$$\widehat{W}_3 = g_0^{(n)}(Y^n) \tag{6}$$

and under $\mathcal{H} = 1$ it uses a decoding function $g_1^{(n)} : \mathcal{Y}^n \times \mathcal{K}_1 \times \mathcal{K}_2 \to \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3$ to produce the triple of guesses

$$(\widehat{W}_1, \widehat{W}_2, \widehat{W}_3) = g_1^{(n)}(Y^n, S_1, S_2). \tag{7}$$

The decoder performance associated with a tuple of encoding and decoding functions $(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, g_0^{(n)}, g_1^{(n)})$ is measured by the error probabilities under the two hypotheses:

$$P_{e,0} \triangleq \Pr\left(\widehat{W}_3 \neq W_3 \middle| \mathcal{H} = 0\right), \tag{8}$$
$$P_{e,1} \triangleq \Pr\left(\bigcup_{\ell=1}^{3} \widehat{W}_\ell \neq W_\ell \middle| \mathcal{H} = 1\right). \tag{9}$$

On the other side, the warden observes the message W_3 and the channel outputs Z^n, which are generated from inputs X_1^n, X_2^n, X_3^n according to an arbitrary but given discrete and memoryless channel law $\Gamma_{Z|X_1 X_2 X_3}$. By the uniform nature of the messages and the secret-keys, for each $w_3 \in \mathcal{M}_3$ and $W_3 = w_3$, the warden's output distribution under $\mathcal{H} = 1$ is

$$\widehat{Q}_{\mathcal{C}, w_3}^n(z^n) \triangleq \frac{1}{\mathsf{M}_1 \mathsf{M}_2 \mathsf{K}_1 \mathsf{K}_2} \Bigg[\sum_{(w_1, s_1)} \sum_{(w_2, s_2)}$$
$$\Gamma_{Z|X_1 X_2 X_3}^{\otimes n}(z^n | x_1^n(w_1, s_1), x_2^n(w_2, s_2), x_3^n(w_3)) \Bigg], \tag{10}$$

and under $\mathcal{H} = 0$, it is

$$\Gamma_{Z|X_1 X_2 X_3}^{\otimes n}(z^n | 0^n, 0^n, x_3^n(w_3)). \tag{11}$$

For any $w_3 \in \mathcal{M}_3$, the covertness constraint at the warden is defined by means of the divergence

$$\delta_{n, w_3} \triangleq \mathbb{D}\left(\widehat{Q}_{\mathcal{C}, w_3}^n \big\| \Gamma_{Z|X_1 X_2 X_3}^{\otimes n}(\cdot | 0^n, 0^n, x_3^n(w_3))\right). \tag{12}$$

This divergence can be related to the warden's detection error probabilities by standard arguments [10, Section 11.8].

IV. CODING SCHEME, MAIN RESULT, AND NUMERICAL SIMULATIONS

A. Coding Scheme

Our coding scheme multiplexes τ different[2] phases and codes $t = 1, 2, \ldots, \tau$. The need of multiple phases stems from the multi-objective nature of our communication that not only wishes to minimize various error probabilities but also the divergence between the output distribution observed at the warden under the two hypotheses (so as to reduce the warden's detection capability). While certain phases will provide small probabilities of error, others will induce small divergences. The combined scheme over all phases then induces an optimal overall-tradeoff between small probabilities of error and small divergences.

Code construction: Our code construction has the following parameters:

- a sequence of positive numbers $\{\omega_n\}_{n \in \mathbb{N}}$ satisfying

$$\lim_{n \to \infty} \omega_n = 0, \tag{13a}$$
$$\lim_{n \to \infty} \left(\omega_n \sqrt{n} - \log n\right) = \infty; \tag{13b}$$

- a probability distribution P_T over $\{1, \ldots, \tau\}$;
- non-negative values $\{\rho_{1,t}, \rho_{2,t}\}_{t=1}^{\tau}$;

---

[2] We will see that $\tau = 6$ suffices.

- conditional probability distributions $P_{X_3|T=t}$ over $\mathcal{X}_3$, for $t = 1, \ldots, \tau$.

For any blocklength n we split the entire blocklength n into τ transmission phases $t = 1, 2, \ldots, \tau$, where the t-th phase is of length $n_t := \lfloor n \cdot P_T(t) \rfloor$.

We pick pairs of non-detectable multiple-access codes [12] $\{\mathcal{C}_{1,t}, \; \mathcal{C}_{2,t}\}$ for $t = 1, \ldots, \tau$, for the transmission of the two covert messages W_1 and W_2. By construction [9], [12], [13], codebooks $\mathcal{C}_{1,t}$ and $\mathcal{C}_{2,t}$ contain codewords $\{x_{1,t}^{n_t}(W_1, S_1)\}$ and $\{x_{2,t}^{n_t}(W_2, S_2)\}$ that depend on the respective messages as well as the corresponding secret-keys S_1 and S_2.

An important parameter of the covert-communication codes $\mathcal{C}_{1,t}$ and $\mathcal{C}_{2,t}$ is the average number of 1-symbols in the codewords. For each $t = 1, \ldots, \tau$, we choose each codebook $\mathcal{C}_{1,t}$ to consist of codewords containing approximately $\rho_{1,t}\omega_n\sqrt{n_t}$ 1-symbols and $\mathcal{C}_{2,t}$ to consist of codewords containing approximately $\rho_{2,t}\omega_n\sqrt{n_t}$ 1-symbols.

Standard (non-covert) single-user codes $\mathcal{C}_{3,t}$, for $t = 1, \ldots, \tau$ are used for the transmission of the non-covert message W_3 in the different phases. The codewords $x_{3,t}^{n_t}(W_3)$ depend only on message W_3. A key parameter of these codes is again the frequency of the various symbols in the codewords, which we call $P_{X_3|T=t}$ for codebook $\mathcal{C}_{3,t}$. It is fixed and independent of the blocklength.

Encoding: User 3 forms the concatenation of codewords

$$x_3^3(W_3) := x_{3,1}^{n_1}(W_3), \; x_{3,2}^{n_2}(W_3), \; \ldots, \; x_{3,\tau}^{n_\tau}(W_3) \qquad (14)$$

and sends the resulting string over the channel.

Under $\mathcal{H} = 0$, Users 1 and 2 send the all-zero sequences $x_1^n = 0^n$ and $x_2^n = 0^n$. Under $\mathcal{H} = 1$, Users 1 and 2 concatenate the codewords from the different codebooks

$$
\begin{aligned}
x_\ell^n(W_\ell, S_\ell) := \\
x_{\ell,1}^{n_1}(W_\ell, S_\ell), \; x_{\ell,2}^{n_2}(W_\ell, S_\ell), \; \ldots, \; x_{\ell,\tau}^{n_\tau}(W_\ell, S_\ell), \\
\forall \ell \in \{1, 2\} \qquad (15)
\end{aligned}
$$

and send the resulting strings over the channel.

The encoding process under $\mathcal{H} = 1$ is depicted in Figure 2. Notice that the number of 1-symbols varies from one user to the other and it also varies over the τ phases. In particular, the covertness constraint imposes on Users 1 and 2 to transmit a limited number of 1-symbols thereby making the codewords sparse, which is not the case for User 3.

Decoding: The legitimate receiver employs the following successive decoding procedure:

1) The receiver first decodes the non-covert message W_3 based on the entire output sequence Y^n and using all τ codebooks $\mathcal{C}_{3,1}, \ldots, \mathcal{C}_{3,\tau}$, and assuming that Users 1 and 2 send the all-zero sequences. Various decoding algorithms can be employed, for example joint typicality decoding or a maximum likelihood decoding rule based on *all the τ codebooks*.

   During this decoding steps it is assumed that both Users 1 and 2 send the all-zero codewords, irrespective of whether $\mathcal{H} = 0$ or $\mathcal{H} = 1$. In fact, even under $\mathcal{H} = 1$ this assumption will not distort the problem too much
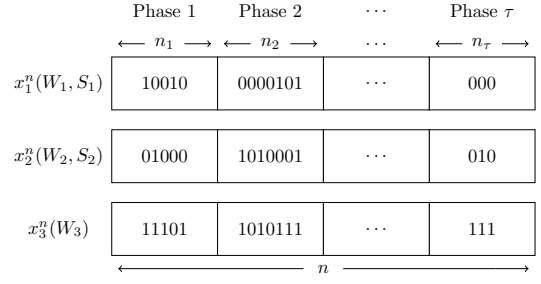
Fig. 2: Our encoding process under $\mathcal{H} = 1$ for binary input alphabets at all users. Under $\mathcal{H} = 0$, the covert users 1 and 2 send the all-zero sequence.
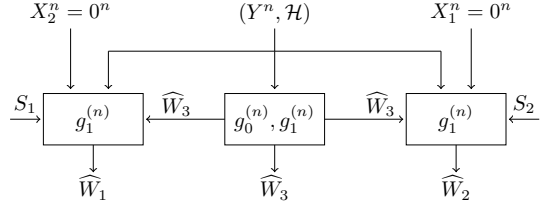
Fig. 3: Under $\mathcal{H} = 1$, the non-covert message W_3 is decoded first, followed by parallel decoding of the covert messages using an Interference as Noise scheme. Under $\mathcal{H} = 0$, we decode only W_3.

because the number of 1-symbols is small (in the order of $\omega_n\sqrt{n}$) anyways.

2) Having decoded the message W_3, the receiver proceeds with decoding either message W_1 or message W_2. Assume it starts by decoding W_1 using a standard decoding rule based on all τ covert-codes $\{\mathcal{C}_{1,t}\}$. For this decoding step the receiver can assume that $x_2^n = 0^n$. A potential decoding rule is to combine the likelihoods of the τ codewords to take a decision on the transmitted message W_1. A second alternative is to look for an index j satisfying

$$\log\left(\frac{\Gamma_{Y|X_1X_2X_3}^{\otimes n}(Y^n|x_1^n(j, S_1), 0^n, X_3^n(\widehat{W}_3))}{\Gamma_{Y|X_1X_2X_3}^{\otimes n}(Y^n|0^n, 0^n, X_3^n(\widehat{W}_3))}\right) > \eta_1, \qquad (16)$$

for a suitably chosen constant η_1. Here, $\widehat{W}_3$ denotes the receiver's guess of W_3.

If such an index exists and is unique, the receiver declares W_1 to be equal to this index. Otherwise it declares and error and stops.

3) The receiver proceeds to decode the second covert message W_2 in a way that is analogous to the decoding of message W_1, but now x_1^n (and not x_2^n) is assumed to be the all-zero codeword.

Steps 2) and 3) can be inverted or even be ran in parallel. It is however important that Step 1) is executed first. The decoding process is also depicted in Figure 3.

B. Generalization of the Coding Scheme

We propose a slight generalization of our coding scheme including two new parameters $\phi_1, \phi_2 \in (0, 1]$. In our descrip-

tion, we assume $\phi_1 \geq \phi_2$, otherwise we will switch the roles of Users 1 and 2.

In the generalized scheme, communication at Users 1 and 2 is only over a fraction ϕ_1 of each phase; during the remaining $(1 - \phi_1)$ fraction of each phase both users simply send the all-zero symbols. User 3 acts as before. Specifically, during the first ϕ_1 fraction of each phase, User 1 communicates as described in the previous section, and accordingly, the codebooks $\{C_{1,t}\}$ contain codewords of lengths $n_t \phi_1$. Instead, codebooks $\{C_{2,t}\}$ contain codewords of length $n_t \phi_2 < n_t \phi_1$. In fact, User 2 sends the appropriate codeword from these codebooks during the first $n_t \phi_2$ channel uses of each phase, and during the next $n_t(\phi_1 - \phi_2)$ channel uses it sends i.i.d. symbols $\{X_{2,i}\}$ drawn according to the same distribution as used in the code construction of the phase. This ensures a homogeneous expected divergence at the warden under the two hypotheses during the first $n_t \phi_1$ channel uses of each phase. During the last $(1 - \phi_1)$ fraction of each phase the divergence is zero because under both hypotheses Users 1 and 2 both send the all-zero symbol.

It can be shown that the described modifications yield a factor ϕ_1 for the logarithmic message size that can be reliably transmitted at User 1 and for the divergence, while they yield a factor ϕ_2 for the logarithmic message size that can be reliably sent by User 2.

C. Main Results

For ease of notation, define

$$\Gamma^Y_{x_1 x_2 x_3}(y) \triangleq \Gamma_{Y|X_1 X_2 X_3}(y \mid x_1, x_2, x_3), \quad (17)$$

$$\Gamma^Z_{x_1 x_2 x_3}(z) \triangleq \Gamma_{Z|X_1 X_2 X_3}(z \mid x_1, x_2, x_3), \quad (18)$$

and

$$D^{(1)}_Y(x_3) \triangleq \mathbb{D}\left(\Gamma^Y_{10x_3} \| \Gamma^Y_{00x_3}\right), \quad (19)$$

$$D^{(2)}_Y(x_3) \triangleq \mathbb{D}\left(\Gamma^Y_{01x_3} \| \Gamma^Y_{00x_3}\right), \quad (20)$$

$$D^{(1)}_Z(x_3) \triangleq \mathbb{D}\left(\Gamma^Z_{10x_3} \| \Gamma^Z_{00x_3}\right), \quad (21)$$

$$D^{(2)}_Z(x_3) \triangleq \mathbb{D}\left(\Gamma^Z_{01x_3} \| \Gamma^Z_{00x_3}\right), \quad (22)$$

$$D^{(\ell)}_{Z-Y}(x_3) \triangleq D^{(\ell)}_Z(x_3) - D^{(\ell)}_Y(x_3), \quad \forall \ell \in \{1, 2\}. \quad (23)$$

Also, we define for each $x_3 \in \mathcal{X}_3$ and $\rho_1, \rho_2 \geq 0$:

$$\chi^2(\rho_1, \rho_2, x_3) \triangleq \sum_{z \in \mathcal{Z}} \left[\frac{\rho_1}{\rho_1 + \rho_2} \frac{\Gamma^Z_{10x_3}(z)}{\Gamma^Z_{00x_3}(z)} \right.$$
$$\left. + \frac{\rho_2}{\rho_1 + \rho_2} \frac{\Gamma^Z_{01x_3}(z)}{\Gamma^Z_{00x_3}(z)} - 1 \right]^2. \quad (24)$$

Theorem 1: Choose an arbitrary set of

- a positive integer τ;
- a positive real number $\phi_1, \phi_2 \in [0, 1]$;
- a joint distribution $P_{X_3 T}$ over $\mathcal{X}_3 \times \{1, \ldots, \tau\}$;
- non-negative numbers $\{\rho_{1,t}, \rho_{2,t}\}_{t=1}^\tau$;
- a non-negative sequence $\{\omega_n\}_{n=1}^\infty$ satisfying (13).

For any $\epsilon > 0$, arbitrary small positive numbers $\xi_m \in (0, 1)$ for all $m \in \{1, \ldots, 6\}$, and sufficiently large blocklength n, it

is possible to find codes, $C_{1,t}, C_{2,t}, C_{3,t}$ of blocklengths $n_t = \lfloor n \cdot P_T(t) \rfloor$, for $t = 1, \ldots, \tau$, and message sizes

$$\log(\mathsf{M}_1) = \phi_1 \cdot (1 - \xi_1) \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}}\left[\rho_{1,T} D^{(1)}_Y(X_3)\right], \quad (25)$$

$$\log(\mathsf{M}_2) = \phi_2 \cdot (1 - \xi_2) \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}}\left[\rho_{2,T} D^{(2)}_Y(X_3)\right], \quad (26)$$

$$\log(\mathsf{M}_3) = (1 - \xi_3) n I(X_3; Y \mid X_1 = 0, X_2 = 0, T). \quad (27)$$

$$\log(\mathsf{K}_1) = \phi_1 \cdot (1 - \xi_4) \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}}\left[\rho_{1,T} D^{(1)}_{Z-Y}(X_3)\right], (28)$$

$$\log(\mathsf{K}_2) = \phi_2 \cdot (1 - \xi_5) \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}}\left[\rho_{2,T} D^{(2)}_{Z-Y}(X_3)\right], (29)$$

so that the encoding/decoding scheme described in the previous subsection achieves probability of error $P_e \leq \epsilon$ and average warden divergence

$$\frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3}$$

$$\leq \phi_1 (1 + \xi_6) \frac{\omega_n^2}{2} \mathbb{E}\left[(\rho_{1,T} + \rho_{2,T})^2 \cdot \chi^2(\rho_{1,T}, \rho_{2,T}, X_3)\right]. (30)$$

Proof: The proof is omitted. It follows from the schemes described in the previous sections IV-A and IV-B. ∎

The logarithmic scalings of the covert-message sizes M_1 and M_2 are at most square-root-n scalings (because ω_n vanishes), indicating that the number of covert bits that can be transmitted is only in the order of square-root-n. However, we have the usual linear-in-n behavior for the logarithmic scaling of the non-covert-message sizes. Communication of covert messages is thus of zero-rate while non-covert messages are communicated at standard positive rates. To obtain meaningful quantities, we will therefore define classical rates for User 3, while for the covert users we scale the logarithms of the message sizes by $\sqrt{n}$, and call the resulting asymptotic limits square-root rates.

The secret-key and covert-message square-root-scalings all depend on the vanishing sequence ω_n. Increasing ω_n proportionally increases the permissible covert-message size but also quadratically increases the average divergence at the warden. To eliminate this dependence, we normalize the covert-message rates and the secret-key rates by the square-root of the average warden-divergence, leading to

$$r_\ell := \lim_{n \to \infty} \frac{\log(\mathsf{M}_\ell)}{\sqrt{n \mathbb{E}_{W_3}[\delta_{n,W_3}]}}, \quad \ell \in \{1, 2\}, \quad (31)$$

$$R_3 := \lim_{n \to \infty} \frac{\log(\mathsf{M}_3)}{n} \quad (32)$$

$$k_\ell := \lim_{n \to \infty} \frac{\log(\mathsf{K}_\ell)}{\sqrt{n \mathbb{E}_{W_3}[\delta_{n,W_3}]}}, \quad \ell \in \{1, 2\}. \quad (33)$$

We then obtain the following asymptotic capacity result for our setup with mixed covert and non-covert users.

Theorem 2: There exists a sequence of encodings and decodings functions satisfying

$$\lim_{n \to \infty} P_{e,\mathcal{H}} = 0, \quad \forall \mathcal{H} \in \{0, 1\}, \quad (34a)$$

$$\lim_{n \to \infty} \delta_{n,w_3} = 0, \quad \forall w_3 \in \mathcal{M}_3, \quad (34b)$$

if, and only if, for all $\ell \in \{1, 2\}$:

$$r_\ell = \sqrt{2}\beta_\ell \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T} D_Y^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T} + \rho_{2,T})^2 \chi^2(\rho_{1,T}, \rho_{2,T}, X_3)\right]}}, \tag{35}$$

$$R_3 \leq \mathbb{I}(X_3; Y \mid X_1 = 0, X_2 = 0, T), \tag{36}$$

$$k_\ell \geq \sqrt{2}\beta_\ell \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T} D_{Z-Y}^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T} + \rho_{2,T})^2 \chi^2(\rho_{1,T}, \rho_{2,T}, X_3)\right]}}, \tag{37}$$

for some pmf $P_{X_3 T}$ over $\mathcal{X}_3 \times \{1, \ldots, 6\}$, positive parameters $\{\rho_{1,t}, \rho_{2,t}\}_{t=1}^6$, and $(\beta_1, \beta_2) \in (0, 1]^2$.

Proof: The "if"-direction follows from Theorem 1 by choosing $\phi_1 = \beta_1^2$ and $\phi_2 = \beta_1\beta_2$ when $\phi_1 \geq \phi_2$ and by setting $\phi_2 = \beta_2^2$ and $\phi_1 = \beta_1\beta_2$ otherwise. The "only if"-direction is sketched in Section V. ∎

D. Numerical examples

Consider binary input alphabets at all users, $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \{0, 1\}$, and the following channels to the legitimate receiver and the warden. Here the various rows correspond to the different triples (x_1, x_2, x_3) in lexicographic order and the columns to the six y- or z-values.

$$\Gamma_{Y|X_1 X_2 X_3} = \begin{bmatrix} 0.28 & 0.26 & 0.02 & 0.01 & 0.18 & 0.25 \\ 0.12 & 0.36 & 0.29 & 0.06 & 0.11 & 0.06 \\ 0.17 & 0.14 & 0.25 & 0.10 & 0.13 & 0.21 \\ 0.05 & 0.15 & 0.31 & 0.28 & 0.01 & 0.20 \\ 0.08 & 0.39 & 0.02 & 0.25 & 0.18 & 0.08 \\ 0.05 & 0.21 & 0.13 & 0.28 & 0.03 & 0.30 \\ 0.15 & 0.05 & 0.10 & 0.17 & 0.33 & 0.20 \\ 0.05 & 0.25 & 0.10 & 0.20 & 0.10 & 0.30 \end{bmatrix}, \tag{38}$$

$$\Gamma_{Z|X_1 X_2 X_3} = \begin{bmatrix} 0.15 & 0.11 & 0.57 & 0.01 & 0.06 & 0.10 \\ 0.15 & 0.41 & 0.12 & 0.15 & 0.06 & 0.11 \\ 0.23 & 0.02 & 0.01 & 0.48 & 0.10 & 0.16 \\ 0.14 & 0.17 & 0.21 & 0.12 & 0.24 & 0.12 \\ 0.01 & 0.12 & 0.19 & 0.15 & 0.19 & 0.34 \\ 0.10 & 0.11 & 0.15 & 0.14 & 0.18 & 0.32 \\ 0.05 & 0.15 & 0.15 & 0.20 & 0.10 & 0.35 \\ 0.10 & 0.10 & 0.27 & 0.13 & 0.20 & 0.20 \end{bmatrix}. \tag{39}$$

Figure 4 illustrates the rate-region in Theorem 2 under the additional constraint on the secret-key rates $k_1 \leq 0.8, k_2 \leq 0.8$ (solid line) and the corresponding reduced rate-region when one imposes $T = 1$ (dashed line), i.e., when in our scheme communication takes place only over a single phase. The obtained results prove that the covert capacity, i.e., achievable square-root rates, is improved when the users can communicate over different phases and using different codes in the various phases.

Figure 5 illustrates the maximum covert-user square-root rate r_2 as function of the secret-key rate k_2, i.e., when one optimizes $P_{X_3 T}$ (solid line). This is compared to scenarios where non-covert User 3 sends constant symbols $X_3 = 0$ (dashed line) or $X_3 = 1$ (dash-dotted line). The observed performance underlines that the presence of non-covert User 3 can increase the covert capacity at Users 1 and 2.

Figure 6 illustrates the rate-region in Theorem 2 at different rates r_1 for the covert User 1, showcasing the trade-off between the different users at secret-key rates $k_1 \leq 0.8$ and $k_2 \leq 0.8$.
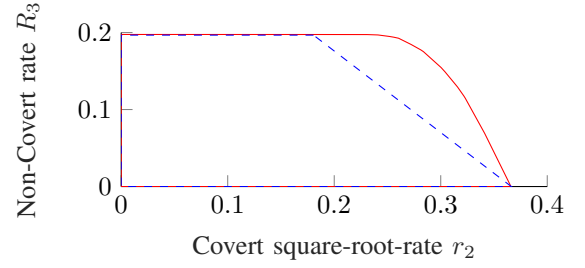

Fig. 4: Rate-region (r_2, R_3) for secret-key rates $k_1 \leq 0.8, k_2 \leq 0.8$ and $r_1 = 0.5$ (solid line) and a degenerate region when restricting to $|\mathcal{T}| = 1$ (dashed line).
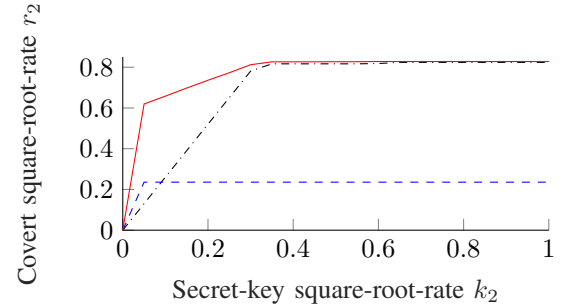

Fig. 5: Covert rate r_2 as function of secret-key rate k_2 when optimizing over $P_{X_3 T}$ (solid line) and when choosing $X_3 = 0$ or $X_3 = 1$ deterministically (dashed and dash-dotted lines) for a covert rate $r_1 = 0.1$ and a secret-key rate $k_1 \leq 0.8$.
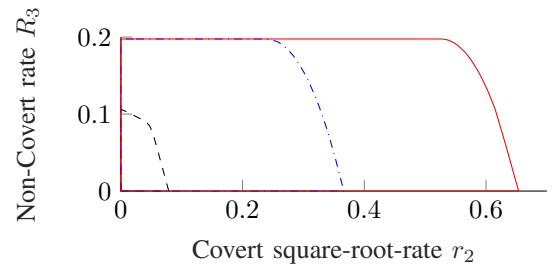

Fig. 6: Rate-region (r_2, R_3) for secret-key rates $k_1 \leq 0.8, k_2 \leq 0.8$ and different rates: $r_1 = 0.75$ (dashed line), $r_1 = 0.5$ (dash-dotted line) and $r_1 = 0.25$ (solid line).

V. PROOF OF CONVERSE TO THEOREM 2

Fix a sequence of encodings and decodings satisfying (34). Recalling the definition of $\widetilde{Q}_{\mathcal{C}, w_3}^n(z^n)$ in (10) and denoting

the i-th component of the codeword $x_3^n(w_3)$ by $x_{3,i}(w_3)$, we obtain on average over the random code and message W_3:

$$\mathbb{E}[\delta_{n,W_3}]$$

$$\overset{(a)}{=} \sum_{i=1}^{n} \mathbb{E}_{W_3}\left[\mathbb{D}\left(\widehat{Q}_{\mathcal{C},w_3}^{(i)} \| \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_{3,i}(W_3))]\right)\right] \tag{40}$$

$$\overset{(b)}{\geq} n\mathbb{E}_{P_{TX_3}}\left[\frac{(\alpha_{n,T,1}+\alpha_{n,T,2})^2}{2}\chi^2(\alpha_{n,T,1},\alpha_{n,T,2},X_{3,T})\right.$$

$$\left. + o\left(\max_{\ell\in\{1,2\}}\{\alpha_{n,T,\ell}\}\right)\right], \tag{41}$$

where T is uniform over $\{1,\ldots,n\}$. Here, (a) holds by the memoryless nature of the channel and by defining $\widehat{Q}_{\mathcal{C},w_3}^{(i)}$ as the i-th marginal of $\widehat{Q}_{\mathcal{C},w_3}^n$; (b) holds by an extension of [12, Lemma 1], and upon defining

$$\alpha_{n,i,\ell} \triangleq \frac{1}{\mathsf{M}_\ell\mathsf{K}_\ell}\sum_{w_\ell=1}^{\mathsf{M}_\ell}\sum_{s_\ell=1}^{\mathsf{K}_\ell}\mathbb{1}\{x_{\ell,i}(w_\ell,s_\ell)=1\}, \quad \ell\in\{1,2\}. \tag{42}$$

Notice that by (40) each $\alpha_{n,i,\ell} \to 0$ as $n \to \infty$. And thus by standard arguments and an extension of [12, Lemma 2]:

$$\log(\mathsf{M}_1) \leq n\mathbb{E}_{P_{TX_3}}\left[\alpha_{n,T,1}D_Y^{(1)}(X_{3,T}) + o(1)\right] + 1, \tag{43}$$

and similarly

$$\log(\mathsf{M}_2) \leq n\mathbb{E}_{P_{TX_3}}\left[\alpha_{n,T,2}D_Y^{(2)}(X_{3,T}) + o(1)\right] + 1. \tag{44}$$

Moreover for all $\ell\in\{1,2\}$,

$$\log(\mathsf{M}_\ell\mathsf{K}_\ell) \geq n\mathbb{E}_{P_{TX_3}}\left[\alpha_{n,T,\ell}D_Z^{(\ell)}(X_{3,T}) + o(1)\right] \tag{45}$$

Define next

$$\rho_{n,T,\ell} \triangleq \frac{\alpha_{n,T,\ell}}{\mathbb{E}[\alpha_{n,T,1}+\alpha_{n,T,1}]}, \quad \ell\in\{1,2\}, \tag{46}$$

and notice that by (41), (43), and (44):

$$\frac{\log(\mathsf{M}_\ell)}{\sqrt{n\mathbb{E}_{W_3}[\delta_{n,W_3}]}}$$

$$= \frac{\beta_\ell\mathbb{E}_{P_{TX_3}}\left[\rho_{n,T,\ell}D_Y^{(\ell)}(X_{3,T})\right]}{\mathbb{E}_{P_{TX_3}}\left[\frac{(\rho_{n,T,1}+\rho_{n,T,2})^2}{2}\chi^2(\rho_{n,T,1},\rho_{n,T,2},X_{3,T})\right]} + o(1),$$

$$\tag{47}$$

for some $\beta_\ell\in[0,1]$. Combined with (43)–(44), this yields:

$$\sqrt{n\mathbb{E}_{W_3}[\delta_{n,W_3}]} \leq$$

$$\frac{n}{\beta_\ell}\mathbb{E}_{P_{TX_3}}\left[\frac{(\rho_{n,T,1}+\rho_{n,T,2})^2}{2}\chi^2(\rho_{n,T,1},\rho_{n,T,2},X_{3,T})\right]$$

$$+ o(1) \tag{48}$$

which can be combined with (45) to establish that

$$\frac{\log(\mathsf{M}_\ell\mathsf{K}_\ell)}{\sqrt{n\mathbb{E}_{W_3}[\delta_{n,W_3}]}} \geq$$

$$\beta_\ell\frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{n,T,\ell}D_Z^{(\ell)}(X_{3,T})\right]}{\mathbb{E}_{P_{TX_3}}\left[\frac{(\rho_{n,T,1}+\rho_{n,T,2})^2}{2}\chi^2(\rho_{n,T,1},\rho_{n,T,2},X_{3,T})\right]} + o(1). \tag{49}$$

By standard arguments and since the probabilities that $X_{1,T}$ and $X_{2,T}$ differ from 0 vanish as $n \to \infty$:

$$\frac{1}{n}\log(\mathsf{M}_3) \leq I(X_{3,T};Y_T|X_{1,T}=0,X_{2,T}=0,T) + o(1). \tag{50}$$

Combining all arguments establishes the converse result.

VI. Summary and Discussion

This paper establishes the fundamental limits of a multi-access communication setup with two covert users and one non-covert user communicating to the same receiver in presence of a warden. Both covert users also share a common secret-key of fixed key rate with the receiver. Our results highlight that multiplexing different codes in different phases is crucial to exhaust the entire tradeoff of achievable covert and non-covert rates. Moreover, our results also show that the presence of the non-covert user can potentially improve the covert-capacity under a stringent secret-key rate constraint.

In a straightforward way, our results can also be extended to multiple users with arbitrary finite alphabets (i.e., $\mathcal{X}_1$ and $\mathcal{X}_2$ not necessarily binary).

References

[1] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A survey on iot security: Application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82 721–82 743, 2019.

[2] 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.

[3] 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.

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

[5] 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.

[6] K. S. Kumar Arumugam and M. R. Bloch, "Embedding covert information in broadcast communications," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2787–2801, 2019.

[7] D. Kibloff, S. M. Perlaza, and L. Wang, "Embedding covert information on a given broadcast code," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2169–2173.

[8] S. W. Kim and H. Q. Ta, "Covert communications over multiple overt channels," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 1112–1124, 2022.

[9] A. Bounhar, M. Sarkiss, and M. Wigger, "Mixing a covert and a non-covert user," in *2023 IEEE International Symposium on Information Theory (ISIT)*, 2023, pp. 2577–2582.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Ed.* Wiley, 2006.

[11] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.

[12] 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.

[13] 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.