Whispering Secrets in a Crowd: Leveraging Non-Covert Users for Covert Communications

Abdelaziz Bounhar, Mireille Sarkiss, Michèle Wigger

Abstract

This paper establishes the fundamental limits of a multi-access system where multiple users communicate to a legitimate receiver in presence of an external warden. Only a specific subset of the users, called covert users, needs their communication to remain undetected to the warden, while the remaining non-covert users have no such constraint. The fundamental limits show a tradeoff between the different rates that are simultaneously achievable at the various users in function of the secret-key rates that the different users share with the legitimate receiver. Interestingly, the presence of the non-covert users can enhance the capacities of the covert users, especially under stringent secret-key budgets. Our findings underscore the essential requirement of employing a multiplexing (coded time-sharing) strategy to exhaust the fundamental region of all rates that are simultaneously achievable at the covert and the non-covert users. As a side-product of our results, we also establish the covert-capacity secret-key tradeoff for standard single-user and multi-access covert communication systems (without non-covert users), i.e., the largest covert rates that are achievable under given secret-key rate budgets. Previous works had only established the minimum secret-key rates required at largest covert rates.

I. Introduction

Physical layer security leverages information-theoretic techniques and the characteristics of wireless channels to establish secure communication preventing an attacker to intercept or decipher the transmitted data. A recent technique within physical layer security is *covert communication*, which requires conveying information without being detected by adversaries (wardens), by other users, or by network monitoring nodes. Such communication setups are relevant in future IoT and sensor networks, e.g., when adversaries should not be able to detect certain monitoring activities. To maintain communication undetectable (covert), users must remain silent for most of the time, which inherently allows them to transmit only a small number of information bits. In the IoT context, such a small number of bits per device suffices for many applications, and as such, covert communication seems an adequate approach to ensure secure IoT communication. Covert communication is also inherently much more energy-efficient than conventional cryptographic algorithms, which is particularly beneficial for IoT devices with stringent battery limitations.

The early work of [2] first characterized the fundamental limits of covert communications over AWGN channels. It showed that it is possible to communicate covertly and reliably as long as the message satisfies the so-called *square-root law*, i.e., the number of communicated information bits scales like the square-root of the number of channel uses. (Recall that without covertness constraints reliable communication is possible when the number of information bits scales linearly in the number of channel uses.) Similar square-root laws were subsequently identified as the fundamental limits of covert communications for various other channels and setups [2, 3, 4, 5]. More specifically, [5] considered communication over arbitrary Discrete Memoryless Channels (DMC) and assumed the existence of an arbitrary large secret-key shared between the transmitter and the legitimate receiver. In contrast, [4] assumed the more general setup of rate-limited secret-keys. In particular, it determined the minimum secret-key rate required to communicate at the largest possible covert data-rate. In this work we strengthen this result and characterize the required key-rate for any covert-rate, not only the maximum rate. Or rather, equivalently, we characterize the largest covert data-rate that is achievable under any given key-rate budget. In all these works, covert communication takes place in the *square-root law regime*. It has been shown that rates beyond this regime are possible when the warden has uncertainty about the channel statistics [6, 7, 8, 9] or in the presence of a jammer [10, 11, 12].

Network covert communication with either multiple transmitters or multiple legitimate receivers has also been studied [13, 14, 15, 16]. In particular, [13] characterized the limits of covert communication over a single-transmitter two-receiver Broadcast Channel (BC) when the transmitter sends a common non-covert message to both receivers and a covert message to only one of them. The transmission of this covert message should not be detected by the non-intended receiver. The same communication scenario was also studied in [14] but assuming that the code used to send the common message is fixed and given and cannot be optimized to facilitate the embedding of the covert message. The BC setup with multiple legitimate receivers and an external warden was also studied from a communication-theoretic perspective [15, 16], where it was empirically shown that the detection error probability at the warden vanishes faster in the increasing number of legitimate receivers. The fundamental limits of covert communication over a multi-transmitter single-receiver Discrete Memoryless Multi-Access Channel (DMMAC)

This paper was presented in part at the 2024 IEEE International Conference on Communications [1].

were established in [17], assuming that communication from all transmitters has to remain undetected by the external warden. Their work characterized the set of all covert data-rates that are simultaneously achievable from the various transmitters to the legitimate receiver and the secret-key rates that are required to achieve the set of largest possible covert data-rates. In this manuscript, we extend these findings and determine the minimum secret-key rates that are required to attain any achievable set of covert data-rates, not only the largest data-rates. To this end, we consider a slightly more general model than in [17], where the different transmitters have access to individual local randomness.

Additionally, the current work extends the result in [17] to a scenario that mixes covert and non-covert transmissions. The covertness constraint thus imposes that the external warden cannot distinguish between the following two hypotheses:

$$\mathcal{H} = 0: \quad \text{only non-covert users transmit} \tag{1}$$

$$\mathcal{H} = 1: \quad \text{all users transmit.} \tag{2}$$

While the rates of non-covert transmissions are defined in the usual way, i.e., as

$$R_\ell = \frac{\log_2 \mathsf{M}_\ell}{n}, \tag{3}$$

for $\log_2 \mathsf{M}_\ell$ denoting the number of information bits sent by user ℓ and n the blocklength, the rates of the covert users and their secret-key rates are defined according to the square-root law and normalized by the detection capability of the warden:

$$r_\ell = \frac{\log_2 \mathsf{M}_\ell}{\sqrt{n\delta_n}} \tag{4}$$

and

$$k_\ell = \frac{\log_2 \mathsf{K}_\ell}{\sqrt{n\delta_n}} \tag{5}$$

for $\log_2 \mathsf{K}_\ell$ denoting the number of secret-key bits shared between the covert user ℓ and the legitimate receiver and δ_n (as defined precisely later) denoting an average divergence between the output distributions that the warden observes under the two hypotheses (i.e., covert users transmitting or not). Note that in [4, 5] it was shown that the definitions in (4) and (5) are meaningful in a setup with only covert users.

In this work, we determine the set of all covert, non-covert, and secret-key rates as defined in (3)–(5) that are simultaneously achievable over a given DMMAC with an external warden. In particular, we identify the rates that are simultaneously achievable without any shared secret-key at all. In our setup, we assume that the external warden has access to the non-covert messages. Our achievability result is thus even robust under such a strong assumption regarding the warden, and trivially remains valid also for less-powerful wardens.

Our fundamental tradeoff-region exhibits interesting tradeoffs between the covert and non-covert rates. In fact, for general DMMACs, the largest covert rates are only achievable under reduced non-covert rates and vice versa. Interestingly, this tradeoff even depends on the achievable key rates as we show through our theoretical findings and through numerical examples. The described tradeoffs stem from the inherent tension regarding the choice of the statistics of the inputs at the *non-covert* users: certain statistics induce DMMACs from the covert users to the legitimate receiver that allow for high covert rates, while other statistics allow for high transmission rates for the *non-covert* messages themselves. In contrast, the input statistics of the covert users do not influence the communication rates at the non-covert users, because to ensure undetectability the number of non-zero symbols has to stay low (sub-linear in the blocklength) and thus the statistics of the covert users asymptotically has no influence on the channel experienced by the non-covert transmissions.

To establish the fundamental tradeoff between the set of achievable rates, we propose a random coding scheme and an information-theoretic converse result. The coding scheme multiplexes various instances of a general scheme over multiple phases, where in each phase a different set of parameters is employed. Multiplexing allows to attain a somehow limited form of collaboration between the distributed multi-access transmitters, despite the fact that they convey independent messages. As we show, multiplexing is indeed required to exhaust the set of all achievable rates in our setup. The same holds also for non-covert communication over a DMMAC [18]. In our mixed covert/non-covert setup, multiplexing is required even when there is only a single covert and a single non-covert user. This contrasts the findings for the DMMAC when transmission from all users needs to be covert. In such case, no multiplexing is required and instead the data stream from each user can be transmitted using a standalone single-user coding scheme and in the decoding of a given covert message all other transmissions are ignored [19]. In our coding scheme that we employ in a given phase, we use a similar coding idea for the covert users, while we use a standard DMMAC coding scheme for the non-covert users. In the decoding of the non-covert users, the legitimate receiver can simply ignore the covert users, since they anyways remain silent most of the time.

Notice that the study of mixed covert users and non-covert users is novel in this line of work, and so are the results on the fundamental limits. Previous works had only considered setups with only non-covert or only covert users. However, even in the works with only covert users, the minimum secret-key rate required to achieve covertness was only determined in the special case of maximum covert transmission rates. The present work establishes the required secret-key rates for all achievable covert rates, also when they are reduced. As already mentioned, our work also presents new results for the classical single-user and multi-access covert communication scenarios as studied in [4, 17]. These new results are crucial to determine the set of covert rates that are achievable over DMCs and DMMACs under any given secret-key budgets per user, which cannot be obtained from the previous results in [4, 17].

To simplify notation, in most of this manuscript we restrict to binary input alphabets at the covert users and to two covert and a single non-covert users. All results and proofs extend however to arbitrary input alphabets and arbitrary number of covert and non-covert users in a straightforward way. For conciseness, we only present the main results for these extended setups in our manuscript. As a final extension, we also present the fundamental tradeoff between the message and secret-key rates that are simultaneously achievable in a three-transmitter and two-receiver Discrete Memoryless Interference Channel (DMIC) with an external warden where two of the transmitters send individual covert messages to their corresponding legitimate receivers and a third transmitter sends a common non-covert message to both legitimate receivers. Since in our mixed covert-/non-covert DMMAC scheme, covert messages were decoded independently, it can also be applied to this DMIC setup with same rates. Following the same arguments as in the converse proof for the DMMAC, it can then be shown that this scheme is also optimal for the proposed DMIC setup, thus establishing its fundamental limits. In a similar spirit, [19] established the covert capacity of the DMIC with only covert users based on the capacity-achieving covert DMMAC scheme and its analysis [17].

To summarize, our main contributions in this manuscript are:

- We characterize the fundamental limits of non-covert rates, covert rates, and secret-key rates that are simultaneously achievable over a DMMAC with an external warden (Theorems 1, 2, 3 and 4). The related simulation and numerical examples allow us to conclude that the presence of non-covert users can enhance the set of achievable rates of covert users.
- We establish a connection of our result to the jammer assisted covert communication (Corollary 1).
- We extend our findings to the DMIC by establishing the fundamental limits of non-covert, covert, and secret-key rates over a three-user DMIC with one non-covert and two covert users (Theorem 5).
- For single-user and DMMACs with only covert users, we establish the set of all covert rates that are achievable under a given secret-key rate budget (Corollaries 2 and 3). Previous results had only determined the required secret-key rates for maximum covert rates. Our results determine the required secret-key rates for any achievable set of covert rates. Notice that for the DMMAC (but not for single-user channels), our scheme achieving minimum secret-key rates requires local randomness at a part of the transmitters.

.

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 $\mathcal{S}$. The set $[|1,p|]$ denotes the set of positive integers between 1 and p, i.e. $[|1,p|] = \{1,2,\ldots,p\}$. 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 $P^{\otimes n}(x^n) = \prod_{i=1}^n P(x_i)$. For two distributions P and Q on $\mathcal{X}$, $\mathbb{D}(P\|Q) = \sum_{x\in\mathcal{X}} P(x)\log\left(\frac{P(x)}{Q(x)}\right)$ is the relative entropy. For two distributions P and Q on $\mathcal{X}$, we say that P is absolutely continuous with respect to (w.r.t.) Q, noted $P \ll Q$, if for all $x\in\mathcal{X}$ we have $P(x)=0$ if $Q(x)=0$, which is equivalent to the condition that the support of P is included in the support of Q. The logarithm and exponential functions are in base e and motivated by continuity of the function $t\log t$ we define $0\log(0)=0$. We use Landau notation, i.e., for a function $f(n)$ we write $f(n)=o(g(n))$ if the ratio $f(n)/g(n)$ vanishes as $n\to\infty$, and we write $f(n)=\mathcal{O}(g(n))$ if the cumulation points of the ratio $f(n)/g(n)$ are within a bounded interval.

Paper Outline: We present our main problem setup, namely a three-user Multiple Access Channel (MAC) with two covert and one non-covert users, in the following Section II. The corresponding results are presented in Section III. Section IV extends these results to arbitrary number of covert and non-covert users and arbitrary input alphabets, and to the interference channel (IC). Section V concludes the article. The technical proofs are deferred to appendices.

II. The Mixed Covert/Non-Covert Three-User MAC: Setup

Consider the setup depicted in Figure 1 where three users communicate to a legitimate receiver in the presence of a warden. Both Users 1 and 2 wish to communicate covertly, i.e., the warden should not be able to detect their communication. User 3 does not mind being detected by the warden, and we shall even assume that the warden knows its transmitted message.

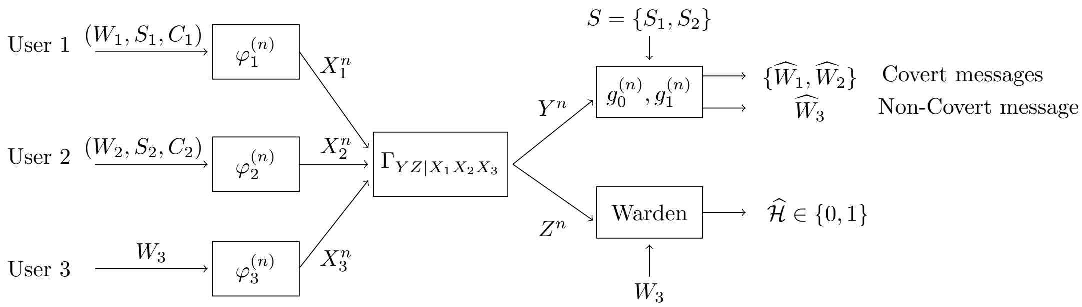


Fig. 1: Multi-access communication where communications from Users 1 and 2 have to remain undetectable to the external warden.

We thus have two hypotheses $\mathcal{H} = 0$ and $\mathcal{H} = 1$, where under $\mathcal{H} = 0$ only User 3 sends a message while under $\mathcal{H} = 1$ all users send individual messages to the legitimate receiver. For simplicity, we assume that Users 1 and 2 produce inputs in the binary alphabet $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$. 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 output alphabets $\mathcal{Y}$ and $\mathcal{Z}$. These outputs are produced by a discrete and memoryless multi-access channel (DMMAC) with transition law $\Gamma_{YZ|X_1X_2X_3}(\cdot, \cdot|\cdot, \cdot, \cdot)$. This means, if Users 1–3 send input symbols $x_{1,i}$, $x_{2,i}$, and $x_{3,i}$ in time slot i, then the legitimate receiver and the warden observe symbols Y_i and Z_i according to the pmf $\Gamma_{YZ|X_1X_2X_3}(\cdot, \cdot|x_{1,i}, x_{2,i}, x_{3,i})$, irrespective of the previously observed outputs and inputs.

Define the message, key sets, and sets of local randomness

$$\mathcal{M}_1 \triangleq \{1, \ldots, \mathsf{M}_1\}, \tag{6}$$
$$\mathcal{M}_2 \triangleq \{1, \ldots, \mathsf{M}_2\}, \tag{7}$$
$$\mathcal{M}_3 \triangleq \{1, \ldots, \mathsf{M}_3\}, \tag{8}$$
$$\mathcal{K}_1 \triangleq \{1, \ldots, \mathsf{K}_1\}, \tag{9}$$
$$\mathcal{K}_2 \triangleq \{1, \ldots, \mathsf{K}_2\}, \tag{10}$$
$$\mathcal{G}_1 \triangleq \{1, \ldots, \mathsf{G}_1\}, \tag{11}$$
$$\mathcal{G}_2 \triangleq \{1, \ldots, \mathsf{G}_2\}, \tag{12}$$

for given positive numbers M_1, M_2, M_3, K_1, K_2, G_1, and G_2, and let the messages W_1, W_2, W_3, the keys S_1 and S_2, and the local randomness C_1 and C_2 be uniform over $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{K}_1, \mathcal{K}_2, \mathcal{G}_1$, and $\mathcal{G}_2$, respectively. Secret-key S_1 is known to User 1 and to the legitimate receiver, and message W_1 and local randomness C_1 are known to User 1 only. Similarly, secret-key S_2 is known to User 2 and to the legitimate receiver and message W_2 and local randomness C_2 are known to User 2 only. In contrast, message W_3 is known to User 3 and the warden.

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

$$X_1^n = 0^n, \tag{13}$$
$$X_2^n = 0^n, \tag{14}$$

whereas User 3 applies some encoding function $\varphi_3^{(n)} \colon \mathcal{M}_3 \to \mathcal{X}_3^n$ to its message W_3 and sends the resulting codeword

$$X_3^n = \varphi_3^{(n)}(W_3) \tag{15}$$

over the channel.

Under $\mathcal{H} = 1$:
User 1 applies some encoding function $\varphi_1^{(n)} \colon \mathcal{M}_1 \times \mathcal{K}_1 \times \mathcal{G}_1 \to \mathcal{X}_1^n$ to its message W_1, its secret-key S_1 and its local randomness C_1, and sends the resulting codeword

$$X_1^n = \varphi_1^{(n)}(W_1, S_1, C_1) \tag{16}$$

over the channel. Similarly, User 2 applies some encoding function $\varphi_2^{(n)} \colon \mathcal{M}_2 \times \mathcal{K}_2 \times \mathcal{G}_2 \to \mathcal{X}_2^n$ to its message W_2, its secret-key S_2 and its local randomness C_2, and sends the resulting codeword

$$X_2^n = \varphi_2^{(n)}(W_2, S_2, C_2) \tag{17}$$

over the channel.

User 3 constructs its channel inputs in the same way as before, see (15), since it is not necessarily aware of whether $\mathcal{H}=0$ or $\mathcal{H}=1$.

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 observed sequence of outputs Y^n and its knowledge of the keys (S_1,S_2). Thus, under $\mathcal{H}=0$ it uses a decoding function $g_0^{(n)}\colon \mathcal{Y}^n \to \mathcal{M}_3$ to produce the single guess

$$\widehat{W}_3 = g_0^{(n)}(Y^n) \tag{18}$$

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

$$(\widehat{W}_1,\widehat{W}_2,\widehat{W}_3) = g_1^{(n)}(Y^n,S_1,S_2). \tag{19}$$

Decoding performance of a tuple of encoding and decoding functions $(\varphi_1^{(n)},\varphi_2^{(n)},\varphi_3^{(n)},g_0^{(n)},g_1^{(n)})$ is measured by the error probabilities under the two hypotheses:

$$P_{e,0} \triangleq \Pr\left(\widehat{W}_3 \neq W_3 \middle|\, \mathcal{H}=0\right), \tag{20}$$

$$P_{e,1} \triangleq \Pr\left(\widehat{W}_3 \neq W_3 \text{ or } \widehat{W}_2 \neq W_2 \text{ or } \widehat{W}_1 \neq W_1 \middle| \mathcal{H}=1\right). \tag{21}$$

Communication is subject to a covertness constraint at the warden, which observes the channel outputs Z^n as well as the correct message W_3. (Obviously, covertness assuming that the warden knows W_3 implies also covertness in the setup where it does not know W_3.) For a given codebook $\mathcal{C}$ and for each $w_3 \in \mathcal{M}_3$ and $W_3 = w_3$, we define the warden's output distribution under $\mathcal{H}=1$

$$\widehat{Q}^n_{\mathcal{C},w_3}(z^n) \triangleq \frac{1}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\mathsf{G}_1\mathsf{G}_2} \sum_{(w_1,s_1,c_1)} \sum_{(w_2,s_2,c_2)} \Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|x_1^n(w_1,s_1,c_1),x_2^n(w_2,s_2,c_2),x_3^n(w_3)), \tag{22}$$

and under $\mathcal{H}=0$

$$\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|0^n,0^n,x_3^n(w_3)), \tag{23}$$

and the divergence between these two distributions:

$$\delta_{n,w_3} \triangleq \mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3} \,\|\, \Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))\right), \quad \forall w_3 \in \mathcal{M}_3. \tag{24}$$

The choice of this measure for covertness is motivated by the fact that any binary hypothesis test at the warden satisfies [20] $\alpha+\beta \geq 1-\delta_{n,w_3}$, for α and β the probabilities of miss-detection and false alarm, respectively. Ensuring a negligible δ_{n,w_3} for any $w_3 \in \mathcal{M}_3$ is thus sufficient to achieve covertness irrespective of the message that is transmitted by the non-covert user.

Our main interest and focus will be on coding schemes $\{(\varphi_1^{(n)},\varphi_2^{(n)},\varphi_3^{(n)},g_0^{(n)},g_1^{(n)})\}_{n=1}^{\infty}$ that guarantee reliable transmission and undetectability at the warden in the sense of:

$$\lim_{n\to\infty} P_{e,\mathcal{H}} = 0, \qquad \mathcal{H} \in \{0,1\}, \tag{25}$$

$$\lim_{n\to\infty} \delta_{n,w_3} = 0, \qquad \forall w_3 \in \mathcal{M}_3. \tag{26}$$

Our problem is thus multi-objective as we not only wish to have reliable communication from all the users to the legitimate receiver, i.e. vanishing error probabilities (20) and (21), but also a vanishing detectability capability at the warden (24).

Remark 1 (Special Cases of Our Setup). *The setup we introduced in this section includes canonical scenarios as special cases. In fact, when the inputs X_2 and X_3 do not influence the outputs at the legitimate receiver, Users 2 and 3 can communicate reliably only when $\mathsf{M}_2=\mathsf{M}_3=1$ and the setup reduces to a covert single-user DMC (from User 1) as studied in [4, 5]. If only the input of the non-covert user X_3 does not influence the outputs, then we fall back to the two-user covert DMMAC studied in [17].*

If the non-covert user has no data to transmit, i.e., if $\mathsf{M}_3=1$, then the setup specializes to a covert two-user DMMAC with a friendly jammer (User 3) whose inputs are revealed to the legitimate receiver and the warden.

We shall formally define these special cases in our results Section III-D.

III. The Mixed Covert/Non-Covert Three User MAC: Results

We make the following assumptions to avoid degenerate conditions. For any $x_3 \in \mathcal{X}_3$:

$$\Gamma_{Y|X_1X_2X_3}(\cdot|0,1,x_3) \ll \Gamma_{Y|X_1X_2X_3}(\cdot|0,0,x_3), \tag{27a}$$
$$\Gamma_{Y|X_1X_2X_3}(\cdot|1,0,x_3) \ll \Gamma_{Y|X_1X_2X_3}(\cdot|0,0,x_3), \tag{27b}$$
$$\Gamma_{Z|X_1X_2X_3}(\cdot|0,1,x_3) \ll \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_3), \tag{27c}$$
$$\Gamma_{Z|X_1X_2X_3}(\cdot|1,0,x_3) \ll \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_3), \tag{27d}$$
$$\Gamma_{Z|X_1X_2X_3}(\cdot|0,1,x_3) \neq \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_3), \tag{27e}$$
$$\Gamma_{Z|X_1X_2X_3}(\cdot|1,0,x_3) \neq \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_3). \tag{27f}$$

Conditions (27a) and (27b) exclude the situation in which the legitimate receiver can detect certain input symbols with probability 1, in which case it has been shown that one can communicate $\mathcal{O}(\sqrt{n}\log(n))$ covert and reliable bits in a blocklength n, see [4, Appendix G]. Conditions (27c) and (27d) prevent the warden to distinguish the covert users' inputs with probability 1. Finally, (27e) and (27f) prevent that the warden's output distribution does not depend on the covert users' inputs, in which case, covertness is trivial.

A. Useful Definitions

Before presenting our results we make the following definitions. Let $\{\omega_n\}_{n=1}^{\infty}$ be a sequence satisfying

$$\lim_{n\to\infty} \omega_n = 0, \tag{28a}$$
$$\lim_{n\to\infty} \left(\omega_n\sqrt{n} - \log n\right) = \infty, \tag{28b}$$

and define

$$\alpha_n \triangleq \frac{\omega_n}{\sqrt{n}}. \tag{29}$$

We further define for any $x_3 \in \mathcal{X}_3$ the abbreviations

$$\mathbb{D}_Y^{(1)}(x_3) \triangleq \mathbb{D}\left(\Gamma_{Y|X_1X_2X_3}(\cdot|1,0,x_3) \,||\, \Gamma_{Y|X_1X_2X_3}(\cdot|0,0,x_3)\right), \tag{30a}$$
$$\mathbb{D}_Y^{(2)}(x_3) \triangleq \mathbb{D}\left(\Gamma_{Y|X_1X_2X_3}(\cdot|0,1,x_3) \,||\, \Gamma_{Y|X_1X_2X_3}(\cdot|0,0,x_3)\right), \tag{30b}$$
$$\mathbb{D}_Y^{(1,2)}(x_3) \triangleq \mathbb{D}\left(\Gamma_{Y|X_1X_2X_3}(\cdot|1,1,x_3) \,||\, \Gamma_{Y|X_1X_2X_3}(\cdot|0,0,x_3)\right), \tag{30c}$$
$$\mathbb{D}_Z^{(1)}(x_3) \triangleq \mathbb{D}\left(\Gamma_{Z|X_1X_2X_3}(\cdot|1,0,x_3) \,||\, \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_3)\right), \tag{30d}$$
$$\mathbb{D}_Z^{(2)}(x_3) \triangleq \mathbb{D}\left(\Gamma_{Z|X_1X_2X_3}(\cdot|0,1,x_3) \,||\, \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_3)\right), \tag{30e}$$
$$\mathbb{D}_Z^{(1,2)}(x_3) \triangleq \mathbb{D}\left(\Gamma_{Z|X_1X_2X_3}(\cdot|1,1,x_3) \,||\, \Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_3)\right). \tag{30f}$$

Define further the function

$$\chi^2(\rho_1,\rho_2,x_3) \triangleq \sum_{z\in\mathcal{Z}} \frac{\left(\frac{\rho_1}{\rho_1+\rho_2}\Gamma_{Z|X_1X_2X_3}(z|1,0,x_3) + \frac{\rho_2}{\rho_1+\rho_2}\,\Gamma_{Z|X_1X_2X_3}(z|0,1,x_3) - \Gamma_{Z|X_1X_2X_3}(z|0,0,x_3)\right)^2}{\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3)}. \tag{31}$$

We have $\chi^2(\rho_1,\rho_2,x_3)$ the chi-squared distance between the mixture distribution

$$\frac{\rho_1}{\rho_1+\rho_2}\Gamma_{Z|X_1X_2X_3}(\cdot|1,0,x_3) + \frac{\rho_2}{\rho_1+\rho_2}\,\Gamma_{Z|X_1X_2X_3}(\cdot|0,1,x_3), \tag{32}$$

and $\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3)$. Notice that $\chi^2(\rho_1,\rho_2,x_3) = \chi^2\left(\frac{\rho_1}{\rho_1+\rho_2}, \frac{\rho_2}{\rho_1+\rho_2}, x_3\right)$, i.e., any normalization of both ρ_1 and ρ_2 does not change the χ^2 distance.

B. A Basic Coding Scheme

Our first result in Theorem 1 is an achievability result, and we start by explaining the underlying coding scheme. In this subsection we present a special case of the scheme, the general scheme is described in the following Subsection III-C. In this special case, we have a deterministic scheme and we simply omit the local randomness in the notation; the general scheme later can be randomized for one of the users.

Fix a finite alphabet $\mathcal{T}$. Let $\{\omega_n\}$ be a sequence satisfying (28) and define $\{\alpha_n\}$ as in (29). Pick a pmf P_{TX_3} over $\mathcal{T}\times\mathcal{X}_3$ and the conditional pmfs

$$P_{X_{1,n}|T}(1 \mid t) = \rho_{1,t}\alpha_n, \quad t\in\mathcal{T}, \tag{33}$$

$$P_{X_{2,n}|T}(1 \mid t) = \rho_{2,t}\alpha_n, \quad t \in \mathcal{T}. \tag{34}$$

Define the joint pmf

$$P^{(n)}_{TX_1X_2X_3Y} \triangleq P_{TX_3}P_{X_{1,n}|T}P_{X_{2,n}|T}\Gamma_{Y|X_1X_2X_3}. \tag{35}$$

Let also $\mu_n \triangleq n^{-1/3}$.

Notice that while both pmfs $P_{X_{1,n}|T}$ and $P_{X_{2,n}|T}$ used to construct codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ depend on the blocklength n, the pmf $P_{X_3|T}$ used to construct $\mathcal{C}_3$ is independent of n.

Fix a large blocklength n and choose a type-vector $\boldsymbol{\pi} \in \{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\}^{\mathcal{T}}$ with entries summing to 1, i.e. $\|\boldsymbol{\pi}\|_1 = 1$, and so that

$$|\boldsymbol{\pi}(t) - P_T(t)| \leq \frac{1}{n}, \quad \forall t \in \mathcal{T}, \tag{36}$$

as well as $\boldsymbol{\pi}(t) = 0$ whenever $P_T(t) = 0$.

Codebook generation: Let $t^n = (t_1, \ldots, t_n)$ be any sequence of type $\boldsymbol{\pi}$, i.e., so that the empirical frequency of symbol $t \in \mathcal{T}$ in t^n equals $\boldsymbol{\pi}(t)$. The t^n-sequence acts as a multiplexing sequence that indicates which distribution to use in the construction of the different entries of the covert and non-covert codewords. As we see in the following, the i-th distribution used to construct the i-th entries of all codewords is determined by the value of the symbol t_i.

- For user 1, generate a codebook
$$\mathcal{C}_1 = \{x_1^n(1,1), \ldots, x_1^n(\mathsf{M}_1, \mathsf{K}_1)\} \tag{37}$$
by drawing the i-th entry of each codeword $x_1^n(w_1, s_1)$ according to the pmf $P_{X_{1,n}|T}(\cdot|t_i)$ independent of all other entries.
- For user 2, generate a codebook
$$\mathcal{C}_2 = \{x_2^n(1,1), \ldots, x_2^n(\mathsf{M}_2, \mathsf{K}_2)\} \tag{38}$$
by drawing the i-th entry of each codeword $x_2^n(w_2, s_2)$ according to the pmf $P_{X_{2,n}|T}(\cdot|t_i)$ independent of all other entries.
- For user 3, generate a codebook
$$\mathcal{C}_3 = \{x_3^n(1), \ldots, x_3^n(\mathsf{M}_3)\} \tag{39}$$
by drawing the i-th entry of each codeword $x_3^n(w_3)$ according to the pmf $P_{X_3|T}(\cdot|t_i)$ independent of all other entries.

Encoding and Decoding:
If $\mathcal{H} = 1$, Users 1 and 2 send the codewords $x_1^n(W_1, S_1)$ and $x_2^n(W_2, S_2)$ respectively, and if $\mathcal{H} = 0$ they send $x_1^n = 0^n$ and $x_2^n = 0^n$. User 3 sends the same codeword $x_3^n(W_3)$ under both hypotheses.

The legitimate receiver, which observes $Y^n = y^n$ and knows the secret-keys (S_1, S_2) and the true hypothesis $\mathcal{H}$, performs successive decoding starting with message W_3 followed by the messages W_1 and W_2 in case $\mathcal{H} = 1$. (The decoding procedure is also summarized in Figure 2.)

More specifically, under both hypotheses, the legitimate receiver looks for a unique index $w_3 \in \mathcal{M}_3$ satisfying

$$(t^n, x_3^n(w_3), y^n) \in \mathcal{T}^n_{\mu_n}(P_{TX_3Y}). \tag{40}$$

If such a unique index w_3 exists, the legitimate receiver sets $\widehat{W}_3 = w_3$. Otherwise it declares an error and stops.

Only under $\mathcal{H} = 1$ and after decoding the message W_3, the legitimate receiver decodes the messages of the two covert users. These two decoding steps depend on the set

$$\mathcal{A}^n_\eta \triangleq \left\{ (x_1^n, x_2^n, x_3^n, y^n) \in \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{X}_3^n \times \mathcal{Y}^n : \log\left(\frac{\Gamma^{\otimes n}_{Y|X_1X_2X_3}(y^n|x_1^n, x_2^n, x_3^n)}{\Gamma^{\otimes n}_{Y|X_1X_2X_3}(y^n|0^n, 0^n, x_3^n)} \right) \geq \eta \right\}, \tag{41}$$

where η is a given positive constant.

To decode message W_1, the legitimate receiver looks for a unique index w_1 satisfying

$$\left(x_1^n(w_1, S_1), 0^n, x_3^n(\widehat{W}_3), y^n \right) \in \mathcal{A}^n_{\eta_1}, \tag{42}$$

where η_1 is a positive constant that needs to be chosen judiciously. (Details on how to choose η_1 and later η_2 are presented when we analyze the scheme, see Equations (160) and (162) in Appendix A.) If such a unique index w_1 exists, the legitimate receiver sets $\widehat{W}_1 = w_1$. Otherwise it declares an error and stops.

Similarly, to decode message W_2, the legitimate receiver looks for a unique index w_2 satisfying

$$\left(0^n, x_2^n(w_2, S_2), x_3^n(\widehat{W}_3), y^n \right) \in \mathcal{A}^n_{\eta_2}, \tag{43}$$

for a well chosen positive constant η_2. If such a unique index w_2 exists, it sets $\widehat{W}_2 = w_2$, and it declares an error otherwise. We notice that the decoding of each covert message uses the previously decoded non-covert message, but assumes that the other covert users send the all-zero sequence. In fact, the number of non-zero symbols is small in each block, and the all-zero approximation seems not to cause any loss in optimality in terms of achievable rates.

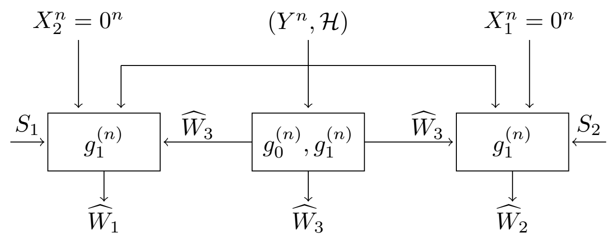


Fig. 2: Under $\mathcal{H} = 1$, the non-covert message W_3 is decoded first, followed by parallel decoding of the two covert messages. Under $\mathcal{H} = 0$ only W_3 is decoded.

C. 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 description, we assume $\phi_1 \geq \phi_2$, otherwise we 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 the time; during the remaining $(1 - \phi_1)$ fraction of time both users simply send the zero symbol. User 3 acts as before and transmits over the entire duration of the blocklength n. See Figure 3 for an illustration of the scheme.

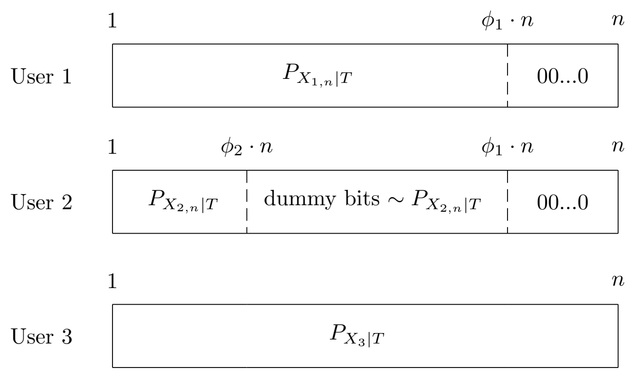


Fig. 3: Illustration of the code construction in the generalized coding scheme for the scenario $\phi_1 \geq \phi_2$.

For each $t \in \mathcal{T}$ let

$$\mathcal{L}(t) := \{j \in [|1, n|]\colon\ t_j = t\}, \tag{44}$$

and choose two disjoint subsets of time-instances $\mathcal{L}_1(t), \mathcal{L}_{1,2}(t) \subseteq \mathcal{L}(t)$ of sizes

$$|\mathcal{L}_1(t)| \approx nP_T(t)(\phi_1 - \phi_2) \tag{45}$$

$$|\mathcal{L}_{1,2}(t)| \approx nP_T(t)\phi_2. \tag{46}$$

Users 1 and 2 send the following channel inputs depending on whether a time slot i lies in the sets $\mathcal{L}_{1,2}(t)$ or $\mathcal{L}_1(t)$ for some t or not.

- In all channel uses $\cup_{t\in\mathcal{T}}\mathcal{L}_{1,2}(t)$: Users 1 and 2 transmit the corresponding symbols from the codewords $x_1^n(W_1, S_1)$ and $x_2^n(W_2, S_2)$ as described in the basic scheme.
- For each $t \in \mathcal{T}$, in channel uses $\mathcal{L}_1(t)$: User 1 transmits the corresponding symbols from the codeword $x_1^n(W_1, S_1)$ and User 2 sends i.i.d. inputs according to $P_{X_{2,n|T=t}}$. Thus, in this scheme only User 2 uses its local randomness C_2 to generate the i.i.d. inputs in channel uses $\cup_{t\in\mathcal{T}}\mathcal{L}_1(t)$.
- In the remaining channel uses (i.e., channel uses neither in $\cup_{t\in\mathcal{T}}\mathcal{L}_{1,2}(t)$ nor in $\cup_{t\in\mathcal{T}}\mathcal{L}_1(t)$), Users 1 and 2 send input 0.

The legitimate receiver decodes message W_3 as before, see (40). To decode message W_1, it applies the decoding rule in (42), but focusing only on channel uses $\cup_{t\in\mathcal{T}}(\mathcal{L}_{1,2}(t) \cup \mathcal{L}_1(t))$. To decode message W_2, it applies the decoding rule in (43), but focusing only on channel uses $\cup_{t\in\mathcal{T}}\mathcal{L}_{1,2}(t)$.

D. Main Results

Our first result is a finite blocklength achievability result based on the coding schemes described in the previous two subsections.

Theorem 1. *Fix any pmf P_{TX_3} over finite alphabets $\mathcal{T} \times \mathcal{X}_3$ and $(T, X_3) \sim P_{TX_3}$, any sequence $\{\omega_n\}_{n=1}^{\infty}$ as in (28), any pair $(\phi_1, \phi_2) \in [0,1]^2$, and any nonnegative tuple $\{(\rho_{1,t}, \rho_{2,t})\}_{t\in\mathcal{T}}$. Then, for any $\epsilon > 0$ and arbitrary small numbers $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6 > 0$ and for sufficiently large blocklengths n, we can find encoding and decoding functions $\{(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, g_0^{(n)}, g_1^{(n)})\}_n$ with message sizes $\mathsf{M}_1, \mathsf{M}_2, \mathsf{M}_3$ and secret-key sizes $\mathsf{K}_1, \mathsf{K}_2$ satisfying*

$$\log(\mathsf{M}_1) = (1 - \xi_1) \cdot \phi_1 \cdot \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}} \left[\rho_{1,T} \mathbb{D}_Y^{(1)}(X_3) \right], \tag{47}$$

$$\log(\mathsf{M}_2) = (1 - \xi_2) \cdot \phi_2 \cdot \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}} \left[\rho_{2,T} \mathbb{D}_Y^{(2)}(X_3) \right], \tag{48}$$

$$\log(\mathsf{M}_3) = (1 - \xi_3) n \mathbb{I}(X_3; Y \mid X_1 = 0, X_2 = 0, T), \tag{49}$$

$$\log(\mathsf{M}_1) + \log(\mathsf{K}_1) = (1 + \xi_4) \cdot \phi_1 \cdot \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}} \left[\rho_{1,T} \mathbb{D}_Z^{(1)}(X_3) \right], \tag{50}$$

$$\log(\mathsf{M}_2) + \log(\mathsf{K}_2) = (1 + \xi_5) \cdot \phi_2 \cdot \omega_n \sqrt{n} \mathbb{E}_{P_{TX_3}} \left[\rho_{2,T} \mathbb{D}_Z^{(2)}(X_3) \right], \tag{51}$$

and so that

$$P_{e,\mathcal{H}} \leq \epsilon, \qquad \forall \mathcal{H} \in \{0,1\} \tag{52}$$

and

$$\frac{1}{\mathsf{M}_3} \sum_{W_3=1}^{\mathsf{M}_3} \delta_{n,W_3} = (1 + \xi_6) \cdot \max(\phi_1; \phi_2) \cdot \frac{\omega_n^2}{2} \mathbb{E}_{P_{TX_3}} \left[(\rho_{1,T} + \rho_{2,T})^2 \chi^2(\rho_{1,T}, \rho_{2,T}, X_3) \right]. \tag{53}$$

(Notice that the parameter ϕ_ℓ influences only the message and key sizes of User ℓ, $\ell \in \{1,2\}$, but not of User $3-\ell$. The expected divergence at the warden depends on the term $\max(\phi_1; \phi_2)$.)

Proof: Appendix A contains the proof in the special case $\phi_1 = \phi_2 = 1$, which is obtained by analyzing the basic coding scheme in Section III-B.

The proof in the general case can be obtained by analyzing the generalized coding scheme in Section III-C. The analysis is the same as in Appendix A, up to some small modifications that allow to introduce the factors ϕ_1 and ϕ_2. We explain the modifications when $\phi_1 \geq \phi_2$, otherwise we switch the indices 1 and 2.

- When analyzing $P_{e,1,1}$ in Appendix A, restrict to channel uses in $\cup_{t\in\mathcal{T}} (\mathcal{L}_{1,2}(t) \cup \mathcal{L}_1(t))$ because only these channel uses are used for the decoding of message W_1. Since approximately a fraction ϕ_1 of the n channel uses are in $\cup_{t\in\mathcal{T}} (\mathcal{L}_{1,2}(t) \cup \mathcal{L}_1(t))$, we obtain the additional factor ϕ_1 in (47).
- When analyzing $P_{e,1,2}$, restrict to channel uses in $\cup_{t\in\mathcal{T}} \mathcal{L}_{1,2}(t)$ because only these channel uses are used for the decoding of message W_2. Since approximately a fraction ϕ_2 of the n channel uses are in $\cup_{t\in\mathcal{T}} \mathcal{L}_{1,2}(t)$, we obtain the additional factor ϕ_2 in (48).
- The details of the resolvability analysis are provided in Appendix B. The modifications for the generalized scheme allow to introduce the factor ϕ_1 for bounds (50) and (53) and the factor ϕ_2 for bounds (51) and (53).

■

Remark 2. *To achieve the performance in Theorem 1, the covert user 1 does not require access to local randomness to achieve all tuples for which $\phi_1 \geq \phi_2$. Similarly, covert user 2 does not require access to local randomness to achieve tuples for which $\phi_2 \geq \phi_1$. (Covert user 2 utilizes common randomness to generate the random inputs in channel uses $\cup_t \mathcal{L}_1(t)$).*

We observe the difference in the scalings of the logarithms of the covert-message size and the key sizes and the scaling of the logarithm of the non-covert message size. While the formers grow in the order of $\omega_n \sqrt{n}$, and thus slowlier than $\sqrt{n}$, the logarithm of the non-covert message size scales linearly in n. Communication from the non-covert user thus admits for a positive communication-rate in the traditional sense (ratio between the number of information bits and channel uses), which is not the case for the communications from the covert users.

It is further interesting to analyze the influence of the sequence $\{\omega_n\}$. The 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. Combined with the observation in the previous paragraph, we conclude that we obtain meaningful rates for the covert-messages by dividing the log of their message sizes by the square-roots of the blocklength n and the square-root of the averaged divergences. This leads to the following definition.

Definition 1. *A non-negative tuple $(r_1, r_2, R_3, k_1, k_2)$ is achievable if there exists a sequence (in the blocklength n) of tuples*[1] $(\mathsf{M}_1, \mathsf{M}_2, \mathsf{M}_3, \mathsf{K}_1, \mathsf{K}_2, \mathsf{G}_1, \mathsf{G}_2)$ *and encoding/decoding functions $(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, g_0^{(n)}, g_1^{(n)})$ satisfying*

$$\lim_{n\to\infty} P_{e,\mathcal{H}} = 0, \qquad \forall \mathcal{H} \in \{0,1\}, \tag{54}$$

[1]Not to overload notation, we did not add a superscript $^{(n)}$ to the parameters $\mathsf{M}_1, \mathsf{M}_2, \mathsf{M}_3, \mathsf{K}_1, \mathsf{K}_2$. They however all depend on the blocklength n.

$$\lim_{n\to\infty} \delta_{n,W_3} = 0, \qquad \forall W_3 \in \mathcal{M}_3, \tag{55}$$

and

$$r_\ell = \liminf_{n\to\infty} \frac{\log(\mathsf{M}_\ell)}{\sqrt{n \frac{1}{\mathsf{M}_3} \sum_{W_3=1}^{\mathsf{M}_3} \delta_{n,W_3}}}, \qquad \forall \ell \in \{1,2\}, \tag{56}$$

$$R_3 = \liminf_{n\to\infty} \frac{\log(\mathsf{M}_3)}{n}, \tag{57}$$

$$k_\ell = \limsup_{n\to\infty} \frac{\log(\mathsf{K}_\ell)}{\sqrt{n \frac{1}{\mathsf{M}_3} \sum_{W_3=1}^{\mathsf{M}_3} \delta_{n,W_3}}}, \qquad \forall \ell \in \{1,2\}. \tag{58}$$

The following theorem determines the set of all achievable rate-key tuples $(r_1, r_2, R_3, k_1, k_2)$.

Theorem 2. *A nonnegative rate-key tuple $(r_1, r_2, R_3, k_1, k_2)$ is achievable if, and only if, there exists a pmf over $\mathcal{T} \times \mathcal{X}_3$ with $\mathcal{T} = \{1, \ldots, 6\}$ and $(T, X_3) \sim P_{TX_3}$, a nonnegative tuple $\{(\rho_{1,t}, \rho_{2,t})\}_{t\in\mathcal{T}}$, and a pair $(\beta_1, \beta_2) \in [0,1]^2$ so that the following inequalities hold:*

$$r_\ell \leq \beta_\ell \sqrt{2} \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T} \mathbb{D}_Y^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T} + \rho_{2,T})^2 \cdot \chi^2(\rho_{1,T}, \rho_{2,T}, X_3)\right]}}, \qquad \forall \ell \in \{1,2\}, \tag{59}$$

$$R_3 \leq \mathbb{I}(X_3; Y \mid X_1 = 0, X_2 = 0, T), \tag{60}$$

$$k_\ell \geq \beta_\ell \sqrt{2} \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T}\left(\mathbb{D}_Z^{(\ell)}(X_3) - \mathbb{D}_Y^{(\ell)}(X_3)\right)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T} + \rho_{2,T})^2 \cdot \chi^2(\rho_{1,T}, \rho_{2,T}, X_3)\right]}}, \qquad \forall \ell \in \{1,2\}, \tag{61}$$

where recall that $\mathbb{D}_Y^{(\ell)}(\cdot)$ and $\mathbb{D}_Z^{(\ell)}(\cdot)$ are defined in (30).

Proof: The achievability proof essentially follows from Theorem 1, by setting $\phi_\ell = \beta_\ell \max(\phi_1;\phi_2)$ and taking $n \to \infty$. For details, see Appendix C. For the proof of the converse, see Appendix D. ■

Lemma 1. *The set of five-dimensional vectors $(r_1, r_2, R_3, k_1, k_2)$ satisfying Inequalities* (59)–(61) *for some choice of pmfs P_{TX_3} and nonnegative pairs $\{(\rho_{1,t}, \rho_{2,t})\}_{t\in\mathcal{T}}$ is a convex set.*

Proof: See Appendix E. ■

Remark 3. *Whenever the numerator in* (61) *is negative, no secret-key is required to establish covert communication in our setup. In particular, whenever $\mathbb{D}_Z^{(1)}(x_3) < \mathbb{D}_Y^{(1)}(x_3)$ and $\mathbb{D}_Z^{(2)}(x_3) < \mathbb{D}_Y^{(2)}(x_3)$ for all $x_3 \in \mathcal{X}_3$, Condition* (61) *is always satisfied.*

Our theorem includes various interesting special cases. For example, when User 3 has no message to transmit ($\mathsf{M}_3 = 1$ and $R_3 = 0$), its inputs act as a jamming sequence that shapes the channel and in addition is known to the warden and the legitimate receiver. We then obtain the following corollary.

Corollary 1 (User 3 acting as a Friendly Jammer)**.** *When User 3 sends no message ($\mathsf{M}_3 = 1$ and $R_3 = 0$), it acts as a friendly jammer whose inputs are known to the warden and to the legitimate receiver. In this case, a rate-key tuple (r_1, r_2, k_1, k_2) for Users 1 and 2 is achievable if, and only if,* (59) *and* (61) *hold for some choice of P_{TX_3} and pairs $\{(\rho_{1,t}, \rho_{2,t})\}_{t\in\mathcal{T}}$ and $(\beta_1, \beta_2) \in [0,1]^2$.*

Our theorem also includes results for the two-user covert DMMAC and the single-user covert DMC as specials cases. In both cases our results are stronger than the previously known findings in [17] and [4], because we not only determine the required key rate at the largest covert communication rates, but at all rates. The special case of the two-user covert DMMAC is obtained if in our setup we assume that either $\mathcal{X}_3 = \{x_3\}$ is a singleton or that the output distributions at the legitimate receiver and the warden $\Gamma_{Y|X_1X_2X_3}$ and $\Gamma_{Z|X_1X_2X_3}$ do not depend on the input X_3. Interestingly, in these cases the cardinality of $\mathcal{T}$ can be set to 1 without loss in optimality. So no multiplexing (coded time-sharing) is needed. Moreover, when $\mathcal{T} = \{t\}$, then the single parameters $\rho_{1,t}$ and $\rho_{2,t}$ can be chosen to sum to 1 without loss in optimality. These observations are made precise in the following corollary and its proof.

Corollary 2 (Only Two Covert Users)**.** *Assume that $\mathcal{X}_3 = \{x_3\}$ or that for any $x_3 \in \mathcal{X}_3$:*

$$\Gamma_{Y|X_1X_2X_3}(y|x_1, x_2, x_3) = \Gamma_{Y|X_1X_2}(y|x_1, x_2), \tag{62a}$$

$$\Gamma_{Z|X_1X_2X_3}(y|x_1,x_2,x_3) = \Gamma_{Z|X_1X_2}(z|x_1x_2). \tag{62b}$$

Then $R_3 = 0$ and a message and secret-key rates tuple (r_1, r_2, k_1, k_2) is achievable if, and only if, there exist nonnegative numbers $\rho_1, \rho_2 \geq 0$ summing to 1 ($\rho_1 + \rho_2 = 1$) and $\beta_1, \beta_2 \in [0,1]$ so that

$$r_1 \leq \beta_1\sqrt{2}\frac{\rho_1\mathbb{D}\left(\Gamma_{Y|X_1X_2}(\cdot|1,0)\,||\,\Gamma_{Y|X_1X_2}(\cdot|0,0)\right)}{\sqrt{\chi^2(\rho_1,\rho_2)}}, \tag{63}$$

$$r_2 \leq \beta_2\sqrt{2}\frac{\rho_2\mathbb{D}\left(\Gamma_{Y|X_1X_2}(\cdot|0,1)\,||\,\Gamma_{Y|X_1X_2}(\cdot|0,0)\right)}{\sqrt{\chi^2(\rho_1,\rho_2)}}, \tag{64}$$

$$k_1 \geq \beta_1\sqrt{2}\frac{\rho_1\left(\mathbb{D}\left(\Gamma_{Z|X_1X_2}(\cdot|1,0)\,||\,\Gamma_{Z|X_1X_2}(\cdot|0,0)\right) - \mathbb{D}\left(\Gamma_{Y|X_1X_2}(\cdot|1,0)\,||\,\Gamma_{Y|X_1X_2}(\cdot|0,0)\right)\right)}{\sqrt{\chi^2(\rho_1,\rho_2)}}, \tag{65}$$

$$k_2 \geq \beta_2\sqrt{2}\frac{\rho_2\left(\mathbb{D}\left(\Gamma_{Z|X_1X_2}(\cdot|0,1)\,||\,\Gamma_{Z|X_1X_2}(\cdot|0,0)\right) - \mathbb{D}\left(\Gamma_{Y|X_1X_2}(\cdot|0,1)\,||\,\Gamma_{Y|X_1X_2}(\cdot|0,0)\right)\right)}{\sqrt{\chi^2(\rho_1,\rho_2)}}, \tag{66}$$

where we use the abbreviations

$$\Gamma_{Y|X_1X_2}(y|x_1,x_2) \triangleq \Gamma_{Y|X_1X_2X_3}(y|x_1,x_2,x_3), \tag{67}$$

$$\Gamma_{Z|X_1X_2}(y|x_1,x_2) \triangleq \Gamma_{Z|X_1X_2X_3}(y|x_1,x_2,x_3), \tag{68}$$

$$\chi^2(\rho_1,\rho_2) \triangleq \sum_{z\in\mathcal{Z}} \frac{\left(\rho_1\cdot\Gamma_{Z|X_1X_2}(z|1,0) + \rho_2\cdot\Gamma_{Z|X_1X_2}(z|0,1) - \Gamma_{Z|X_1X_2}(z|0,0)\right)^2}{\Gamma_{Z|X_1X_2}(z|0,0)}. \tag{69}$$

Proof: We present the proof assuming that $\mathcal{X}_3 = \{x_3\}$ is a singleton. Under the assumptions (62) the proof is similar.

We start by proving that without loss in generality in Theorem 2 one can restrict to constant T. To this end, define $\bar{\rho}_\ell \triangleq \mathbb{E}_{P_T}[\rho_{\ell,T}]$, and notice that when X_3 is a constant x_3, the numerators of (59) and (61) simplify to $\bar{\rho}_\ell \cdot \mathbb{D}_Y^{(\ell)}(x_3)$ and $\bar{\rho}_\ell \cdot \left(\mathbb{D}_Z^{(\ell)}(x_3) - \mathbb{D}_Y^{(\ell)}(x_3)\right)$, respectively. Moreover, in this case, the denominator of (59) and (61) can be lower bounded as:

$$\begin{aligned}
&\mathbb{E}_{P_T}\left[(\rho_{1,T}+\rho_{2,T})^2\cdot\chi^2(\rho_{1,T},\rho_{2,T},x_3)\right] \\
&= \mathbb{E}_{P_T}\left(\sum_{z\in\mathcal{Z}} \frac{\left(\rho_{1,T}\Gamma_{Z|X_1X_2X_3}(z|1,0,x_3) + \rho_{2,T}\Gamma_{Z|X_1X_2X_3}(z|0,1,x_3) - (\rho_{1,T}+\rho_{2,T})\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3)\right)^2}{\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3)}\right) \\
&\geq \sum_{z\in\mathcal{Z}} \frac{\left(\bar{\rho}_1\cdot\Gamma_{Z|X_1X_2X_3}(z|1,0,x_3) + \bar{\rho}_2\cdot\Gamma_{Z|X_1X_2X_3}(z|0,1,x_3) - (\bar{\rho}_1+\bar{\rho}_2)\cdot\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3)\right)^2}{\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3)},
\end{aligned} \tag{70}$$

where the inequality holds by the convexity of the square-function and Jensen's inequality.

We conclude that replacing $\rho_{1,t}$ and $\rho_{2,t}$ by the respective expectations $\bar{\rho}_1$ and $\bar{\rho}_2$ (and thus T by a constant) does not change the numerator of the right-hand sides of (59) and (61), while it divides all the denominators by the same factor $\sqrt{\gamma} \geq 1$, for γ the ratio between the left- and right-hand sides of (70). Dividing β_ℓ by $\sqrt{\gamma}$, we can recover the same constraints on the rates and keys, from which we started. There is thus no reason to consider non-constant random variables T.

Notice further that for a single $T = t$ the rate and key expressions in Theorem 2 only depend on the normalized coefficients $\frac{\rho_{1,t}}{\rho_{1,t}+\rho_{2,t}}$ and $\frac{\rho_{2,t}}{\rho_{1,t}+\rho_{2,t}}$ but not on the absolute values of $\rho_{1,t}$ and $\rho_{2,t}$, because $\chi^2(\rho_{1,t},\rho_{2,t},x_3)$ also only depends on the ratios $\frac{\rho_{1,t}}{\rho_{1,t}+\rho_{2,t}}$ and $\frac{\rho_{2,t}}{\rho_{1,t}+\rho_{2,t}}$ but not on their absolute values. This implies that without loss in generality we can restrict to $\rho_{1,t} + \rho_{2,t} = 1$, which establishes the above corollary. ∎

Remark 4. *Corollary 2 strengthens the results in [17] for two covert users because in Corollary 2 we characterize the required secret-key rates for all achievable covert rates, not only the ones on the boundary of the achievable region. We recall that we need local randomness only at one of the users.*

In a similar way, we can recover the fundamental limits for a single-user communication system. We start from Corollary 2 (where $\rho_1 + \rho_2 = 1$) and assume that for any x_1, x_2, x_3:

$$\Gamma_{Y|X_1X_2X_3}(y|x_1,x_2,x_3) = \Gamma_{Y|X_1}(y|x_1), \tag{71a}$$

$$\Gamma_{Z|X_1X_2X_3}(y|x_1,x_2,x_3) = \Gamma_{Z|X_1}(z|x_1), \tag{71b}$$

for some conditional pmf $\Gamma_{Y|X_1}$. This immediately implies that Users 2 and 3 cannot communicate reliably, i.e., we can restrict to $r_2 = k_2 = R_3 = 0$. Moreover, under Assumption (71b), we have

$$\chi^2(\rho_1,\rho_2) = \sum_{z\in\mathcal{Z}} \frac{\left(\rho_1\cdot\Gamma_{Z|X_1}(z|1) + \rho_2\cdot\Gamma_{Z|X_1}(z|0) - (\rho_1+\rho_2)\cdot\Gamma_{Z|X_1}(z|0)\right)^2}{\Gamma_{Z|X_1}(z|0)} \tag{72}$$

$$= \sum_{z\in\mathcal{Z}} \frac{\left(\rho_1 \cdot \Gamma_{Z|X_1}(z|1) - \rho_1 \cdot \Gamma_{Z|X_1}(z|0)\right)^2}{\Gamma_{Z|X_1}(z|0)} \tag{73}$$

$$= \rho_1^2 \cdot \sum_{z\in\mathcal{Z}} \frac{\left(\Gamma_{Z|X_1}(z|1) - \Gamma_{Z|X_1}(z|0)\right)^2}{\Gamma_{Z|X_1}(z|0)}, \tag{74}$$

where in (72) we used that $\rho_1+\rho_2 = 1$. Therefore, the ρ_1-factor cancels in constraints (59) and (61) and the following corollary is obtained. Define

$$\mathbb{D}_Y \triangleq \mathbb{D}\left(\Gamma_{Y|X_1}(\cdot|1) \mid\mid \Gamma_{Y|X_1}(\cdot|0)\right), \tag{75}$$

$$\mathbb{D}_Z \triangleq \mathbb{D}\left(\Gamma_{Z|X_1}(\cdot|1) \mid\mid \Gamma_{Z|X_1}(\cdot|0)\right), \tag{76}$$

$$\chi^2 \triangleq \sum_{z\in\mathcal{Z}} \frac{\left(\Gamma_{Z|X_1}(z|1) - \Gamma_{Z|X_1}(z|0)\right)^2}{\Gamma_{Z|X_1}(z|0)}. \tag{77}$$

We have the following corollary.

Corollary 3 (Only a Single Covert User)**.** *Assume (71) and $r_2 = k_2 = R_3 = 0$. Then a message and secret-key rate pair for User 1 (r_1, k_1) is achievable if, and only if, there exists a number $\beta_1 \in [0, 1]$ so that*

$$r_1 \leq \beta_1 \sqrt{2} \frac{\mathbb{D}_Y}{\sqrt{\chi^2}}, \tag{78}$$

$$k_1 \geq \beta_1 \sqrt{2} \frac{\mathbb{D}_Z - \mathbb{D}_Y}{\sqrt{\chi^2}}. \tag{79}$$

Defining the secret-key covert-capacity tradeoff $r_1^\star(k_1)$ *as the largest rate achievable given a key-rate budget* k_1*,*

$$r_1^\star(k_1) = \max\left\{r_1 : (r_1, k_1) \text{ is achievable}\right\}, \tag{80}$$

we have:

$$r_1^\star(k_1) = \min\left\{k_1 \frac{\mathbb{D}_Y}{\max\{\mathbb{D}_Z - \mathbb{D}_Y, 0\}},\ \sqrt{2}\frac{\mathbb{D}_Y}{\sqrt{\chi^2}}\right\}. \tag{81}$$

For channels where $\mathbb{D}_Z > \mathbb{D}_Y$, the secret-key covert-capacity tradeoff is thus linearly increasing in the secret-key rate $k_1 \in \left[0, \sqrt{2}\frac{\mathbb{D}_Z - \mathbb{D}_Y}{\sqrt{\chi^2}}\right]$, and saturates to the largest covert rate for all larger secret-key rates, see Figure 4. In particular, for 1 additional key bit, one can transmit $\frac{\mathbb{D}_Y}{\mathbb{D}_Z - \mathbb{D}_Y}$ covert message bits. For channels where $\mathbb{D}_Z \leq \mathbb{D}_Y$ the covert capacity is constant and does not require any positive secret-key rate.

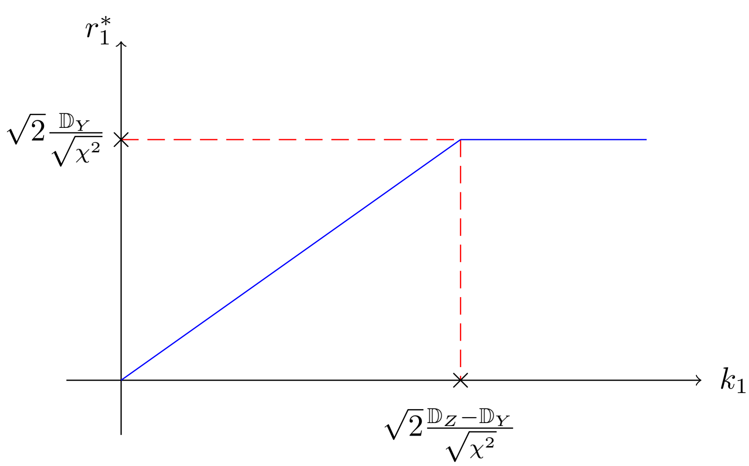


Fig. 4: The secret-key covert-capacity tradeoff for the scenario $\mathbb{D}_Z > \mathbb{D}_Y$.

Again, Corollary 3 not only recovers the result in [4] but even strengthens it, because in [4] the required secret-key rate is characterized only for covert capacity and not for achievable covert rates. In particular, based on the results in [4] it is not possible to characterize the secret-key covert capacity tradeoff $r_1^\star(k_1)$.

Remark 5. *Corollary 3 can be achieved with a deterministic scheme where the single user does not have access to additional local randomness.*

E. Numerical examples

Consider binary input alphabets for all three users, i.e. $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \{0, 1\}$, and output alphabets $\mathcal{Y} = \mathcal{Z} = \{1, \ldots, 6\}$. Assume the channel transition laws

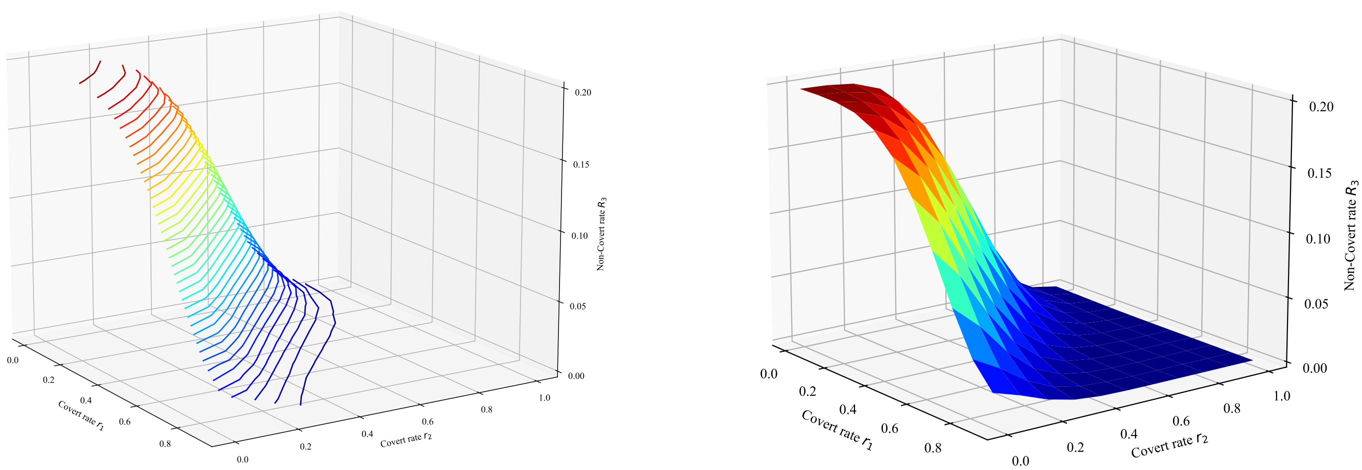
$$\Gamma_{Y|X_1X_2X_3} = \begin{bmatrix} 0.28 & 0.26 & 0.02 & 0.01 & 0.18 & 0.25 \\ 0.12 & 0.36 & 0.29 & 0.06 & 0.11 & 0.06 \\ 0.17 & 0.14 & 0.25 & 0.10 & 0.13 & 0.21 \\ 0.05 & 0.15 & 0.31 & 0.28 & 0.01 & 0.20 \\ 0.08 & 0.18 & 0.02 & 0.25 & 0.39 & 0.08 \\ 0.05 & 0.21 & 0.13 & 0.28 & 0.03 & 0.30 \\ 0.15 & 0.05 & 0.10 & 0.17 & 0.33 & 0.20 \\ 0.05 & 0.25 & 0.10 & 0.20 & 0.10 & 0.30 \end{bmatrix}, \tag{82a}$$

and

$$\Gamma_{Z|X_1X_2X_3} = \begin{bmatrix} 0.15 & 0.11 & 0.57 & 0.01 & 0.06 & 0.10 \\ 0.15 & 0.41 & 0.12 & 0.15 & 0.06 & 0.11 \\ 0.23 & 0.02 & 0.01 & 0.48 & 0.10 & 0.16 \\ 0.14 & 0.17 & 0.21 & 0.12 & 0.24 & 0.12 \\ 0.01 & 0.12 & 0.19 & 0.15 & 0.19 & 0.34 \\ 0.10 & 0.11 & 0.15 & 0.14 & 0.18 & 0.32 \\ 0.05 & 0.15 & 0.15 & 0.20 & 0.10 & 0.35 \\ 0.10 & 0.10 & 0.27 & 0.13 & 0.20 & 0.20 \end{bmatrix}, \tag{82b}$$

where the six columns correspond to the six output symbols $1, \ldots, 6$ and the eight rows correspond to the eight distinct triples (x_1, x_2, x_3) in increasing alphabetical ordering, i.e., $(0, 0, 0), (0, 0, 1), \ldots, (1, 1, 1)$. Notice that above channels satisfy Conditions (27).

Figure 5 illustrates the set of achievable triples (r_1, r_2, R_3) according to Theorem 2 for the channels in (82) and maximum secret-key rate budgets $k_1, k_2 \leq 0.8$.



(a) Contour plot of the rate-region (r_1, r_2, R_3) (b) 3D-rate region (r_1, r_2, R_3)

Fig. 5: Rate-region (r_1, r_2, R_3) for the channels in (82) and secret-key rate budgets $k_1, k_2 \leq 0.8$.

For better visualization, in Figure 6, we also show the two-dimensional tradeoff between the rates (r_2, R_3) for $k_1, k_2 \leq 0.8$ and r_1 fixed to 0.25, 0.5, or 0.75. Furthermore, in Figure 7, we plot the tradeoff between the two covert users (r_1, r_2) for secret-key rates $k_1, k_2 \leq 0.8$ at different values of $R_3 \in \{0.1965, 0.15, 0.05\}$.

It is also interesting to study the influence of the non-covert user on the set of achievable (r_2, k_2) rate-key pairs. To this end, in Figure 8 we plot the maximum achievable rate r_2 as a function of the secret-key rate k_2 without accounting for rates r_1 or r_3. We compare this maximum rate to the maximum rate achievable for deterministic inputs $X_3 = 0$ and $X_3 = 1$. We observe that the gain in optimizing over a *randomized* input X_3 achieves larger gains than a simple convex-hull operation.

Finally, in Figure 9 we show that a multiplexing (coded time-sharing) strategy is better than a single phase transmission.

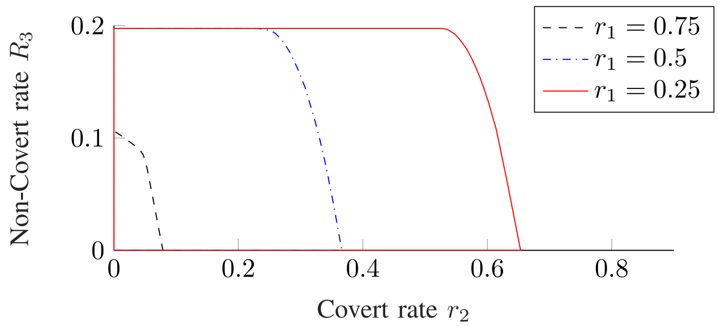


Fig. 6: Rate-region (r_2, R_3) for secret-key rates $k_1, k_2 \leq 0.8$ and different rates r_1.

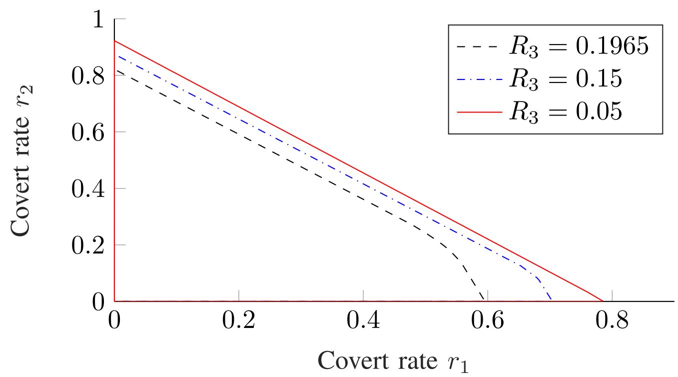


Fig. 7: Rate-region (r_1, r_2) for secret-key rate $k_1, k_2 \leq 0.8$ and different rates R_3.

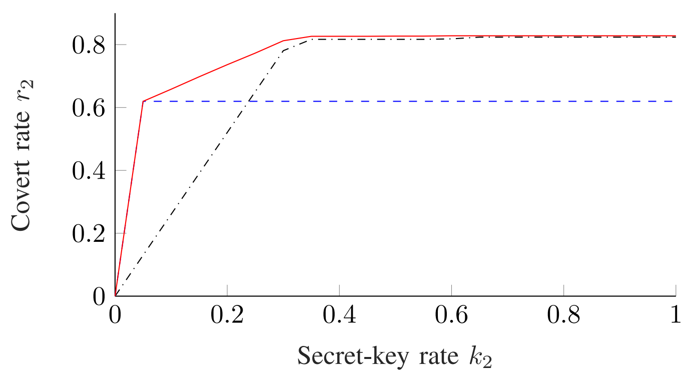


Fig. 8: Covert rate r_2 as a function of the secret-key rate k_2 when optimizing over P_{X_3T} (solid line) and when choosing $X_3 = 0$ or $X_3 = 1$ deterministically (dashed and dash-dotted lines) for a covert rate $r_1 = 0.1$ and a secret-key rate $k_1 \leq 0.8$.

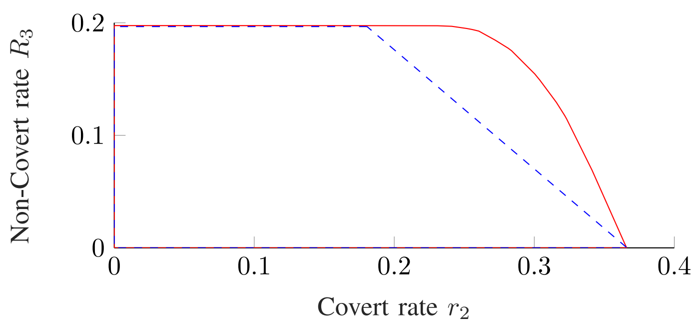


Fig. 9: Rate-region (r_2, R_3) for secret-key rates $k_1, k_2 \leq 0.8$ and $r_1 = 0.5$ in function of the allowed cardinality $|\mathcal{T}|$: we have $|\mathcal{T}| = 6$ for the solid line and a degenerate region with $|\mathcal{T}| = 1$ for the dashed line.

IV. Extensions

In this section, we broaden the scope of our findings from the multi-access scenario with 2 covert users and 1 non-covert user and with binary covert input alphabets $\{0,1\}$ to:

- arbitrary numbers of covert and non-covert users;
- arbitrary finite input alphabets at the covert users; and
- the interference channel with two covert users and one non-covert user.

A. Arbitrary Number of Covert and Non-Covert Users

Consider the setup depicted in Figure 10 where L_{c} covert users and L_{nc} non-covert users communicate individual messages $W_1,\dots,W_L$ for $L \triangleq L_{\text{c}}+L_{\text{nc}}$, to a legitimate receiver in the presence of a warden. Each message W_ℓ is uniformly distributed over the set $\mathcal{M}_\ell \triangleq \{1,\dots,\mathsf{M}_\ell\}$, $\ell \in \{1,\dots,L\}$, and independent of all other messages. Covert users can secure their transmissions at hand of secret-keys S_ℓ which are uniform over the sets $\{1,\dots,\mathsf{K}_\ell\}$ and independent of each other and the messages, and also with independent local randomness C_ℓ, which is uniform over $\{1,\dots,\mathsf{G}_\ell\}$. The legitimate receiver and the warden observe channel outputs produced by a memoryless interference channel with finite output alphabets $\mathcal{Y}$ and $\mathcal{Z}$ and transition law $\Gamma_{YZ|X_1\cdots X_L}$. Similarly to Section II, it is assumed that the messages $W_{L_{\text{c}}+1},\dots,W_L$ transmitted by the non-covert users are known to the warden.

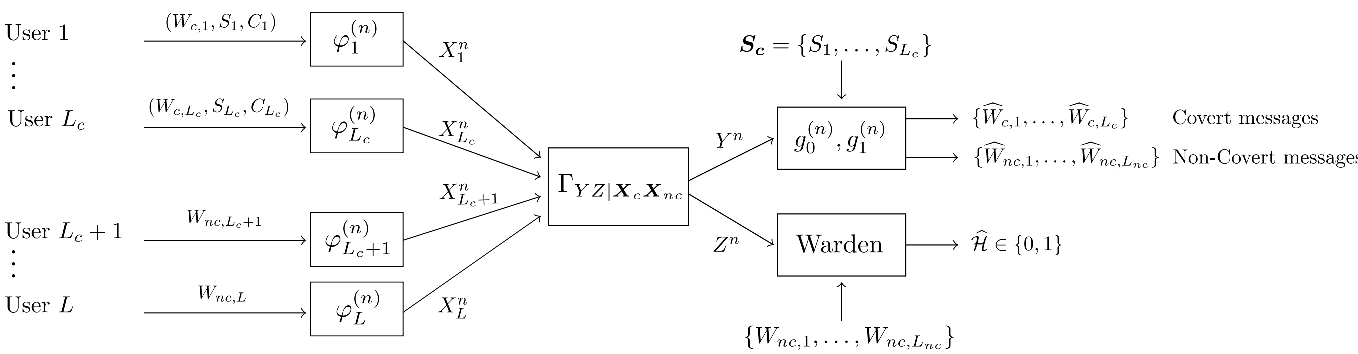


Fig. 10: Multi-access communication model with $L_{\text{c}}\geq 0$ covert users and $L_{\text{nc}}\geq 0$ non-covert users.

Each covert user $\ell=1,\dots,L_{\text{c}}$ produces its channel inputs $X_\ell^n \triangleq (X_{\ell,1},\dots,X_{\ell,n})$ in the binary input alphabet $\mathcal{X}_\ell=\{0,1\}$, while each non-covert user $\ell=L_{\text{c}}+1,\dots,L$ produces inputs X_ℓ^n pertaining to an arbitrary finite alphabet $\mathcal{X}_\ell$. Under $\mathcal{H}=0$ only the non-covert users transmit their messages:

$$X_\ell^n(W_\ell)=\varphi_\ell^{(n)}(W_\ell), \quad \ell\in\{L_{\text{c}}+1,\dots,L\}, \tag{83}$$

for some encoding function $\varphi_\ell^{(n)}(\cdot)$ on appropriate domains, while the covert users send the all-zero sequences:

$$X_\ell^n = 0^n \quad \ell\in\{1,\dots,L_{\text{c}}\}. \tag{84}$$

In this case, the output distribution induced at the warden is given by $\Gamma^{\otimes n}_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(\cdot \mid \mathbf{0}^n, \boldsymbol{X}^n_{\text{nc}}(\boldsymbol{W}_{\text{nc}}))$, where we define

$$\boldsymbol{W}_{\text{nc}} \triangleq (W_{L_{\text{c}}+1},\dots,W_L) \tag{85}$$

and where

$$\boldsymbol{X}^n_{\text{nc}}(\boldsymbol{W}_{\text{nc}}) \triangleq (X^n_{L_{\text{c}}+1}(W_{L_{\text{c}}+1}),\dots,X^n_L(W_L)). \tag{86}$$

Under $\mathcal{H}=1$ all users communicate. That means, (83) continues to hold, but (84) has to be replaced by

$$X_\ell^n(W_\ell,S_\ell)=\varphi_\ell^{(n)}(W_\ell,S_\ell,C_\ell), \quad \ell\in\{1,\dots,L_{\text{c}}\}, \tag{87}$$

for appropriate encoding functions. Thus, under $\mathcal{H}=1$, the output distribution induced at the warden is given by

$$\widehat{Q}^n_{\mathcal{C},\boldsymbol{w}_{\text{nc}}}(z^n) \triangleq \frac{1}{\prod_{\ell=1}^{L_{\text{c}}} \mathsf{M}_\ell\mathsf{K}_\ell\mathsf{C}_\ell} \sum_{\boldsymbol{w}_{\text{c}},\boldsymbol{s}_{\text{c}},\boldsymbol{c}_{\text{c}}} \Gamma^{\otimes n}_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}\left(z^n|\boldsymbol{X}^n_{\text{c}}(\boldsymbol{w}_{\text{c}},\boldsymbol{s}_{\text{c}}),\boldsymbol{X}^n_{\text{nc}}(\boldsymbol{w}_{\text{nc}})\right). \tag{88}$$

where we define $\boldsymbol{w}_{\text{c}} \triangleq (w_1,\dots,w_{L_{\text{c}}})$, $\boldsymbol{s}_{\text{c}} \triangleq (s_1,\dots,s_{L_{\text{c}}})$, and $\boldsymbol{c}_{\text{c}} \triangleq (c_1,\dots,c_{L_{\text{c}}})$, and

$$\boldsymbol{X}^n_{\text{c}}(\boldsymbol{w}_{\text{c}},\boldsymbol{s}_{\text{c}},\boldsymbol{c}_{\text{c}}) \triangleq (X_1^n(w_1,s_1,c_1),\dots,X^n_{L_{\text{c}}}(w_{L_{\text{c}}},s_{L_{\text{c}}},c_{L_{\text{c}}})). \tag{89}$$

The error probabilities under the two hypotheses are defined as:

$$P_{e,0} \triangleq \Pr\left(\bigcup_{\ell=L_c+1,\ldots,L} \widehat{W}_\ell \neq W_\ell \,\middle|\, \mathcal{H}=0 \right), \tag{90}$$

$$P_{e,1} \triangleq \Pr\left(\bigcup_{\ell=1,\ldots,L} \widehat{W}_\ell \neq W_\ell \,\middle|\, \mathcal{H}=1 \right). \tag{91}$$

For a given non-covert messages vector $\boldsymbol{w}_{\text{nc}}$, covertness at the warden is measured by the divergence

$$\delta_{n,\boldsymbol{w}_{\text{nc}}} \triangleq \mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},\boldsymbol{w}_{\text{nc}}} \,\middle\|\, \Gamma^{\otimes n}_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(z^n \mid \mathbf{0}^n, \boldsymbol{X}^n_{\text{nc}}(\boldsymbol{w}_{\text{nc}})) \right). \tag{92}$$

Definition 2. *A tuple $(r_1,\ldots,r_{L_c},R_{L_c+1},\ldots,R_L,k_1,\ldots,k_{L_c})$ is achievable if there exists a sequence (in the blocklength n) of tuples $(\mathsf{M}_1,\ldots,\mathsf{M}_L,\mathsf{K}_1,\ldots,\mathsf{K}_{L_c},\mathsf{G}_1,\ldots,\mathsf{G}_{L_c})$ and encoding/decoding functions $\{(\varphi_1^{(n)},\ldots,\varphi_L^{(n)},g_0^{(n)},g_1^{(n)})\}$ satisfying*

$$\lim_{n\to\infty} P_{e,\mathcal{H}} = 0, \qquad \forall \mathcal{H}\in\{0,1\} \tag{93}$$

$$\lim_{n\to\infty} \delta_{n,\boldsymbol{W}_{\text{nc}}} = 0, \qquad \forall \boldsymbol{W}_{\text{nc}} \in \mathcal{M}_{L_c+1}\times\ldots\times\mathcal{M}_L \tag{94}$$

and

$$r_\ell = \liminf_{n\to\infty} \frac{\log(\mathsf{M}_\ell)}{\sqrt{n\mathbb{E}_{\boldsymbol{W}_{\text{nc}}}\left[\delta_{n,\boldsymbol{W}_{\text{nc}}}\right]}}, \qquad \ell\in\{1,\ldots,L_c\}, \tag{95}$$

$$\tag{96}$$

$$R_\ell = \liminf_{n\to\infty} \frac{\log(\mathsf{M}_\ell)}{n}, \qquad \ell\in\{L_c+1,\ldots,L\}, \tag{97}$$

$$k_\ell = \limsup_{n\to\infty} \frac{\log(\mathsf{K}_\ell)}{\sqrt{n\mathbb{E}_{\boldsymbol{W}_{\text{nc}}}\left[\delta_{n,\boldsymbol{W}_{\text{nc}}}\right]}}, \qquad \ell\in\{1,\ldots,L_c\}, \tag{98}$$

where $\boldsymbol{W}_{\text{nc}}$ is uniformly distributed over the set of all possible vectors $\boldsymbol{w}_{\text{nc}}$.

For conciseness, and similarly to (30a), (30b), (30d), (30e), for any $\ell\in\{1,\ldots,L_c\}$, and tuple $\boldsymbol{x}_{\text{nc}}\in\mathcal{X}_{L_c+1}\times\ldots\times\mathcal{X}_L$, we define the abbreviations

$$\mathbb{D}_Y^{(\ell)}(\boldsymbol{x}_{\text{nc}}) = \mathbb{D}\left(\Gamma_{Y|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(\cdot\mid \boldsymbol{e}_\ell,\boldsymbol{x}_{\text{nc}}) \| \Gamma_{Y|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(\cdot\mid \mathbf{0},\boldsymbol{x}_{\text{nc}})\right), \tag{99}$$

$$\mathbb{D}_Z^{(\ell)}(\boldsymbol{x}_{\text{nc}}) = \mathbb{D}\left(\Gamma_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(\cdot\mid \boldsymbol{e}_\ell,\boldsymbol{x}_{\text{nc}}) \| \Gamma_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(\cdot\mid \mathbf{0},\boldsymbol{x}_{\text{nc}})\right), \tag{100}$$

where $\boldsymbol{e}_\ell$ denotes the ℓ-th canonical basis vector $\boldsymbol{e}_\ell=(0,\ldots,0,1,0,\ldots,0)$ of dimension L_c. Redefine also the set $\mathcal{T}\triangleq\{1,\ldots,L+L_c+1\}$, and for any nonnegative tuple $\boldsymbol{\rho}=(\rho_1,\ldots,\rho_L)$ and $\boldsymbol{x}_{\text{nc}}\in\mathcal{X}_{L_c+1}\times\cdots\mathcal{X}_L$, define (similarly to (31)):

$$\chi^2(\boldsymbol{\rho},\boldsymbol{x}_{\text{nc}}) \triangleq \sum_{z\in\mathcal{Z}} \frac{\left(\sum_{\ell=1}^{L_c}\frac{\rho_\ell}{\|\boldsymbol{\rho}\|_1}\Gamma_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(z\mid\boldsymbol{e}_\ell,\boldsymbol{x}_{\text{nc}}) - \Gamma_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(z\mid\mathbf{0},\boldsymbol{x}_{\text{nc}})\right)^2}{\Gamma_{Z|\boldsymbol{X}_{\text{c}}\boldsymbol{X}_{\text{nc}}}(z\mid\mathbf{0},\boldsymbol{x}_{\text{nc}})}. \tag{101}$$

Theorem 3. *A message and secret-key rate tuple $(r_1,\ldots,r_{L_c},R_{L_c+1},\ldots,R_L,k_1,\ldots,k_{L_c})$ is achievable if, and only if, there exists a random variable T over $\mathcal{T}$ and a random tuple $\boldsymbol{X}_{\text{nc}}\triangleq(X_{L_c+1},\ldots,X_L)$ over $\mathcal{X}_{L_c+1}\times\cdots\times\mathcal{X}_L$ of joint pmf $P_{T\boldsymbol{X}_{\text{nc}}}$, as well as a set of nonnegative tuples $\{\boldsymbol{\rho}_t\triangleq(\rho_{1,t},\ldots,\rho_{L_c,t})\}_{t\in\mathcal{T}}$ and numbers $\beta_1,\ldots,\beta_{L_c}\in[0,1]$, so that the following inequalities hold:*

$$r_\ell \le \beta_\ell\sqrt{2}\frac{\mathbb{E}_{P_{T\boldsymbol{X}_{\text{nc}}}}\left[\rho_{\ell,T}\mathbb{D}_Y^{(\ell)}(\boldsymbol{X}_{\text{nc}})\right]}{\sqrt{\mathbb{E}_{P_{T\boldsymbol{X}_{\text{nc}}}}\left[\|\boldsymbol{\rho}_T\|_1^2\cdot\chi^2(\boldsymbol{\rho}_T,\boldsymbol{X}_{\text{nc}})\right]}}, \quad \forall\ell\in\{1,\ldots,L_c\}, \tag{102}$$

$$\sum_{j\in\mathcal{J}} R_j \le I(\boldsymbol{X}_{\text{nc},\mathcal{J}};Y\mid\boldsymbol{X}_{\text{c}}=0,\boldsymbol{X}_{nc,\mathcal{J}^c},T), \quad \forall\mathcal{J}\subseteq\{L_{c+1},\ldots,L\}, \tag{103}$$

$$k_\ell \ge \beta_\ell\sqrt{2}\frac{\mathbb{E}_{P_{T\boldsymbol{X}_{\text{nc}}}}\left[\rho_{\ell,T}\left(\mathbb{D}_Z^{(\ell)}(\boldsymbol{X}_{\text{nc}})-\mathbb{D}_Y^{(\ell)}(\boldsymbol{X}_{\text{nc}})\right)\right]}{\sqrt{\mathbb{E}_{P_{T\boldsymbol{X}_{\text{nc}}}}\left[\|\boldsymbol{\rho}_T\|_1^2\cdot\chi^2(\boldsymbol{\rho}_T,\boldsymbol{X}_{nc})\right]}}, \quad \forall\ell\in\{1,\ldots,L_c\}. \tag{104}$$

where for any subset of non-covert users $\mathcal{J}\subseteq\{L_{c+1},\ldots,L\}$:

$$\boldsymbol{X}_{\text{nc},\mathcal{J}} = \{X_\ell : \ell\in\mathcal{J}\}. \tag{105}$$

Proof: A straightforward extension of the proof of Theorem 2 and omitted. ■

Lemma 2. *The set of* $(L+L_{\rm c})$*-dimensional vectors* $(r_1,\ldots,r_{L_{\rm c}},R_{L_{\rm c}+1},\ldots,R_L,k_1,\ldots,k_{L_{\rm c}})$ *satisfying Inequalities* (102) − (104) *for some choice of pmfs* $P_{T\boldsymbol{X}_{nc}}$*, tuples* $\{\boldsymbol{\rho}_t\}_{t\in\mathcal{T}}$*, and numbers* $\beta_1,\ldots,\beta_{L_{\rm c}}\in[0,1]$ *is a convex set.*

Proof: Similar to the proof of Lemma 2 in Appendix E and omitted. ■

B. Arbitrary Input Alphabets at the Covert Users

We extend the result in the previous subsection to arbitrary input alphabets $\mathcal{X}_\ell$ at the covert user $\ell\in\{1,\ldots,L_{\rm c}\}$. The only restriction is that each $\mathcal{X}_\ell$ contains the 0 symbol, which we still consider to be the "off-symbol" sent under $\mathcal{H}=0$.

To state our main result for arbitrary covert input alphabets, we extend Definition (101) as follows. Given a tuple $\boldsymbol{\rho}\triangleq(\rho_1,\ldots,\rho_{L_{\rm c}})$, a set of pmfs $\{\psi_\ell(\cdot)\}_{\ell=1}^{L_{\rm c}}$ over $\mathcal{X}_\ell\backslash\{0\}$, and a tuple $\boldsymbol{x}_{\rm nc}$, define:

$$\chi^2(\boldsymbol{\rho},\{\psi_\ell\},\boldsymbol{x}_{\rm nc})\triangleq\sum_{z\in\mathcal{Z}}\frac{\left(\sum_{\ell=1}^{L_{\rm c}}\frac{\rho_\ell}{\|\boldsymbol{\rho}\|_1}\sum_{x_\ell\in\mathcal{X}_\ell}\psi_\ell(x_\ell)\Gamma_{Z|\boldsymbol{X}_{\rm c}\boldsymbol{X}_{\rm nc}}(z\mid x_\ell\cdot\boldsymbol{e}_\ell,\boldsymbol{x}_{\rm nc})-\Gamma_{Z|\boldsymbol{X}_{\rm c}\boldsymbol{X}_{\rm nc}}(z\mid\boldsymbol{0},\boldsymbol{x}_{\rm nc})\right)^2}{\Gamma_{Z|\boldsymbol{X}_{\rm c}\boldsymbol{X}_{\rm nc}}(z\mid\boldsymbol{0},\boldsymbol{x}_{\rm nc})}. \tag{106}$$

In a similar way, we extend the definitions of (99) and (100), which now depend on the non-zero symbol $x_\ell\in\mathcal{X}_\ell$ used by the covert user ℓ. For given $x_\ell\in\mathcal{X}_\ell\backslash\{0\}$ and $\boldsymbol{x}_{\rm nc}$, define:

$$\mathbb{D}_Y^{(\ell)}(x_\ell,\boldsymbol{x}_{\rm nc})=\mathbb{D}\left(\Gamma_{Y|\boldsymbol{X}_{\rm c}\boldsymbol{X}_{\rm nc}}(\cdot\mid x_\ell\cdot\boldsymbol{e}_\ell,\boldsymbol{x}_{\rm nc})\|\Gamma_{Y|\boldsymbol{X}_{\rm c}\boldsymbol{X}_{\rm nc}}(\cdot\mid\boldsymbol{0},\boldsymbol{x}_{\rm nc})\right), \tag{107}$$

$$\mathbb{D}_Z^{(\ell)}(x_\ell,\boldsymbol{x}_{\rm nc})=\mathbb{D}\left(\Gamma_{Z|\boldsymbol{X}_{\rm c}\boldsymbol{X}_{\rm nc}}(\cdot\mid x_\ell\cdot\boldsymbol{e}_\ell,\boldsymbol{x}_{\rm nc})\|\Gamma_{Z|\boldsymbol{X}_{\rm c}\boldsymbol{X}_{\rm nc}}(\cdot\mid\boldsymbol{0},\boldsymbol{x}_{\rm nc})\right). \tag{108}$$

Theorem 4 (Arbitrary Covert Input Alphabets)**.** *A message and secret-key rate tuple* $(r_1,\ldots,r_{L_{\rm c}},R_{L_{\rm c}+1},\ldots,R_L,k_1,\ldots,k_{L_{\rm c}})$ *is achievable if, and only if, there exists a tuple of random variables* $(T,\boldsymbol{X}_{\rm nc})$ *over* $\mathcal{T}\times\mathcal{X}_{L_{\rm c}+1}\times\cdots\times\mathcal{X}_L$ *distributed according to a joint pmf* $P_{T\boldsymbol{X}_{\rm nc}}$*, nonnegative tuples* $\{\boldsymbol{\rho}_t\triangleq(\rho_{1,t},\ldots,\rho_{L_{\rm c},t})\}_{t\in\mathcal{T}}$ *and numbers* $\beta_1,\ldots,\beta_{L_{\rm c}}\in[0,1]$*, and sets of marginal pmfs* $\{\boldsymbol{\psi}_T\triangleq\{\psi_{1,t},\ldots,\psi_{L_{\rm c},t}\}\}_{t\in\mathcal{T}}$ *over* $\mathcal{X}_1\backslash 0,\ldots,\mathcal{X}_{L_{\rm c}}\backslash\{0\}$ *so that the following inequalities hold:*

$$r_\ell\leq\beta_\ell\sqrt{2}\frac{\mathbb{E}_{P_{T\boldsymbol{X}_{\rm nc}}}\left[\rho_{\ell,T}\sum_{x_\ell\in\mathcal{X}_\ell}\psi_{\ell,T}(x_\ell)\cdot\mathbb{D}_Y^{(\ell)}(x_\ell,\boldsymbol{X}_{\rm nc})\right]}{\sqrt{\mathbb{E}_{P_{T\boldsymbol{X}_{\rm nc}}}\left[\|\boldsymbol{\rho}_T\|_1^2\cdot\chi^2(\boldsymbol{\rho}_T,\boldsymbol{\psi}_T,\boldsymbol{X}_{\rm nc})\right]}},\quad\forall\ell\in\{1,\ldots,L_{\rm c}\}, \tag{109}$$

$$\sum_{j\in\mathcal{J}}R_j\leq I(\boldsymbol{X}_{{\rm nc},\mathcal{J}};Y\mid\boldsymbol{X}_{\rm c}=\boldsymbol{0},\boldsymbol{X}_{{\rm nc},\mathcal{J}^c},T),\quad\forall\mathcal{J}\subseteq\{L_{\rm c}+1,\ldots,L\}, \tag{110}$$

$$k_\ell\geq\beta_\ell\sqrt{2}\frac{\mathbb{E}_{P_{T\boldsymbol{X}_{\rm nc}}}\left[\rho_{\ell,T}\sum_{x_\ell\in\mathcal{X}_\ell}\psi_{\ell,T}(x_\ell)\cdot\left(\mathbb{D}_Z^{(\ell)}(x_\ell,\boldsymbol{X}_{\rm nc})-\mathbb{D}_Y^{(\ell)}(x_\ell,\boldsymbol{X}_{\rm nc})\right)\right]}{\sqrt{\mathbb{E}_{P_{T\boldsymbol{X}_{\rm nc}}}\left[\|\boldsymbol{\rho}_T\|_1^2\cdot\chi^2(\boldsymbol{\rho}_T,\boldsymbol{\psi}_T,\boldsymbol{X}_{\rm nc})\right]}},\quad\forall\ell\in\{1,\ldots,L_{\rm c}\}. \tag{111}$$

Proof: Achievability can be proved by analyzing an extension of the coding schemes proposed in Sections III-B and III-C, where in the code construction the entries of the codewords at a covert User $\ell=1,\ldots,L_{\rm c}$ are drawn conditionally i.i.d. given the t^n sequence according to the conditional pmf

$$P_{X_{\ell,n}|T}(x_\ell|t)=\psi_{\ell,t}(x_\ell)\cdot\rho_{\ell,t}\cdot\alpha_n,\quad x_\ell\in\mathcal{X}_\ell\backslash\{0\} \tag{112}$$

and

$$P_{X_{\ell,n}|T}(0|t)=1-\rho_{\ell,t}\cdot\alpha_n. \tag{113}$$

The converse proof is obtained by generalizing the converse in Appendix D to non-binary input alphabets at the covert users, similarly to [21, Appendix G].
Details of both proofs are omitted. ■

C. Interference channels

In this section, we consider the two-receiver discrete memoryless interference channel (DMIC) in Figure 11. We have two covert users, each sending a message to their respective legitimate receiver and a non-covert user sending a common message to both legitimate receivers. The transition law of the DMIC is denoted $\Gamma_{Y_1Y_2Z|X_1X_2X_3}$.

Encodings are as defined in Section II for the multi-access channel. Decoding now takes place at two different legitimate receivers $\ell\in\{1,2\}$, which both know the hypothesis $\mathcal{H}$.

Under $\mathcal{H}=0$ each of the two legitimate receivers ℓ decodes the common message W_3 by producing the guess

$$\widehat{W}_3^{(\ell)}=g_{0,\ell}^{(n)}(Y_\ell^n) \tag{114}$$

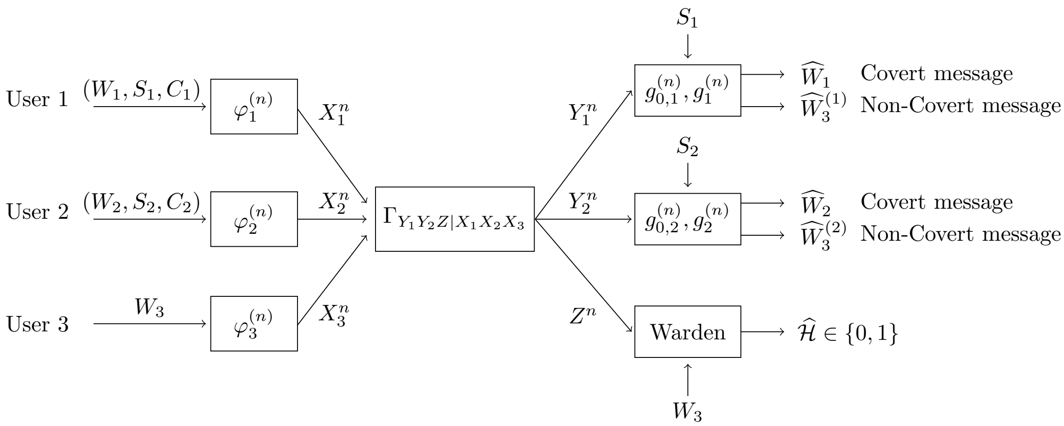


Fig. 11: Interference channel with two covert users and a non-covert user sending a common message to both legitimate receivers.

using a decoding function $g_{0,\ell}^{(n)}(\cdot)$.

Under $\mathcal{H} = 1$, it uses a decoding function $g_\ell^{(n)}\colon \mathcal{Y}^n \times \mathcal{K}_\ell \to \mathcal{M}_\ell \times \mathcal{M}_3$ to produce the pair of guesses

$$(\widehat{W}_\ell, \widehat{W}_3^{(\ell)}) = g_\ell^{(n)}(Y_\ell^n, S_\ell). \tag{115}$$

Achievability is defined analogously to the DMMAC, see Definition 1, but where the definitions of the probabilities of error (20) and (21) need to be replaced by

$$P_{e,0} \triangleq \Pr\left(\widehat{W}_3^{(1)} \neq W_3 \text{ or } \widehat{W}_3^{(2)} \neq W_3 \middle| \mathcal{H} = 0\right), \tag{116}$$

$$P_{e,1} \triangleq \Pr\left(\widehat{W}_3^{(1)} \neq W_3 \text{ or } \widehat{W}_3^{(2)} \neq W_3 \text{ or } \widehat{W}_2 \neq W_2 \text{ or } \widehat{W}_1 \neq W_1 \middle| \mathcal{H} = 1\right). \tag{117}$$

Let $\mathcal{T} \triangleq \{1, \ldots, 7\}$. We have the following theorem for the DMIC.

Theorem 5. *A message and secret-key rates tuple $(r_1, r_2, R_3, k_1, k_2)$ is achievable over the DMIC $\Gamma_{Y_1 Y_2 Z|X_1 X_2 X_3}$ if, and only if, there exists a pair of random variables (T, X_3) over $\mathcal{T} \times \mathcal{X}_3$ distributed according to a pmf P_{TX_3} over $\mathcal{T}$ and $\mathcal{X}_3$, nonnegative tuples $\{(\rho_{1,t}, \rho_{2,t})\}_{t\in\mathcal{T}}$, and $\beta_1, \beta_2 \in [0,1]$, so that for all $\ell \in \{1,2\}$, the following three inequalities hold:*

$$r_\ell \leq \beta_\ell \sqrt{2} \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T} \mathbb{D}_{Y_\ell}^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T} + \rho_{2,T})^2 \chi^2(\rho_{1,T}, \rho_{2,T}, X_3)\right]}}, \tag{118}$$

$$R_3 \leq \min\left(\mathbb{I}(X_3; Y_1 \mid X_1 = 0, X_2 = 0, T); \mathbb{I}(X_3; Y_2 \mid X_1 = 0, X_2 = 0, T)\right), \tag{119}$$

$$k_\ell \geq \beta_\ell \sqrt{2} \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T}\left(\mathbb{D}_Z^{(\ell)}(X_3) - \mathbb{D}_{Y_\ell}^{(\ell)}(X_3)\right)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T} + \rho_{2,T})^2 \chi^2(\rho_{1,T}, \rho_{2,T}, X_3)\right]}}. \tag{120}$$

Notice that here in Theorem 5, we have a different output signal Y_ℓ for each legitimate receiver $\ell \in \{1,2\}$.

Proof: Analogous to the DMMAC proof, but with the following modifications for the achievability and converse proofs.

1) Achievability: In the decoding, (40) needs to be replaced by the following two conditions (one for each legitimate receiver)

$$(t^n, x_3^n(w_3), y_1^n) \in \mathcal{T}_{\mu_n}^n(P_{TX_3Y_1}), \tag{121}$$

$$(t^n, x_3^n(w_3), y_2^n) \in \mathcal{T}_{\mu_n}^n(P_{TX_3Y_2}), \tag{122}$$

and (42) and (43) need to be replaced by

$$(x_1^n(w_1, S_1), 0^n, x_3^n(\widehat{W}_3^{(1)}), y_1^n) \in \mathcal{A}_{\eta_1}^n, \tag{123}$$

and

$$(0^n, x_2^n(w_2, S_2), x_3^n(\widehat{W}_3^{(2)}), y_2^n) \in \mathcal{A}_{\eta_2}^n. \tag{124}$$

Accordingly, in the analysis, the DMMAC output Y also need to be replaced by either of the two DMIC outputs Y_1 or Y_2.

2) Converse: Replace the DMMAC output Y by the two DMIC outputs Y_ℓ when deriving the bounds on r_ℓ and on k_ℓ, and perform the steps to bound R_3 twice: once using output Y_1 instead of Y and once using output Y_2. ■

Notice that now the rate of the non-covert message is the minimum rate at which the legitimate receivers can decode the non-covert message. This is because (121) requires condition

$$R_3 \leq \mathbb{I}(X_3; Y_1 \mid X_1 = 0, X_2 = 0, T) \tag{125}$$

for reliable decoding, whereas (122) requires

$$R_3 \leq \mathbb{I}(X_3; Y_2 \mid X_1 = 0, X_2 = 0, T). \tag{126}$$

V. Conclusion

This paper characterizes the fundamental limits of a multi-access communication setup with covert users (sharing common secret-keys of fixed rates with the legitimate receiver) and non-covert users, all communicating with the same legitimate receiver in presence of a warden. Our findings illustrate an intricate interplay among three pivotal quantities: the covert users' rates, the non-covert users' rate and the secret-key rates. Similarly to multi-access scenarios without covert constraints, our results emphasize the necessity of a multiplexing (coded time-sharing) strategy in the code constructions, so as to exhaust the entire tradeoff of achievable covert and non-covert rates. This holds even with only a single non-covert user. Our results also prove that the presence of the non-covert users can increase the covert-capacity under a stringent secret-key rate constraint.

The scenario considered in this work contains as interesting special cases covert communication over a single-user DMC or over a DMMAC (without non-covert users). Our results also imply new findings for these previously studied special cases. In fact, with our new results, we can characterize the minimum secret-key rates that are required to achieve any set over covert data-rates, while previous works only characterized the secret-key rates required to transmit at largest covert rates. With our new results we can thus determine the set of covert rates that are achievable over a DMC or a DMMAC under a given stringent secret-key budget, which was not possible with the previous findings. Notice that while for the single-user DMC our findings apply both to setups with and without local randomness at the transmitter, for the DMMAC our scheme requires local randomness at least at some of the transmitters when we do not communicate at largest possible covert data-rates.

We further showed that our findings on the mixed covet and non-covert DMMAC naturally also extend to related setups including the interference channels or channels with a friendly jammer. Further interesting research directions include studies of fading channels, or non-synchronized transmissions.

Appendix A
Proof of Theorem 1

A. Analysis of the Decoding Error Probability of the Scheme in Section III-B

In this section, we analyze the expected error probabilities $\mathbb{E}_{\mathcal{C}}[P_{e,0}]$ and $\mathbb{E}_{\mathcal{C}}[P_{e,1}]$, where expectations are with respect to the random codebooks.

1) Analysis of $\mathbb{E}_{\mathcal{C}}[P_{e,0}]$: By standard arguments and because for all $t \in \mathcal{T}$, $P_{X_{1,n}}(0 \mid t)$ and $P_{X_{2,n}}(0 \mid t)$ tend to 1 as $n \to \infty$:

$$\lim_{n\to\infty} \mathbb{E}_{\mathcal{C}}[P_{e,0}] = 0 \tag{127}$$

whenever

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\mathsf{M}_3) \leq I(X_3; Y \mid X_1 = 0, X_2 = 0, T). \tag{128}$$

2) Analysis of $\mathbb{E}_{\mathcal{C}}[P_{e,1}]$: Define the probabilities

$$P_{e,1,1} \triangleq \Pr\left(\widehat{W}_1 \neq W_1 \middle| \mathcal{H} = 1\right), \tag{129}$$

$$P_{e,1,2} \triangleq \Pr\left(\widehat{W}_2 \neq W_2 \middle| \mathcal{H} = 1\right), \tag{130}$$

$$P_{e,1,3} \triangleq \Pr\left(\widehat{W}_3 \neq W_3 \middle| \mathcal{H} = 1, \widehat{W}_1 = W_1, \widehat{W}_2 = W_2\right), \tag{131}$$

and notice that by the union bound we have

$$P_{e,1} \leq P_{e,1,1} + P_{e,1,2} + P_{e,1,3}. \tag{132}$$

In the following, we analyze each of the three summands separately.

Analyzing $\mathbb{E}_{\mathcal{C}}[P_{e,1,1}]$:

By the symmetry of the code construction and the uniformity of the messages and the key, we can assume that $W_1 = 1$, $S_1 = 1$ and $W_3 = w_3$. Then, we have:

$$\mathbb{E}_{\mathcal{C}}[P_{e,1,1}] \leq \Pr[(X_1^n(1,1), 0^n, X_3^n(w_3), Y^n) \notin \mathcal{A}_{\eta_1}^n] + \sum_{w_1=2}^{\mathsf{M}_1} \Pr[(X_1^n(w_1,1), 0^n, X_3^n(w_3), Y^n) \in \mathcal{A}_{\eta_1}^n] \tag{133}$$

We first analyze a specific term in the summation of (133). For any $w_1 \in [|2, \mathsf{M}_1|]$ we have:

$$\begin{aligned}
&\Pr[(X_1^n(w_1,1), 0^n, X_3^n(w_3), Y^n) \in \mathcal{A}_{\eta_1}^n] && (134)\\
&= \mathbb{E}_{X_1^n(w_1,1), X_3^n(w_3), Y^n} \left[\left\{ \mathbb{1}(X_1^n(w_1,1), 0^n, X_3^n(w_3), Y^n) \in \mathcal{A}_{\eta_1}^n \right\} \right] && (135)\\
&\overset{(a)}{\leq} e^{-\eta_1} \mathbb{E}_{X_1^n(w_1,1), X_3^n(w_3), Y^n} \left[\frac{\Gamma_{Y|X_1X_2X_3}^{\otimes n}(Y^n | X_1^n(w_1,1), 0^n, X_3^n(w_3))}{\Gamma_{Y|X_1X_2X_3}^{\otimes n}(Y^n | 0^n, 0^n, X_3^n(w_3))} \underbrace{\mathbb{1}\left\{ (X_1^n(w_1,1), 0^n, X_3^n(w_3), Y^n) \in \mathcal{A}_{\eta_1}^n \right\}}_{\leq 1} \right] && (136)\\
&\overset{(b)}{\leq} e^{-\eta_1} \mathbb{E}_{X_1^n(w_1,1), X_3^n(w_3), Y^n} \left[\frac{\Gamma_{Y|X_1X_2X_3}^{\otimes n}(Y^n | X_1^n(w_1,1), 0^n, X_3^n(w_3))}{\Gamma_{Y|X_1X_2X_3}^{\otimes n}(Y^n | 0^n, 0^n, X_3^n(w_3))} \right] && (137)\\
&= e^{-\eta_1} \prod_{i=1}^{n} \mathbb{E}_{X_{1,i}(w_1,1), X_{3,i}(w_3), Y_i} \left[\frac{\Gamma_{Y|X_1X_2X_3}(Y_i | X_{1,i}(w_1,1), 0, X_{3,i}(w_3))}{\Gamma_{Y|X_1X_2X_3}(Y_i | 0, 0, X_{3,i}(w_3))} \right] && (138)\\
&= e^{-\eta_1} \prod_{t \in \mathcal{T}} \left(\mathbb{E}_{P_{X_3Y|T=t} P_{X_{1,n}|T=t}} \left[\frac{\Gamma_{Y|X_1X_2X_3}(Y | X_1, 0, X_3)}{\Gamma_{Y|X_1X_2X_3}(Y | 0, 0, X_3)} \right] \right)^{n\boldsymbol{\pi}(t)} && (139)\\
&\overset{(c)}{=} e^{-\eta_1} \prod_{t \in \mathcal{T}} \left(\mathbb{E}_{P_{X_3Y|T=t}} \left[\frac{\Gamma_{Y|X_2X_3}^{(t)}(Y | 0, X_3)}{\Gamma_{Y|X_1X_2X_3}(Y | 0, 0, X_3)} \right] \right)^{n\boldsymbol{\pi}(t)}. && (140)
\end{aligned}$$

Here, (a) holds by the definition of the set $\mathcal{A}_{\eta_1}^n$ in (41); (b) by replacing the indicator function by the all-one function and (c) upon defining

$$\Gamma_{Y|X_2X_3}^{(t)}(y|x_2, x_3) \triangleq \sum_{x_1 \in \{0,1\}} P_{X_{1,n}|T}(x_1 \mid t) \Gamma_{Y|X_1X_2X_3}(y|x_1, x_2, x_3). \tag{141}$$

For any $t \in \mathcal{T}$ we have:

$$\mathbb{E}_{P_{X_3Y|T=t}} \left[\frac{\Gamma_{Y|X_2X_3}^{(t)}(Y|0, X_3)}{\Gamma_{Y|X_1X_2X_3}(Y|0,0,X_3)} \right] = 1 - \rho_{1,t}\alpha_n + \rho_{1,t}\alpha_n \cdot \mathbb{E}_{P_{X_3Y|T=t}} \left[\frac{\Gamma_{Y|X_1X_2X_3}(Y|1,0,X_3)}{\Gamma_{Y|X_1X_2X_3}(Y|0,0,X_3)} \right] \tag{142}$$

Under our assumption (27a) that for any x_3 we have $\Gamma_{Y|X_1=1,X_2=0,X_3=x_3} \ll \Gamma_{Y|X_1=0,X_2=0,X_3=x_3}$, we can conclude that $\frac{\Gamma_{Y|X_1X_2X_3}(y|1,0,x_3)}{\Gamma_{Y|X_1X_2X_3}(y|0,0,x_3)}$ is uniformly upper-bounded for all realizations of y and x_3, i.e.,

$$\frac{\Gamma_{Y|X_1X_2X_3}(y|1,0,x_3)}{\Gamma_{Y|X_1X_2X_3}(y|0,0,x_3)} \leq \Delta_Y, \tag{143}$$

for some finite $\Delta_Y > 0$. We continue with this upper bound to deduce

$$\begin{aligned}
&\mathbb{E}_{P_{X_3Y|T=t}} \left[\frac{\Gamma_{Y|X_1X_2X_3}(Y|1,0,X_3)}{\Gamma_{Y|X_1X_2X_3}(Y|0,0,X_3)} \right] \\
&= \mathbb{E}_{P_{X_3|T=t}} \Bigg[(1-\rho_{1,t}\alpha_n)(1-\rho_{2,t}\alpha_n) \sum_y \Gamma_{Y|X_1X_2X_3}(y|1,0,X_3) + \rho_{1,t}\alpha_n(1-\rho_{2,t}\alpha_n) \sum_y \frac{\Gamma_{Y|X_1X_2X_3}(y|1,0,X_3)^2}{\Gamma_{Y|X_1X_2X_3}(y|0,0,X_3)} \\
&\qquad + (1-\rho_{1,t}\alpha_n)\rho_{2,t}\alpha_n \sum_y \frac{\Gamma_{Y|X_1X_2X_3}(y|1,0,X_3)}{\Gamma_{Y|X_1X_2X_3}(y|0,0,X_3)} \cdot \Gamma_{Y|X_1X_2X_3}(y|0,1,X_3) \\
&\qquad + \rho_{1,t}\rho_{2,t}\alpha_n^2 \sum_y \frac{\Gamma_{Y|X_1X_2X_3}(y|1,0,X_3)}{\Gamma_{Y|X_1X_2X_3}(y|0,0,X_3)} \cdot \Gamma_{Y|X_1X_2X_3}(y|1,1,X_3) \Bigg] && (144)\\
&\leq \mathbb{E}_{P_{X_3|T=t}} \left[(1-\rho_{2,t}\alpha_n)(1-\rho_{1,t}\alpha_n) + \left((\rho_{1,t}+\rho_{2,t})\alpha_n - \rho_{1,t}\rho_{2,t}\alpha_n^2 \right) \Delta_Y \right] && (145)\\
&= (1-\rho_{2,t}\alpha_n)(1-\rho_{1,t}\alpha_n) + \left((\rho_{1,t}+\rho_{2,t})\alpha_n + \rho_{1,t}\rho_{2,t}\alpha_n^2 \right) \Delta_Y && (146)
\end{aligned}$$

which with (142) yields:

$$\mathbb{E}_{P_{X_3Y|T=t}}\left[\frac{\Gamma^{(t)}_{Y|X_2X_3}(Y|0,X_3)}{\Gamma_{Y|X_1X_2X_3}(Y|0,0,X_3)}\right] \leq 1-\rho_{1,t}(\rho_{1,t}+\rho_{2,t})\alpha_n^2(1-\Delta_Y)+\mathcal{O}(\alpha_n^3). \tag{147}$$

Combining (140) with (147), we obtain:

$$\sum_{w_1=2}^{\mathsf{M}_1} \mathop{\mathbb{E}}_{X_1^n(w_1,1),X_3^n(w_3),Y^n}\left[\mathbb{1}\left\{(X_1^n(w_1,1),0^n,X_3^n(w_3),y^n)\in\mathcal{A}^n_{\eta_1}\right\}\right]$$

$$\leq \mathsf{M}_1 e^{-\eta_1}\prod_{t\in\mathcal{T}}\left(1-\rho_{1,t}(\rho_{1,t}+\rho_{2,t})\alpha_n^2(1-\Delta_Y)+\mathcal{O}(\alpha_n^3)\right)^{n\boldsymbol{\pi}(t)} \tag{148}$$

$$= \mathsf{M}_1 e^{-\eta_1} e^{n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\log\left(1-\rho_{1,t}(\rho_{1,t}+\rho_{2,t})\alpha_n^2(1-\Delta_Y)+\mathcal{O}(\alpha_n^3)\right)} \tag{149}$$

$$\leq \mathsf{M}_1 e^{-\eta_1} e^{-n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\left(\rho_{1,t}(\rho_{1,t}+\rho_{2,t})\alpha_n^2(1-\Delta_Y)+\mathcal{O}(\alpha_n^3)\right)} \tag{150}$$

$$\leq \mathsf{M}_1 e^{-\eta_1} e^{-\omega_n^2\sum_{t\in\mathcal{T}}(P_T(t)+\mu_n)[\rho_{1,t}(\rho_{1,t}+\rho_{2,t})(1-\Delta_Y)+\mathcal{O}(\alpha_n)]} \tag{151}$$

Notice that the term in the last exponent tends to 0 as $n\to\infty$ because $\omega_n\to 0$ and the sum over t is bounded. We conclude that the term in (151), and thus the sum in (133), tend to 0 as $n\to\infty$, if

$$\lim_{n\to\infty}(\log\mathsf{M}_1-\eta_1)=-\infty. \tag{152}$$

We next bound the first summand on the right-hand side of (133). To this end, start by noticing the following:

$$\Pr\left[\log\left(\frac{\Gamma^{\otimes n}_{Y|X_1X_2X_3}(Y^n|X_1^n(1,1),0^n,X_3^n(w_3))}{\Gamma^{\otimes n}_{Y|X_1X_2X_3}(Y^n|0^n,0^n,X_3^n(w_3))}\right)\leq\eta_1\right]=\Pr\left[\sum_{i=1}^n\log\left(\frac{\Gamma_{Y|X_1X_2X_3}(Y_i|X_{1,i},0,X_{3,i})}{\Gamma_{Y|X_1X_2X_3}(Y_i|0,0,X_{3,i})}\right)\leq\eta_1\right] \tag{153}$$

where $X_{1,i},X_{3,i}$ and Y_i denote the i-th entries of $X_1^n(1,1),X_3^n(w_3)$, and Y^n, and probabilities are with respect to the randomness in the code construction and the channel. For each $i\in\{1,\ldots,n\}$, consider the random variable

$$\Xi_i\triangleq\log\left(\frac{\Gamma_{Y|X_1X_2X_3}(Y_i|X_{1,i},0,X_{3,i})}{\Gamma_{Y|X_1X_2X_3}(Y_i|0,0,X_{3,i})}\right), \tag{154}$$

where as above, the tuple $(X_{1,i},X_{3,i},Y_i)$ follows the joint pmf $P_{X_{1,t_i}}(x_1)P_{X_3|T=t_i}(x_3)\Gamma_{Y|X_1X_2X_3}(y|x_1,0,x_3)$. Let

$$\Lambda_Y\triangleq\min_{(x_1,x_3,y)}\log\left(\frac{\Gamma_{Y|X_1X_2X_3}(y|x_1,0,x_3)}{\Gamma_{Y|X_1X_2X_3}(y|0,0,x_3)}\right), \tag{155}$$

where the minimum is only over triples (x_1,x_3,y) for which the ratio $\frac{\Gamma_{Y|X_1X_2X_3}(y|x_1,0,x_3)}{\Gamma_{Y|X_1X_2X_3}(y|0,0,x_3)}$ is non-zero. (Since we prevent the ratio inside the log from being 0 and the sets $\mathcal{X}_1,\mathcal{X}_3,\mathcal{Y}$ are all finite, the minimum must exist.) By above definition and Assumption (27a), we have $\Pr[|\Xi_i|\leq c]=1$ for $c\triangleq\max\{|\log\Lambda_Y|,|\log\Delta_Y|\}$. Moreover, the first and second moments of the random variable Ξ_i, for $i\in\{1,\ldots,n\}$, satisfy

$$\mathbb{E}\left[\Xi_i\right]\overset{(a)}{=}\rho_{1,t_i}\alpha_n\mathbb{E}_{P_{X_3|T=t_i}}\cdot\left[\mathbb{D}_Y^{(1)}(X_3)\right] \tag{156}$$

$$\mathbb{E}\left[\Xi_i^2\right]\overset{(b)}{\leq}\rho_{1,t_i}\alpha_n\cdot(\log\Delta_Y)^2, \tag{157}$$

where (a) and (b) hold because when $x_1=0$ the log term in the definition of Ξ_i, (154), is zero. We can thus apply a large deviation argument to bound the probability

$$\Pr\left[\frac{1}{n}\sum_{i=1}^n\Xi_i<\alpha_n\mathbb{E}_{\boldsymbol{\pi}P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Y^{(1)}(X_3)\right]-a\right], \tag{158}$$

for any $a>0$. Since above probability coincides with the probabilities in (153) for $\eta_1=n\left(\alpha_n\mathbb{E}_{\boldsymbol{\pi}P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Y^{(1)}(X_3)\right]-a\right)$, we obtain by Bernstein's inequality:

$$\Pr\left[\log\left(\frac{\Gamma^{\otimes n}_{Y|X_1X_2X_3}(Y^n|X_1^n(1,1),0^n,X_3^n(w_3))}{\Gamma^{\otimes n}_{Y|X_1X_2X_3}(Y^n|0^n,0^n,X_3^n(w_3))}\right)\leq\eta_1\right]\leq 2e^{-\frac{na^2}{\alpha_n\mathbb{E}_{\boldsymbol{\pi}}[\rho_{1,T}](\log\Delta_Y)^2+2/3ac}}, \tag{159}$$

for any $\eta_1\geq n\left(\alpha_n\mathbb{E}_{\boldsymbol{\pi}P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Y^{(1)}(X_3)\right]-a\right)$. Noting that $\lim_{n\to\infty}\boldsymbol{\pi}(t)\to P_T(t)$, we specialize above bound to the choice

$$\eta_1=(1-\mu_1)n\alpha_n\mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Y^{(1)}(X_3)\right] \tag{160}$$

for an arbitrary $\mu_1 \in (0,1)$, in which case a scales as α_n and the exponent scales as $-n\alpha_n = -\omega_n\sqrt{n}$ and thus tends to $-\infty$ in the limit as $n \to \infty$. As a consequence, the probability (159), and thus (153), vanish exponentially fast in the blocklength n as $n \to \infty$.

Choosing further

$$\log(\mathsf{M}_1) = (1-\xi_1)\alpha_n n \mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Y^{(1)}(X_3)\right], \tag{161}$$

for any small number $\xi_1 > \mu_1$ by (152), we finally conclude that the entire probability $\mathbb{E}_{\mathcal{C}}[P_{e,1,1}]$ on the right-hand side of (133) vanishes exponentially fast in the blocklength.

Analyzing $\mathbb{E}_{\mathcal{C}}[P_{e,1,2}]$:

By symmetry, the same steps allow one to conclude also that for any small numbers $\mu_2 > 0$ and $\xi_2 > \mu_2$ and under the choices

$$\eta_2 \triangleq (1-\mu_2)\alpha_n n \mathbb{E}_{P_{TX_3}}\left[\rho_{2,T}\mathbb{D}_Y^{(2)}(X_3)\right], \tag{162}$$

$$\log(\mathsf{M}_2) = (1-\xi_2)\alpha_n n \mathbb{E}_{P_{TX_3}}\left[\rho_{2,T}\mathbb{D}_Y^{(2)}(X_3)\right], \tag{163}$$

the probability of decoding error $\mathbb{E}_{\mathcal{C}}[P_{e,1,2}]$ of Message W_2 vanishes exponentially fast in the blocklength n.

Analyzing $\mathbb{E}_{\mathcal{C}}[P_{e,1,3}]$:

As in the analysis of $P_{e,0}$, we deduce that under condition (128):

$$\lim_{n\to\infty} \mathbb{E}_{\mathcal{C}}[P_{e,1,3}] = 0 \tag{164}$$

B. Channel Resolvability Analysis

1) Auxiliary Lemma and Definitions: We will need the following lemma, which is an immediate consequence of [17, Lemma 1].

Lemma 3. *For each blocklength n, consider two pmfs $P_{X_{1,n}}$ and $P_{X_{2,n}}$ over the binary alphabets $\mathcal{X}_1 = \mathcal{X}_2 = \{0,1\}$ respectively, such that*

$$\lim_{n\to\infty} P_{X_{\ell,n}}(1) = 0, \quad \ell \in \{1,2\}. \tag{165}$$

Let $\mathcal{X}_3$, $\mathcal{Z}$, and $\Gamma_{Z|X_1X_2X_3}$ be as defined earlier. Then, for all sufficiently large values of n, the conditional pmfs

$$\Gamma_{Z|X_3}(z \mid x_3) \triangleq \sum_{(x_1,x_2)\in\mathcal{X}_1\times\mathcal{X}_2} P_{X_{1,n}}(x_1)P_{X_{2,n}}(x_2)\Gamma_{Z|X_1X_2X_3}(z|x_1,x_2,x_3), \quad x_3 \in \mathcal{X}_3,\ z \in \mathcal{Z}, \tag{166}$$

satisfy

$$\mathbb{D}(\Gamma_{Z|X_3}(\cdot \mid x_3) \,\|\, \Gamma_{Z|X_1X_2X_3}(\cdot \mid 0,0,x_3) \tag{167}$$

$$= (1+o(1)) \cdot \frac{\left(P_{X_{1,n}}(1)+P_{X_{2,n}}(1)\right)^2}{2}\chi^2\left(\frac{P_{X_{1,n}}(1)}{P_{X_{1,n}}(1)+P_{X_{2,n}}(1)}, \frac{P_{X_{2,n}}(1)}{P_{X_{1,n}}(1)+P_{X_{2,n}}(1)}, x_3\right). \tag{168}$$

The following definitions will be useful in our resolvability analysis. For any $t \in \mathcal{T}$, define the averaged channels

$$\Gamma^{(t)}_{Z|X_3}(z|x_3) \triangleq \sum_{(x_1,x_2)\in\mathcal{X}_1\times\mathcal{X}_2} P_{X_{1,n}|T}(x_1 \mid t)P_{X_{2,n}|T}(x_2 \mid t)\Gamma_{Z|X_1X_2X_3}(z|x_1,x_2,x_3), \tag{169}$$

$$\Gamma^{(t)}_{Z|X_2X_3}(z|x_2,x_3) \triangleq \sum_{x_1} P_{X_{1,n}|T}(x_1 \mid t)\Gamma_{Z|X_1X_2X_3}(z|x_1,x_2,x_3), \tag{170}$$

$$\Gamma^{(t)}_{Z|X_1X_3}(z|x_1,x_3) \triangleq \sum_{x_2} P_{X_{2,n}|T}(x_2 \mid t)\Gamma_{Z|X_1X_2X_3}(z|x_1,x_2,x_3), \tag{171}$$

and the corresponding product channels

$$\tilde{\Gamma}^n_{Z|X_3}(z^n \mid x_3^n) \triangleq \prod_{i=1}^n \Gamma^{(t_i)}_{Z|X_3}(z_i \mid x_{3,i}), \tag{172}$$

$$\tilde{\Gamma}^n_{Z|X_2X_3}(z^n \mid x_2^n, x_3^n) \triangleq \prod_{i=1}^n \Gamma^{(t_i)}_{Z|X_2X_3}(z_i \mid x_{2,i}, x_{3,i}), \tag{173}$$

$$\tilde{\Gamma}^n_{Z|X_1X_3}(z^n \mid x_1^n, x_3^n) \triangleq \prod_{i=1}^n \Gamma^{(t_i)}_{Z|X_1X_3}(z_i \mid x_{1,i}, x_{3,i}). \tag{174}$$

2) The Proof: Recall that the warden's output distribution under $\mathcal{H}=1$ for a given codebook $\mathcal{C}$ and message $w_3 \in \mathcal{M}_3$:

$$\widehat{Q}^n_{\mathcal{C},w_3}(z^n) = \frac{1}{\mathsf{M}_1\mathsf{M}_2}\frac{1}{\mathsf{K}_1\mathsf{K}_2}\sum_{w_1=1}^{\mathsf{M}_1}\sum_{s_1=1}^{\mathsf{K}_1}\sum_{w_2=1}^{\mathsf{M}_2}\sum_{s_2=1}^{\mathsf{K}_2}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|x_1^n(w_1,s_1),x_2^n(w_1,s_2),x_3^n(w_3)). \tag{175}$$

In this section, we show the limit

$$\mathbb{E}_{\mathcal{C}}\left[\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))\right)\right] \to 0 \qquad \forall w_3 \in \mathcal{M}_3, \tag{176}$$

where expectation is with respect to the random code construction. Fix a message $w_3 \in \mathcal{W}_3$ and a codeword $x_3^n(w_3)$. We start by expanding the divergence of interest as follows:

$$\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))\right) = \mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}\right) + \mathbb{D}\left(\tilde{\Gamma}^n_{Z|X_3}\middle\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))\right) \tag{177}$$

$$+\sum_{z^n}\left(\widehat{Q}^n_{\mathcal{C},w_3}(z^n) - \tilde{\Gamma}^n_{Z|X_3}(z^n \mid x_3^n(w_3))\right)\log\left(\frac{\tilde{\Gamma}^n_{Z|X_3}(z^n \mid x_3^n(w_3))}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))}\right). \tag{178}$$

Defining ∇_0 as the minimum probability in the support of $\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3(w_3))$:

$$\nabla_0 = \min_{z,x_3\in\text{supp}(\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3(w_3)))}\Gamma_{Z|X_1X_2X_3}(z|0,0,x_3(w_3)), \tag{179}$$

by Pinsker's inequality[2], we can conclude the following:

$$\left|\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))\right) - \mathbb{D}\left(\tilde{\Gamma}^n_{Z|X_3}\middle\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))\right)\right| \tag{180}$$

$$\leq \mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}\right) + n\log\left(\frac{1}{\nabla_0}\right)\sqrt{\frac{1}{2}\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}\right)}. \tag{181}$$

We shall separately analyze the divergences $\mathbb{D}(\tilde{\Gamma}^n_{Z|X_3}(\cdot \mid x_3^n(w_3))\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3)))$ and $\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}\right)$, or more precisely, their expectations over the choices of the codebooks.

Analysis of the expected divergence $\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}(\cdot \mid x_3^n)\right)$*:* Let Z^n be the output sequence observed at the warden under $\mathcal{H}=1$ and W_3 the message of the non-covert user. Given that $W_3 = w_3$ and for given codebooks $\mathcal{C}$, we then have $Z^n \sim \widehat{Q}^n_{\mathcal{C},w_3}$. Consider the average (over the codebooks) expected divergence

$$\mathop{\mathbb{E}}_{\mathcal{C}}\left[\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}(\cdot \mid X_3^n(w_3))\right)\right]$$

$$= \mathop{\mathbb{E}}_{\{X_1^n(w_1,s_1)\},\{X_2^n(w_2,s_2)\},X_3^n(w_3)}\left[\sum_{z^n}\widehat{Q}^n_{\mathcal{C},w_3}(z^n)\log\left(\frac{\widehat{Q}^n_{\mathcal{C},w_3}(z^n)}{\tilde{\Gamma}^n_{Z|X_3}(z^n \mid X_3^n(w_3))}\right)\right] \tag{182}$$

$$\stackrel{(a)}{=} \mathop{\mathbb{E}}_{\{X_1^n(w_1,s_1)\},\{X_2^n(w_2,s_2)\},X_3^n(w_3)}\left[\mathop{\mathbb{E}}_{Z^n}\left[\log\left(\frac{\sum_{(w_1',w_2',s_1',s_2')}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(w_1',s_1'),X_2^n(w_2',s_2'),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n \mid X_3^n(w_3))}\right)\right]\right] \tag{183}$$

$$\stackrel{(b)}{\leq} \mathop{\mathbb{E}}_{\substack{X_1^n(1,1),\\X_2^n(1,1),\\X_3^n(w_3),Z^n}}\left[\log\left(\mathop{\mathbb{E}}_{\substack{\{X_1^n(w_1,s_1)\}\setminus X_1^n(1,1),\\\{X_2^n(w_2,s_2)\}\setminus X_2^n(1,1)}}\left[\frac{\sum_{(w_1',w_2',s_1',s_2')}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(w_1',s_1'),X_2^n(w_2',s_2'),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n \mid X_3^n(w_3))}\right]\right)\right] \tag{184}$$

$$= \mathbb{E}\left[\log\left(\sum_{(w_1,w_2,s_1,s_2)}\mathop{\mathbb{E}}_{\substack{\{X_1^n(w_1,s_1)\}\setminus X_1^n(1,1),\\\{X_2^n(w_2,s_2)\}\setminus X_2^n(1,1)}}\left[\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(w_1',s_1'),X_2^n(w_2',s_2'),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n \mid X_3^n(w_3))}\right]\right)\right] \tag{185}$$

$$\stackrel{(c)}{=} \mathbb{E}\left[\log\left(\frac{(\mathsf{M}_1\mathsf{K}_1-1)(\mathsf{M}_2\mathsf{K}_2-1)}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2} + \frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n \mid X_3^n(w_3))}\right.\right.$$

[2]For any two distributions P and Q on the same alphabet $\mathcal{X}$ we have $\mathbb{V}(P,Q) \leq \sqrt{\frac{\mathbb{D}(P,Q)}{2}}$.

$$+\sum_{(w_2',s_2')\neq(1,1)} \mathop{\mathbb{E}}_{X_2^n(w_2',s_2')}\left[\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(w_2',s_2'),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right]$$

$$+\sum_{(w_1',s_1')\neq(1,1)} \mathop{\mathbb{E}}_{X_1^n(w_1',s_1')}\left[\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(w_1',s_1'),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right]\Bigg)\Bigg] \tag{186}$$

$$\stackrel{(d)}{=}\mathbb{E}\Bigg[\log\Bigg(\frac{(\mathsf{M}_1\mathsf{K}_1-1)(\mathsf{M}_2\mathsf{K}_2-1)}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2}+\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}$$

$$+\frac{(\mathsf{M}_2\mathsf{K}_2-1)\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n|X_1^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}+\frac{(\mathsf{M}_1\mathsf{K}_1-1)\tilde{\Gamma}^n_{Z|X_2X_3}(Z^n|X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}_{Z|X_3}(Z^n\mid X_3^n(w_3))}\Bigg)\Bigg] \tag{187}$$

$$\leq\mathbb{E}\Bigg[\log\Bigg(1+\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}$$

$$+\frac{\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n|X_1^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}+\frac{\tilde{\Gamma}^n_{Z|X_2X_3}(Z^n|X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\Bigg)\Bigg], \tag{188}$$

where in before equation (c) the warden's output sequence Z^n is generated from $X_1^n(W_1,S_1)$, $X_2^n(W_2,S_2)$, and $X_3^n(w_3)$ according to the memoryless channel law $\Gamma^{\otimes n}_{Z|X_1X_2X_3}$, and starting with (c) it is generated according to the same channel law but based on the random codewords $X_1^n(1,1)$, $X_2^n(1,1)$, and $X_3^n(w_3)$.

Above sequence of (in)equalities are justified as follows:

- (a) holds by rewriting the summation as an expectation over Z^n;
- (b) holds by applying Jensen's inequality over all expectations except the expectations over $X_1^n(W_1,S_1), X_2^n(W_2,S_2), X_3^n(w_3), Z^n$ and by assuming that $W_1=W_2=S_1=S_2=1$ and thus Z^n is generated from $X_1^n(1,1)$, $X_2^n(1,1)$, and $X_3^n(w_3)$ according to $\Gamma^{\otimes n}_{Z|X_1X_2X_3}$. This assumption is without loss of optimality by the symmetry of the code construction;
- (c),(d) hold by the linearity of expectation and because for $(w_1',s_1')\neq(1,1)$ and $(w_2',s_2')\neq 1$

$$\mathop{\mathbb{E}}_{X_1^n(w_1',s_1'),X_2^n(w_2',s_2')}\left[\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(w_1',s_1'),X_2^n(w_2',s_2'),X_3^n(w_3))\right]=\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3)) \tag{189}$$

and

$$\mathop{\mathbb{E}}_{X_1^n(w_1',s_1')}\left[\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(w_1',s_1'),X_2^n(1,1),X_3^n(w_3))\right]=\tilde{\Gamma}^n_{Z|X_2X_3}(Z^n\mid X_2^n(1,1),X_3^n(w_3)) \tag{190}$$

$$\mathop{\mathbb{E}}_{X_2^n(w_2',s_2')}\left[\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(w_2',s_2'),X_3^n(w_3))\right]=\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n\mid X_1^n(1,1),X_3^n(w_3)). \tag{191}$$

Define for any triple $\boldsymbol{\theta}=(\theta_0,\theta_1,\theta_2)$ the set

$$\mathcal{B}^n_{\boldsymbol{\theta}}\triangleq\Bigg\{(x_1^n,x_2^n,x_3^n,z^n)\in\mathcal{X}_1^n\times\mathcal{X}_2^n\times\mathcal{X}_3^n\times\mathcal{Z}^n:\log\left(\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n\mid x_1^n,x_2^n,x_3^n)}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n\mid 0^n,0^n,x_3^n)}\right)\leq\theta_0,$$

$$\log\left(\frac{\tilde{\Gamma}^n_{Z|X_1X_3}(z^n\mid x_1^n,x_3^n)}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n\mid 0^n,0^n,x_3^n)}\right)\leq\theta_1,$$

$$\log\left(\frac{\tilde{\Gamma}^n_{Z|X_2X_3}(z^n\mid x_2^n,x_3^n)}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n\mid 0^n,0^n,x_3^n)}\right)\leq\theta_2\Bigg\}, \tag{192}$$

and denote by A and B the events $\{(X_1^n(1,1),X_2^n(1,1),X_3^n(w_3),Z^n)\in\mathcal{B}^n_{\boldsymbol{\theta}}\}$ and $\{(X_1^n(1,1),X_2^n(1,1),X_3^n(w_3),Z^n)\notin\mathcal{B}^n_{\boldsymbol{\theta}}\}$ respectively. Using the total law of expectation we rewrite (188) as:

$$\mathop{\mathbb{E}}_{\{X_1^n(w_1',s_1')\}_{(w_1',s_1')},\{X_2^n(w_2',s_2')\}_{(w_2',s_2')},X_3^n(w_3)}\left[\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}(\cdot\mid X_3^n(w_3))\right)\right]$$

$$\leq\mathbb{E}\Bigg[\log\Bigg(1+\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}+\frac{\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n\mid X_1^n,X_3^n)}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}$$

$$+\frac{\tilde{\Gamma}^n_{Z|X_2X_3}(Z^n\mid X_2^n,X_3^n)}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\Bigg)\Bigg|(X_1^n(1,1),X_2^n(1,1),X_3^n(w_3),Z^n)\in\mathcal{B}^n_{\boldsymbol{\theta}}\Bigg]\cdot\Pr[A]$$

$$+\mathbb{E}\left[\log\left(1+\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}+\frac{\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n\mid X_1^n,X_3^n)}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right.\right.$$
$$\left.\left.+\frac{\tilde{\Gamma}^n_{Z|X_2X_3}(z^n\mid X_2^n,X_3^n)}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right)\middle|(X_1^n(1,1),X_2^n(1,1),X_3^n(w_3),Z^n)\notin\mathcal{B}^n_{\boldsymbol{\theta}}\right]\cdot\Pr[B]. \tag{193}$$

To bound the first summand, we observe:

$$\mathbb{E}\left[\log\left(1+\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}+\frac{\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n\mid X_1^n,X_3^n)}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right.\right.$$
$$\left.\left.+\frac{\tilde{\Gamma}^n_{Z|X_2X_3}(Z^n\mid X_2^n,X_3^n)}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right)\middle|(X_1^n(1,1),X_2^n(1,1),X_3^n(w_3),Z^n)\in\mathcal{B}^n_{\boldsymbol{\theta}}\right]\cdot\Pr[A]$$
$$\stackrel{(a)}{\leq}\mathbb{E}\left[\log\left(1+\frac{e^{\theta_0}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}+\frac{e^{\theta_1}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right.\right.$$
$$\left.\left.+\frac{e^{\theta_2}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right)\middle|(X_1^n(1,1),X_2^n(1,1),X_3^n(w_3),Z^n)\in\mathcal{B}^n_{\boldsymbol{\theta}}\right]\cdot\Pr[A] \tag{194}$$
$$\leq\frac{e^{\theta_0}}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2}\mathbb{E}\left[\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}{\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right]+\frac{e^{\theta_1}}{\mathsf{M}_1\mathsf{K}_1}\mathbb{E}\left[\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}{\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right]$$
$$+\frac{e^{\theta_2}}{\mathsf{M}_2\mathsf{K}_2}\mathbb{E}\left[\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}{\tilde{\Gamma}^n_{Z|X_3}(Z^n\mid X_3^n(w_3))}\right] \tag{195}$$
$$\stackrel{(b)}{=}\frac{e^{\theta_0}}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2}\cdot\sum_{x_3^n}\sum_{z^n}P^{\otimes n}_{X_3|T}(x_3^n\mid t^n)\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|0^n,0^n,x_3^n)$$
$$+\frac{e^{\theta_1}}{\mathsf{M}_1\mathsf{K}_1}\cdot\sum_{x_3^n}\sum_{z^n}P^{\otimes n}_{X_3|T}(x_3^n\mid t^n)\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|0^n,0^n,x_3^n)$$
$$+\frac{e^{\theta_2}}{\mathsf{M}_2\mathsf{K}_2}\cdot\sum_{x_3^n}\sum_{z^n}P^{\otimes n}_{X_3|T}(x_3^n\mid t^n)\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|0^n,0^n,x_3^n) \tag{196}$$
$$\stackrel{(c)}{=}\frac{e^{\theta_0}}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2}+\frac{e^{\theta_1}}{\mathsf{M}_1\mathsf{K}_1}+\frac{e^{\theta_2}}{\mathsf{M}_2\mathsf{K}_2}, \tag{197}$$

where (a) holds by the definition of the set $\mathcal{B}^n_{\boldsymbol{\theta}}$ in (192); (b) holds because $(X_3^n(w_3),Z^n)\sim\tilde{\Gamma}^n_{Z|X_3}P^{\otimes n}_{X_3|T}$; and (c) holds because for all $t\in\mathcal{T}$ the term $P_{X_3|T}(x_3\mid t)\Gamma_{Z|X_1X_2X_3}(z\mid 0,0,x_3)$ denotes a valid probability distribution over $\mathcal{X}_3\times\mathcal{Z}$ and hence sums to 1.

To bound the second summand in (193), remark that by the definition of $\tilde{\Gamma}^n_{Z|X_3}$ in (172), for any pair (x_3^n,z^n):

$$\tilde{\Gamma}^n_{Z|X_3}(z^n\mid x_3^n)=\prod_{i=1}^n\sum_{(x_{1,i},x_{2,i})\in\mathcal{X}_1\times\mathcal{X}_2}P_{X_{1,n}}(x_{1,i}\mid t_i)P_{X_{2,n}}(x_{2,i}\mid t_i)\Gamma_{Z|X_1X_2X_3}(z|x_{1,i},x_{2,i},x_{3,i}) \tag{198}$$
$$=\prod_{i=1}^n\Big[\rho_{1,t_i}\rho_{2,t_i}\alpha_n^2\Gamma_{Z|X_1X_2X_3}(z_i|1,1,x_{3,i})+\rho_{1,t_i}\alpha_n(1-\rho_{2,t_i}\alpha_n)\Gamma_{Z|X_1X_2X_3}(z_i|1,0,x_{3,i})$$
$$+(1-\rho_{1,t_i}\alpha_n)\rho_{2,t_i}\alpha_n\Gamma_{Z|X_1X_2X_3}(z_i|0,1,x_{3,i})$$
$$+(1-\rho_{1,t_i}\alpha_n)(1-\rho_{2,t_i}\alpha_n)\Gamma_{Z|X_1X_2X_3}(z_i|0,0,x_{3,i})\Big] \tag{199}$$
$$\geq\prod_{i=1}^n(1-\rho_{1,t_i}\alpha_n)(1-\rho_{2,t_i}\alpha_n)\Gamma_{Z|X_1X_2X_3}(z_i|0,0,x_{3,j}) \tag{200}$$
$$\geq\prod_{i=1}^n(1-\rho_{1,t_i}\alpha_n)(1-\rho_{2,t_i}\alpha_n)\nabla_0. \tag{201}$$

We can thus conclude that for any $(x_1^n,x_2^n,x_3^n,z^n)\in\mathcal{X}_1^n\times\mathcal{X}_2^n\times\mathcal{X}_3^n\times\mathcal{Z}^n$ the following bound holds:

$$\log\left(1+\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|x_1^n,x_2^n,x_3^n)}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(z^n\mid x_3^n)}+\frac{\tilde{\Gamma}^n_{Z|X_1X_3}(z^n|x_1^n,x_3^n)}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^n_{Z|X_3}(z^n\mid x_3^n)}+\frac{\tilde{\Gamma}^n_{Z|X_2X_3}(z^n|x_2^n,x_3^n)}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(z^n\mid x_3^n)}\right)$$

$$= \log\left(\frac{1}{\tilde{\Gamma}^n_{Z|X_3}(z^n \mid x_3^n)}\right) + \log\Bigg(\tilde{\Gamma}^n_{Z|X_3}(z^n \mid x_3^n) + \frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|x_1^n,x_2^n,x_3^n)}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2} + \frac{\tilde{\Gamma}^n_{Z|X_1X_3}(z^n|x_1^n,x_3^n)}{\mathsf{M}_1\mathsf{K}_1} + \frac{\tilde{\Gamma}^n_{Z|X_2X_3}(z^n|x_2^n,x_3^n)}{\mathsf{M}_2\mathsf{K}_2}\Bigg) \tag{202}$$

$$\overset{(a)}{\leq} \log\left(\frac{1}{\prod_{i=1}^n(1-\rho_{1,t_i}\alpha_n)(1-\rho_{2,t_i}\alpha_n)\nabla_0}\right) + \log(4) \tag{203}$$

$$= \sum_{i=1}^n \log\left(\frac{1}{(1-\rho_{1,t_i}\alpha_n)(1-\rho_{2,t_i}\alpha_n)\nabla_0}\right) + \log(4) \tag{204}$$

$$\leq \sum_{i=1}^n \log\left(\frac{1}{(1-\rho_{1,t_i}\alpha_n)(1-\rho_{2,t_i}\alpha_n)\nabla_0}\right) + n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\log(4) \tag{205}$$

$$= \sum_{t\in\mathcal{T}} n\boldsymbol{\pi}(t)\log\left(\frac{4}{(1-\rho_{1,t}\alpha_n)(1-\rho_{2,t}\alpha_n)\nabla_0}\right). \tag{206}$$

Therefore,

$$\mathbb{E}\Bigg[\log\Bigg(1 + \frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n \mid X_3^n(w_3))} + \frac{\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n|X_1^n(1,1),X_3^n(w_3))}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n \mid X_3^n(w_3))} + \frac{\tilde{\Gamma}^n_{Z|X_2X_3}(Z^n|X_2^n(1,1),X_3^n(w_3))}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^n_{Z|X_3}(Z^n \mid X_3^n(w_3))}\Bigg)\Bigg|(X_1^n(1,1),X_2^n(1,1),X_3^n(w_3),Z^n)\notin\mathcal{B}^n_{\boldsymbol{\theta}}\Bigg]$$

$$\leq \sum_{t\in\mathcal{T}} n\boldsymbol{\pi}(t)\log\left(\frac{4}{(1-\rho_{1,t}\alpha_n)(1-\rho_{2,t}\alpha_n)\nabla_0}\right). \tag{207}$$

We proceed to analyze the probability of the event B, which by the union bound can be upper bounded as follows:

$$\Pr[B] \leq \Pr\left[\log\left(\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n \mid X_1^n(1,1),X_2^n(1,1),X_3^n(w_3))}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}\right)\geq\theta_0\right] + \Pr\left[\log\left(\frac{\tilde{\Gamma}^n_{Z|X_1X_3}(Z^n \mid X_1^n(1,1),X_3^n(w_3))}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}\right)\geq\theta_1\right] + \Pr\left[\log\left(\frac{\tilde{\Gamma}^n_{Z|X_2X_3}(Z^n \mid X_2^n(1,1),X_3^n(w_3))}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(Z^n|0^n,0^n,X_3^n(w_3))}\right)\geq\theta_2\right] \tag{208}$$

$$= \Pr\left[\sum_{i=1}^n\log\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z_i|X_{1,i},X_{2,i},X_{3,i})}{\Gamma_{Z|X_1X_2X_3}(Z_i|0,0,X_{3,i})}\right)\geq\theta_0\right] + \Pr\left[\sum_{i=1}^n\log\left(\frac{\Gamma^{(t_i)}_{Z|X_1X_3}(Z_i|X_{1,i},X_{3,i})}{\Gamma_{Z|X_1X_2X_3}(Z_i|0,0,X_{3,i})}\right)\geq\theta_1\right] + \Pr\left[\sum_{i=1}^n\log\left(\frac{\Gamma^{(t_i)}_{Z|X_2X_3}(Z_i|X_{2,i},X_{3,i})}{\Gamma_{Z|X_1X_2X_3}(Z_i|0,0,X_{3,i})}\right)\geq\theta_2\right] \tag{209}$$

To show that the quantity in (209) vanishes we will make use of Bernstein's inequality. First we notice the expectations

$$\sum_{i=1}^n\mathbb{E}\left[\log\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z_i|X_{1,i},X_{2,i},X_{3,i})}{\Gamma_{Z|X_1X_2X_3}(Z_i \mid 0,0,X_{3,i})}\right)\right]$$

$$= \sum_{t\in\mathcal{T}} n\boldsymbol{\pi}(t)\mathbb{E}\left[\log\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z \mid X_1,X_2,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z \mid 0,0,X_3)}\right)\middle| T=t\right] \tag{210}$$

$$\overset{(a)}{=} n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\mathbb{E}_{P_{X_3|T=t}}\left[\rho_{1,t}\rho_{2,t}\alpha_n^2\mathbb{D}_Z^{(1,2)}(X_3) + \rho_{1,t}\alpha_n(1-\rho_{2,t}\alpha_n)\mathbb{D}_Z^{(1)}(X_3) + (1-\rho_{1,t}\alpha_n)\rho_{2,t}\alpha_n\mathbb{D}_Z^{(2)}(X_3)\middle| T=t\right] \tag{211}$$

$$= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\mathbb{E}_{P_{X_3|T=t}}\left[\rho_{1,t}\alpha_n\mathbb{D}_Z^{(1)}(X_3) + \rho_{2,t}\alpha_n\mathbb{D}_Z^{(2)}(X_3)\middle| T=t\right] + n\mathcal{O}(\alpha_n^2) \tag{212}$$

$$= n\alpha_n\mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3) + \rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right] + n\mathcal{O}(\alpha_n^2), \tag{213}$$

where the first expectation is taken with respect to the law $P_{X_{1,n}|T=t}P_{X_{2,n}|T=t}P_{X_3|T=t}\Gamma_{Z|X_1X_2X_3}$. Also, Equality (a) holds because when $(x_1,x_2)=(0,0)$, the log term is zero.

In a similar way, we have:

$$\sum_{i=1}^n\mathbb{E}\left[\log\left(\frac{\Gamma_{Z|X_1X_3}(Z_i|X_{1,i},X_{3,i})}{\Gamma_{Z|X_1X_2X_3}(Z_i \mid 0,0,X_{3,i})}\right)\right]$$

$$= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\,\mathbb{E}\left[\log\left(\frac{\sum_{x_2}P_{X_2|T}(x_2\mid t)\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,x_2,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\right)\middle| T=t\right] \tag{214}$$

$$= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\,\mathbb{E}\left[\log\left(\frac{\rho_{2,t}\alpha_n\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,1,X_3)+(1-\rho_{2,t}\alpha_n)\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\right)\middle| T=t\right] \tag{215}$$

$$= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\,\mathbb{E}\left[\log\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\cdot\left(1+\rho_{2,t}\alpha_n\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,1,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}-1\right)\right)\right)\middle| T=t\right] \tag{216}$$

$$\begin{aligned}&= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\,\mathbb{E}\left[\log\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\middle| T=t\right]\\ &\quad+n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\;\mathbb{E}\left[\log\left(1+\rho_{2,t}\alpha_n\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,1,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}-1\right)\right)\middle| T=t\right]\end{aligned} \tag{217}$$

where the expectations are with respect to the law $P_{X_{1,n}|T=t}P_{X_3|T=t}\Gamma^{(t)}_{Z|X_1X_3}$. For the first summand, we have:

$$\begin{aligned}&n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\,\mathbb{E}\left[\log\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\middle| T=t\right]\\ &= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\rho_{1,t}\alpha_n(1-\rho_{2,t}\alpha_n)\underset{P_{X_3|T=t}\Gamma_{Z|X_1=1,X_2=0,X_3}}{\mathbb{E}}\left[\log\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid 1,0,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\right)\middle| T=t\ \right]+\mathcal{O}(\alpha_n^2) &(218)\\ &= \alpha_n n\mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3)\right]+\mathcal{O}(\alpha_n^2), &(219)\end{aligned}$$

because the logarithmic term vanishes when $x_1=0$ and by using the definition of the distribution $\Gamma^{(t)}_{Z|X_1X_3}$. For the second summand, we use the upper bound $\log(1+x)\leq x-\frac{x^2}{2}$ to obtain:

$$\begin{aligned}&n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\;\mathbb{E}_{P_{X_{1,n}|T=t}P_{X_3|T=t}\Gamma^{(t)}_{Z|X_1X_3}}\left[\log\left(1+\rho_{2,t}\alpha_n\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,1,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}-1\right)\right)\middle| T=t\right]\\ &\leq n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\rho_{2,t}\alpha_n\mathbb{E}\left[\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,1,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}-1\middle| T=t\right]+\mathcal{O}(\alpha_n^2) &(220)\\ &= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\rho_{2,t}\alpha_n\mathbb{E}_{P_{X_{1,n}|T=t}P_{X_{3,t}}\Gamma_{Z|X_1,X_2=0,X_3}}\left[\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,1,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,0,X_3)}-1\middle| T=t\ \right]+\mathcal{O}(\alpha_n^2) &(221)\\ &= n\sum_{t\in\mathcal{T}}\boldsymbol{\pi}(t)\rho_{2,t}\alpha_n\mathbb{E}_{P_{X_{1,n}|T=t}P_{X_{3,t}}\Gamma_{Z|X_1,X_2=0,X_3}}\left[\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,1,X_3)-1\middle| T=t\ \right]+\mathcal{O}(\alpha_n^2) &(222)\\ &= +\mathcal{O}(\alpha_n^2), &(223)\end{aligned}$$

where the last equation holds because the expectation $\mathbb{E}_{\Gamma_{Z|X_1=x_1,X_2=0,X_3}}[\Gamma_{Z|X_1X_2X_3}(Z\mid x_1,1,X_3)]=1$ for any value of x_1.

Combining (217), (219), and (223), we obtain:

$$\sum_{i=1}^{n}\mathbb{E}\left[\log\left(\frac{\Gamma_{Z|X_1X_3}(Z_i|X_{1,i},X_{3,i})}{\Gamma_{Z|X_1X_2X_3}(Z_i\mid 0,0,X_{3,i})}\right)\right]\leq\alpha_n n\mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3)\right]+\mathcal{O}(\alpha_n^2), \tag{224}$$

Likewise,

$$\sum_{i=1}^{n}\mathbb{E}\left[\log\left(\frac{\Gamma_{Z|X_2X_3}(Z_i|X_{2,i},X_{3,i})}{\Gamma_{Z|X_1X_2X_3}(Z_i\mid 0,0,X_{3,i})}\right)\right]\leq\alpha_n n\mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right]+\mathcal{O}(\alpha_n^2). \tag{225}$$

Then, notice that the variances satisfy:

$$\mathbb{E}\left[\log^2\left(\frac{\Gamma_{Z|X_1X_2X_3}(Z\mid X_1,X_2,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\right)\right]\overset{(a)}{=}\mathcal{O}(\alpha_n), \tag{226}$$

and,

$$\mathbb{E}\left[\log^2\left(\frac{\Gamma_{Z|X_1X_3}(Z\mid X_1,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\right)\right]\overset{(b)}{\leq}\mathcal{O}(\alpha_n), \tag{227}$$

and

$$\mathbb{E}\left[\log^2\left(\frac{\Gamma_{Z|X_2X_3}(Z\mid X_2,X_3)}{\Gamma_{Z|X_1X_2X_3}(Z\mid 0,0,X_3)}\right)\right]\overset{(c)}{\leq}\mathcal{O}(\alpha_n), \tag{228}$$

where similarly to (157), (a) follows because when $(x_1, x_2) = (0,0)$, the log term is zero, whereas (b) and (c) follows by first splitting the log term and using the Taylor expansion and then using the same argument as for (a). Since $\lim_{n\to\infty} \boldsymbol{\pi}(t) \to P_T(t)$ and $\lim_{n\to\infty} \alpha_n^2 = 0$, Bernstein's inequality allows us to conclude that with the choices

$$\theta_0 \triangleq (1+\xi_4)\alpha_n n \mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3) + \rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right], \tag{229a}$$

$$\theta_1 \triangleq (1+\xi_5)\alpha_n n \mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3)\right], \tag{229b}$$

$$\theta_2 \triangleq (1+\xi_6)\alpha_n n \mathbb{E}_{P_{TX_3}}\left[\rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right], \tag{229c}$$

for any $\xi_4 > 0$, $\xi_5 > 0$, $\xi_6 > 0$ there exists a constant $B_1 > 0$, $B_2 > 0$ and $B_3 > 0$ so that for sufficiently large blocklengths n:

$$\Pr[B] \le e^{-B_1 n\alpha_n} + e^{-B_2 n\alpha_n} + e^{-B_3 n\alpha_n} \tag{230a}$$

Plugging (230) into (209) and further combining it with (193), (197), and (207), we finally obtain

$$\begin{aligned}&\mathop{\mathbb{E}}_{\{X_1^n w_1', s_1')\}_{w_1', s_1'}), \{X_2^n w_2', s_2')\}_{w_2', s_2'}), X_3^n(w_3)}\left[\mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^n_{Z|X_3}(\cdot \mid X_3^n(w_3))\right)\right]\\ &\le \frac{e^{\theta_0}}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2} + \frac{e^{\theta_1}}{\mathsf{M}_1\mathsf{K}_1} + \frac{e^{\theta_2}}{\mathsf{M}_2\mathsf{K}_2} + n\mathbb{E}_T\left[\log\left(\frac{4}{(1-\rho_{1,T}\alpha_n)(1-\rho_{2,T}\alpha_n)\nabla_0}\right)\right]\cdot\left(e^{-B_1 n\alpha_n} + e^{-B_2 n\alpha_n} + e^{-B_3 n\alpha_n}\right)\end{aligned} \tag{231}$$

We notice that the second summand tends to 0 because ne^{-na} decays for any positive $a > 0$. If moreover,

$$\limsup_{n\to\infty}\left(\log(\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2) - (1+\xi_4)\alpha_n n\mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3) + \rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right]\right) = -\infty, \tag{232a}$$

$$\limsup_{n\to\infty}\left(\log(\mathsf{M}_1\mathsf{K}_1) - (1+\xi_5)\alpha_n n\mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3)\right]\right) = -\infty, \tag{232b}$$

$$\limsup_{n\to\infty}\left(\log(\mathsf{M}_2\mathsf{K}_2) - (1+\xi_6)\alpha_n n\mathbb{E}_{P_{TX_3}}\left[\rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right]\right) = -\infty, \tag{232c}$$

then also the first summand of (231) tends to 0 exponentially fast.

Notice that for small values ξ_4, ξ_5, ξ_6 and large blocklengths n Constraint (232a) is redundant in view of the per-user secret-key constraints (232b) and (232c).

Analysis of divergence $\mathbb{D}(\tilde{\Gamma}^n_{Z|X_3}(\cdot \mid x_3^n(w_3))\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n, 0^n, x_3^n(w_3)))$*:*

Recall that $\tilde{\Gamma}^n_{Z|X_3}$ and $\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n, 0^n, x_3^n(w_3))$ are both product distributions and thus

$$\begin{aligned}&\mathbb{D}(\tilde{\Gamma}^n_{Z|X_3}(\cdot \mid x_3^n(w_3))\|\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0,0,x_3^n(w_3)))\\ &= \sum_{t\in\mathcal{T}}\sum_{\substack{i\in\{1,\ldots,t\}:\\ t_i=t}} \mathbb{D}(\Gamma^{(t)}_{Z|X_3}(z \mid x_3)\|\Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_{3,i}(w_3))) \qquad (233)\\ &= \sum_{(t,x_3)\in\mathcal{T}}\sum_{\substack{i\in\{1,\ldots,t\}:\\ t_i=t\\ x_{3,i}(t)=x_3}} \mathbb{D}(\Gamma^{(t)}_{Z|X_3}(z \mid x_3)\|\Gamma_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3)) \qquad (234)\\ &\overset{(a)}{=} \sum_{(x_3,t)\in\mathcal{X}_3\times\mathcal{T}} n\lambda_t(x_3)\,(1+o(1))\cdot\frac{(\rho_{1,t}\alpha_n + \rho_{2,t}\alpha_n)^2}{2}\chi^2(\rho_{1,t},\rho_{2,t},x_3)\\ &\overset{(b)}{=} \sum_{(x_3,t)\in\mathcal{X}_3\times\mathcal{T}} \lambda_t(x_3)\,(1+o(1))\cdot\frac{(\rho_{1,t}+\rho_{2,t})^2\cdot\omega_n^2}{2}\chi^2(\rho_{1,t},\rho_{2,t},x_3), \qquad (235)\end{aligned}$$

where in (a) we used the definition

$$\lambda_t(x_3) \triangleq \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{x_{3,j}(w_3) = x_3, t_i = t\} \tag{236}$$

and we applied Lemma 3 for each value of t individually, also using the fact

$$\chi^2(a,b,x_3) = \chi^2\left(\frac{a}{a+b}, \frac{b}{a+b}, x_3\right); \tag{237}$$

and in (b) we used the definition of α_n.

Since both $P_{e,0}$ and $P_{e,1}$ vanish as $n \to \infty$, and by the decoding rule in (40), we can conclude that the sequence of codes in Theorem 1 satisfies for each $(x_3, t) \in \mathcal{X}_3 \times \mathcal{T}$

$$\lim_{n\to\infty} \left|\lambda_t(x_3) - P_T(t)P_{X_3|T}(x_3 \mid t)\right| = 0. \tag{238}$$

Concluding the Divergence Proof: With this observation, combining (246) with (231) and (235), under Condition (231), we obtain:

$$\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3} = (1+o(1))\frac{\omega_n^2}{2}\mathbb{E}_{P_T}\left[(\rho_{1,T}+\rho_{2,T})^2\mathbb{E}_{P_{X_3|T}}\left[\chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]\right]. \tag{239}$$

C. Concluding the Achievability Proof:

By standard averaging arguments it can then be shown that there must exist at least one sequence of codebooks $\{\mathcal{C}_n\}_n$ for which the probabilities of error under the two hypotheses tend to 0 as $n \to \infty$ and with message and secret-key sizes as well as divergence satisfying

$$\log \mathsf{M}_1 = (1-\xi_1)\alpha_n n\mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Y^{(1)}(X_3)\right], \tag{240}$$

$$\log \mathsf{M}_2 = (1-\xi_2)\alpha_n n\mathbb{E}_{P_{TX_3}}\left[\rho_{2,T}\mathbb{D}_Y^{(2)}(X_3)\right], \tag{241}$$

$$\log \mathsf{M}_3 = (1-\xi_3)nI(X_3;Y|X_1=0,X_2=0,T), \tag{242}$$

$$\log(\mathsf{M}_1\mathsf{K}_1) = (1+\xi_5)\alpha_n n\mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3)\right], \tag{243}$$

$$\log(\mathsf{M}_2\mathsf{K}_2) = (1+\xi_6)\alpha_n n\mathbb{E}_{P_{TX_3}}\left[\rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right], \tag{244}$$

$$\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3} = (1+o(1))\frac{\omega_n^2}{2}\mathbb{E}_{P_T}\left[(\rho_{1,T}+\rho_{2,T})^2\mathbb{E}_{P_{X_3|T}}\left[\chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]\right]. \tag{245}$$

APPENDIX B
MODIFICATIONS FOR THE GENERALIZED SCHEME

For simplicity of notation we assume that $\cup_t\mathcal{L}_{1,2}(t) = \{1,\ldots,n_2\}$ and $\cup_t\mathcal{L}_1(t) = \{n_2+1,\ldots,n_1\}$. We denote the first n_2 symbols of the corresponding codewords by $x_1^{n_2}(w_1,s_1)$, $x_2^{n_2}(w_2,s_2)$, and $x_3^{n_2}(w_3)$ and the following n_1-n_2 symbols of the corresponding codewords by $x_{1,n_2+1}^{n_1}(w_1,s_1)$ and $x_{3,n_2+1}^{n_1}(w_3)$.

It is easy to observe that the divergence term δ_{n,w_3} now only depends on the first n_1 channel uses, as the terms corresponding to the last $n-n_1$ channel uses are zero. In analogy to (180) we can then obtain the bound (notice the new blocklength n_1 instead of n):

$$\begin{aligned}&\left|\mathbb{D}\left(\widehat{Q}_{\mathcal{C},w_3}^{n_1}\middle\|\Gamma_{Z|X_1X_2X_3}^{\otimes n_1}(\cdot|0^{n_1},0^{n_1},x_3^{n_1}(w_3))\right) - \mathbb{D}\left(\tilde{\Gamma}_{Z|X_3}^{n_1}\middle\|\Gamma_{Z|X_1X_2X_3}^{\otimes n_1}(\cdot|0^{n_1},0^{n_1},x_3^{n_1}(w_3))\right)\right|\\ &\le \mathbb{D}\left(\widehat{Q}_{\mathcal{C},w_3}^{n_1}\middle\|\tilde{\Gamma}_{Z|X_3}^{n_1}\right) + n_1\log\left(\frac{1}{\nabla_0}\right)\sqrt{\frac{1}{2}\mathbb{D}\left(\widehat{Q}_{\mathcal{C},w_3}^{n_1}\middle\|\tilde{\Gamma}_{Z|X_3}^{n_1}\right)},\end{aligned} \tag{246}$$

where $x_3^{n_1}(w_3)$ denotes the first n_1 symbols of codeword $x_3^n(w_3)$ and $\widehat{Q}_{\mathcal{C},w_3}^{n_1}$ denotes the pmf of the warden's first n_1 output symbols. For the generalized scheme we have:

$$\widehat{Q}_{\mathcal{C},w_3}^{n_1}(z^{n_1}) \triangleq \frac{1}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2}\sum_{(w_1,s_1)}\sum_{(w_2,s_2)}\widehat{Q}_{\mathcal{C},w_1,w_2,w_3,s_1,s_2}^{n_1}(z^{n_1}), \tag{247}$$

where for any valid $(w_1,w_2,w_3,s_1,s_2) \in \mathcal{M}_1\times\mathcal{M}_2\times\mathcal{M}_3\times\mathcal{K}_1\times\mathcal{K}_2$, we have:

$$\begin{aligned}\widehat{Q}_{\mathcal{C},w_1,w_2,w_3,s_1,s_2}^{n_1}(z^{n_1}) &\triangleq \Gamma_{Z|X_1X_2X_3}^{\otimes n_2}(z^{n_2}|x_1^{n_2}(w_1,s_1),x_2^{n_2}(w_2,s_2),x_3^{n_2}(w_3))\\ &\quad\cdot\tilde{\Gamma}_{Z|X_1X_3}^{(n_2\to n_1)}(z_{n_2+1}^{n_1}|x_{1,n_2+1}^{n_1}(w_1,s_1),x_{3,n_2+1}^{n_1}(w_3)),\end{aligned} \tag{248}$$

for $z_{n_2+1}^{n_1} \triangleq (z_{n_2+1},\ldots,z_{n-1})$ and for $\tilde{\Gamma}_{Z|X_1X_3}^{(n_2\to n_1)}$ defined in analogy to $\tilde{\Gamma}_{Z|X_1X_3}^{n}$ in (174) but based on the sequence $t_{n_2+1},\ldots,t_{n_1}$:

$$\tilde{\Gamma}_{Z|X_1X_3}^{(n_2\to n_1)}(z_{n_2+1}^{n_1}|x_{1,n_2+1}^{n_1}(w_1,s_1),x_{3,n_2+1}^{n_1}(w_3)) \triangleq \prod_{i=n_2+1}^{n_1}\Gamma_{Z|X_1X_3}^{(t_i)}(z_i\mid x_{1,i},x_{3,i}). \tag{249}$$

Following similar steps as in (233)–(239), one can bound the second divergence on the left-hand side of (246) as:

$$\mathbb{D}\left(\tilde{\Gamma}_{Z|X_3}^{n_1}\middle\|\Gamma_{Z|X_1X_2X_3}^{\otimes n_1}(\cdot|0^{n_1},0^{n_1},x_3^{n_1}(w_3))\right) = (1+o(1))\frac{n_1}{n}\cdot\frac{\omega_n^2}{2}\mathbb{E}_{P_T}\left[(\rho_{1,T}+\rho_{2,T})^2\mathbb{E}_{P_{X_3|T}}\left[\chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]\right] \tag{250}$$

To bound the divergence term on the right-hand side of (246), we slightly modify the steps in (182)–(197). In particular, we write:

$$\mathop{\mathbb{E}}_{\mathcal{C}}\left[\mathbb{D}\left(\widehat{Q}^{n_1}_{\mathcal{C},w_3}\middle\|\tilde{\Gamma}^{n_1}_{Z|X_3}(\cdot\mid X_3^{n_1}(w_3))\right)\right]$$

$$= \mathbb{E}\left[\log\left(\sum_{(w_1,w_2,s_1,s_2)}\mathop{\mathbb{E}}_{\substack{\{X_1^{n_1}(w_1,s_1)\}\setminus X_1^{n_1}(1,1),\\ \{X_2^{n_2}(w_2,s_2)\}\setminus X_2^{n_2}(1,1)}}\left[\frac{\widehat{Q}^{n_1}_{\mathcal{C},w_1,w_2,w_3,s_1,s_2}(Z^{n_1})}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}\right]\right)\right] \tag{251}$$

$$\leq \mathbb{E}\Bigg[\log\Bigg(\sum_{\substack{(w_1,s_1)\neq(1,1)\\(w_2,s_2)\neq(1,1)}}\mathop{\mathbb{E}}_{\substack{X_1^{n_1}(w_1,s_1)\\X_2^{n_2}(w_2,s_2)}}\left[\frac{\widehat{Q}^{n_1}_{\mathcal{C},w_1,w_2,w_3,s_1,s_2}(Z^{n_1})}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}\right]+\frac{\widehat{Q}^{n_1}_{\mathcal{C},1,1,w_3,1,1}(Z^{n_1})}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}$$
$$+\sum_{(w_2,s_2)\neq(1,1)}\mathop{\mathbb{E}}_{X_2^{n_2}(w_2,s_2)}\left[\frac{\widehat{Q}^{n_1}_{\mathcal{C},1,w_2,w_3,1,s_2}(Z^{n_1})}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}\right]$$
$$+\sum_{(w_1,s_1)\neq(1,1)}\mathop{\mathbb{E}}_{X_1^{n_1}(w_1,s_1)}\left[\frac{\widehat{Q}^{n_1}_{\mathcal{C},w_1,1,w_3,s_1,1}(Z^{n_1})}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}\right]\Bigg)\Bigg] \tag{252}$$

$$\overset{(a)}{=} \mathbb{E}\Bigg[\log\Bigg(\frac{(\mathsf{M}_1\mathsf{K}_1-1)(\mathsf{M}_2\mathsf{K}_2-1)}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2}+\frac{\widehat{Q}^{n_1}_{\mathcal{C},1,1,w_3,1,1}(Z^{n_1})}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}$$
$$+\frac{(\mathsf{M}_2\mathsf{K}_2-1)\tilde{\Gamma}^{n_1}_{Z|X_1X_3}(Z^{n_1}|X_1^{n_1}(1,1),X_3^{n_1}(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}+\frac{(\mathsf{M}_1\mathsf{K}_1-1)\tilde{\Gamma}^{n_2}_{Z|X_2X_3}(Z^{n_2}|X_2^{n_2}(1,1),X_3^{n_2}(w_3))}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_2}_{Z|X_3}(Z^{n_2}\mid X_3^{n_2}(w_3))}\Bigg)\Bigg] \tag{253}$$

$$\leq \mathbb{E}\Bigg[\log\Bigg(1+\frac{\widehat{Q}^{n_1}_{\mathcal{C},1,1,w_3,1,1}(Z^{n_1})}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}$$
$$+\frac{\tilde{\Gamma}^{n_1}_{Z|X_1X_3}(Z^{n_1}|X_1^{n_1}(1,1),X_3^{n_1}(w_3))}{\mathsf{M}_1\mathsf{K}_1\cdot\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}+\frac{\tilde{\Gamma}^{n_2}_{Z|X_2X_3}(Z^{n_2}|X_2^{n_2}(1,1),X_3^{n_2}(w_3))}{\mathsf{M}_2\mathsf{K}_2\cdot\tilde{\Gamma}^{n_2}_{Z|X_3}(Z^{n_2}\mid X_3^{n_2}(w_3))}\Bigg)\Bigg] \tag{254}$$

where in (a) we used that

$$\mathop{\mathbb{E}}_{\substack{X_1^{n_1}(w_1,s_1)\\X_2^{n_2}(w_2,s_2)}}\left[\widehat{Q}^{n_1}_{\mathcal{C},w_1,w_2,w_3,s_1,s_2}(z^{n_1})\right]=\tilde{\Gamma}^{n_1}_{Z|X_3}(z^{n_1}|X_3^{n_1}(w_3)) \tag{255}$$

and

$$\mathop{\mathbb{E}}_{X_1^{n_1}(w_1,s_1)}\left[\widehat{Q}^{n_1}_{\mathcal{C},w_1,1,w_3,s_1,1}(z^{n_2})\right]=\tilde{\Gamma}^{n_2}_{Z|X_2X_3}(z^{n_2}|X_2^{n_2}(1,1),X_3^{n_2}(w_3))\cdot\tilde{\Gamma}^{n_2\to n_1}_{Z|X_3}(z^{n_1}_{n_2+1}|X^{n_1}_{3,n_2+1}(w_3)) \tag{256}$$

$$\mathop{\mathbb{E}}_{X_2^{n_2}(w_2,s_2)}\left[\widehat{Q}^{n_1}_{\mathcal{C},1,w_2,w_3,1,s_2}(z^{n_1})\right]=\tilde{\Gamma}^{n_1}_{Z|X_1X_3}(z^{n_1}|X_1^{n_1}(1,1),X_3^{n_1}(w_3)) \tag{257}$$

and we simplified the last fraction by noting that

$$\frac{\tilde{\Gamma}^{n_2}_{Z|X_2X_3}(z^{n_2}|X_2^{n_2}(1,1),X_3^{n_2}(w_3))\cdot\tilde{\Gamma}^{n_2\to n_1}_{Z|X_3}(z^{n_1}_{n_2+1}|X^{n_1}_{3,n_2+1}(w_3))}{\tilde{\Gamma}^{n_1}_{Z|X_3}(Z^{n_1}\mid X_3^{n_1}(w_3))}=\frac{\tilde{\Gamma}^{n_2}_{Z|X_2X_3}(Z^{n_2}|X_2^{n_2}(1,1),X_3^{n_2}(w_3))}{\tilde{\Gamma}^{n_2}_{Z|X_3}(Z^{n_2}\mid X_3^{n_2}(w_3))}. \tag{258}$$

Define now for any triple $\boldsymbol{\theta}=(\theta_0,\theta_1,\theta_2)$ the set:[3]

$$\mathcal{B}^{n_1}_{\boldsymbol{\theta}}\triangleq\Bigg\{(x_1^{n_1},x_2^{n_2},x_3^{n_1},z^{n_1})\in\mathcal{X}_1^{n_1}\times\mathcal{X}_2^{n_2}\times\mathcal{X}_3^{n_1}\times\mathcal{Z}^{n_1}\colon$$
$$\log\left(\frac{\Gamma^{\otimes n_2}_{Z|X_1X_2X_3}(z^{n_2}\mid x_1^{n_2},x_2^{n_2},x_3^{n_2})\cdot\tilde{\Gamma}^{n_2\to n_1}_{Z|X_1X_3}(z^{n_1}_{n_2+1}\mid x^{n_1}_{1,n_2+1},x^{n_1}_{3,n_2+1})}{\Gamma^{\otimes n_1}_{Z|X_1X_2X_3}(z^{n_1}\mid 0^{n_1},0^{n_1},x_3^{n_1})}\right)\leq\theta_0,$$

[3]Notice that $x_2^{n_2}$ is of length n_2 while the other sequences are of length n_1.

$$\log\left(\frac{\tilde{\Gamma}^{n_1}_{Z|X_1X_3}(z^{n_1} \mid x_1^{n_1}, x_3^{n_1})}{\Gamma^{\otimes n_1}_{Z|X_1X_2X_3}(z^{n_1} \mid 0^{n_1}, 0^{n_1}, x_3^{n_1})}\right) \leq \theta_1,$$
$$\log\left(\frac{\tilde{\Gamma}^{n_2}_{Z|X_2X_3}(z^{n_2} \mid x_2^{n_2}, x_3^{n_2})}{\Gamma^{\otimes n_2}_{Z|X_1X_2X_3}(z^{n_2} \mid 0^{n_2}, 0^{n_2}, x_3^{n_2})}\right) \leq \theta_2 \Bigg\}, \tag{259}$$

and denote by A and B the events $\{(X_1^{n_1}(1,1), X_2^{n_2}(1,1), X_3^{n_1}(w_3), Z^{n_1}) \in \mathcal{B}_{\boldsymbol{\theta}}^{n_1}\}$ and $\{(X_1^{n_1}(1,1), X_2^{n_2}(1,1), X_3^{n_1}(w_3), Z^{n_1}) \notin \mathcal{B}_{\boldsymbol{\theta}}^{n_1}\}$ respectively.

We then continue by similar steps to (193)–(231), where the superscript n to has to be changed to n_1 or n_2 accordingly, and the joint law $\Gamma^{\otimes n}_{Z|X_1X_2X_3}$ to $\Gamma^{\otimes n_2}_{Z|X_1X_2X_3} \cdot \tilde{\Gamma}^{n_2-n_1}_{Z|X_1X_3}$. In particular we apply Bernstein's inequality to analyze the probability of set B, and replace (210) –(225) by the following expresssions:

$$\mathbb{E}\left[\log\left(\frac{\Gamma^{\otimes n_2}_{Z|X_1X_2X_3}(Z^{n_2} \mid X_1^{n_2}(1,1), X_2^{n_2}(1,1), X_3^{n_2}(w_3)) \cdot \tilde{\Gamma}^{n_2\to n_1}_{Z|X_1X_3}(Z^{n_1}_{n_2+1} \mid X^{n_1}_{1,n_2+1}(1,1), X^{n_1}_{3,n_2+1})(w_3))}{\Gamma^{\otimes n_1}_{Z|X_1X_2X_3}(Z^{n_1} \mid 0^{n_1}, 0^{n_1}, X_3^{n_1}(w_3))}\right)\right]$$
$$\leq \alpha_n \left(n_2 \mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3) + \rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right] + (n_1-n_2)\mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3))\right]\right) + n\mathcal{O}(\alpha_n^2) \tag{260a}$$

and

$$\mathbb{E}\left[\log\left(\frac{\tilde{\Gamma}^{n_1}_{Z|X_1X_3}(Z^{n_1} \mid X_1^{n_1}(1,1), X_3^{n_1}(w_3))}{\Gamma^{\otimes n_1}_{Z|X_1X_2X_3}(Z^{n_1} \mid 0^{n_1}, 0^{n_1}, X_3^{n_1}(w_3))}\right)\right] \leq \alpha_n n_1 \mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3)\right] + \mathcal{O}(\alpha_n^2) \tag{260b}$$

$$\mathbb{E}\left[\log\left(\frac{\tilde{\Gamma}^{n_2}_{Z|X_2X_3}(Z^{n_2} \mid X_2^{n_2}(1,1), X_3^{n_2}(w_3))}{\Gamma^{\otimes n_2}_{Z|X_1X_2X_3}(Z^{n_2} \mid 0^{n_2}, 0^{n_2}, X_3^{n_2}(w_3))}\right)\right] \leq \alpha_n n_2 \mathbb{E}_{\pi P_{X_3|T}}\left[\rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right] + \mathcal{O}(\alpha_n^2). \tag{260c}$$

These steps allow us to conclude that whenever

$$\limsup_{n\to\infty}\left(\log(\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2) - (1+\xi_4)\alpha_n \mathbb{E}_{P_{TX_3}}\left[n_1\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3) + n_2\rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right]\right) = -\infty, \tag{261a}$$

$$\limsup_{n\to\infty}\left(\log(\mathsf{M}_1\mathsf{K}_1) - (1+\xi_5)\alpha_n n_1 \mathbb{E}_{P_{TX_3}}\left[\rho_{1,T}\mathbb{D}_Z^{(1)}(X_3)\right]\right) = -\infty, \tag{261b}$$

$$\limsup_{n\to\infty}\left(\log(\mathsf{M}_2\mathsf{K}_2) - (1+\xi_6)\alpha_n n_2 \mathbb{E}_{P_{TX_3}}\left[\rho_{2,T}\mathbb{D}_Z^{(2)}(X_3)\right]\right) = -\infty, \tag{261c}$$

then the divergence on the right-hand side of (246) tends to 0 exponentially fast, and thus the approximation

$$\mathbb{D}\left(\tilde{\Gamma}^{n_1}_{Z|X_3}\middle\|\Gamma^{\otimes n_1}_{Z|X_1X_2X_3}(\cdot|0^{n_1}, 0^{n_1}, x_3^{n_1}(w_3))\right) = (1+o(1))\frac{n_1}{n}\frac{\omega_n^2}{2}\mathbb{E}_{P_T}\left[(\rho_{1,T}+\rho_{2,T})^2 \mathbb{E}_{P_{X_3|T}}\left[\chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]\right] \tag{262}$$

is exponentially tight. We notice that in an asymptotic sense condition (261a) is redundant in view of (261b) and (261c). This concludes the resolvability analysis by considering that $n_1 \approx n\phi_1$ and $n_2 \approx n\phi_2$ and that $\alpha_n = \omega_n\sqrt{n}$.

Appendix C
Achievability Proof to Theorem 2

Start by noticing that without loss in generality in Theorem 1 one can replace Constraint (47) by the difference between constraints (47) and (50) and Constraint (48) by the difference between constraints (48) and (51). Taking $n \to \infty$ with these new constraints proves achievability of the following quintuple $(r_1, r_2, R_3, k_1, k_2)$ for arbitrary pmfs P_{TX_3}, nonnegative tuples $\{(\rho_{1,t}, \rho_{2,t})\}_{t\in\mathcal{T}}$, and pairs $(\phi_1, \phi_2) \in [0,1]^2$:

$$r_\ell = \frac{\phi_\ell}{\sqrt{\max(\phi_1;\phi_2)}}\sqrt{2}\frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T}\mathbb{D}_Y^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T}+\rho_{2,T})^2 \cdot \chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]}}, \qquad \forall \ell \in \{1,2\}, \tag{263}$$

$$R_3 = \mathbb{I}(X_3; Y \mid X_1 = 0, X_2 = 0, T), \tag{264}$$

$$k_\ell = \frac{\phi_\ell}{\sqrt{\max(\phi_1;\phi_2)}}\sqrt{2}\frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T}\left(\mathbb{D}_Z^{(\ell)}(X_3) - \mathbb{D}_Y^{(\ell)}(X_3)\right)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T}+\rho_{2,T})^2 \cdot \chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]}}, \qquad \forall \ell \in \{1,2\}, \tag{265}$$

where we use the definition $0/0 = 0$. Define $\beta_\ell \triangleq \frac{\phi_\ell}{\sqrt{\max(\phi_1;\phi_2)}}$ and notice that it lies in $[0,1]$. Notice further that the equalities on the message rates r_1, r_2, R_3 can be relaxed into $\leq$-inequalities because it is always possible to reduce the message rate by introducing dummy data-bits. Moreover, it is possible to relax the inequalities on the secret-key rates k_1 and k_2 into $\geq$-inequalities because it is always possible to ignore some of the secret-key bits. These latter observations establish achievability of the theorem.

Appendix D
Converse Proof to Theorem 2

The converse proof relies on elements from the proofs of [4, Theorem 3], [5, Theorem 1 and 2] , [22, Theorem 2] and [23, Section 5.2.3].

Consider a sequence of length-n codes with vanishing probability of error $P_{e,\mathcal{H}} \to 0$ and vanishing covertness constraints $\delta_{n,w_3} \to 0$ as the blocklength $n \to \infty$. Consider now a fixed blocklength n, and let X_1^n, X_2^n, X_3^n be the random inputs generated under the chosen codes and Y^n as well as Z^n the corresponding outputs at the legitimate receiver and the warden under $\mathcal{H} = 1$. Define also a random variable T to be uniform over the set of channel uses $[|1, n|]$, independent of the inputs X_1^n, X_2^n, X_3^n and the outputs Y^n and Z^n. With these definitions, since the three users have independent messages and keys, the joint pmf of the time-averaged inputs and outputs has the following form:

$$P_{X_{1,T},X_{2,T},X_{3,T},Y_T,Z_T,T}(x_1,x_2,x_3,y,z,t) = P_T(t)P_{X_1|T}(x_1 \mid t)P_{X_2|T}(x_2 \mid t)P_{X_3|T}(x_3 \mid t)\Gamma_{YZ|X_1X_2X_3}(y,z \mid x_1,x_2,x_3). \tag{266}$$

For our converse proof, we also define $\alpha_{n,i,\ell}$ as the probability of $X_{\ell,i}$ equal 1:

$$\alpha_{n,i,\ell} \triangleq \Pr[X_{\ell,i} = 1], \quad i \in \{1,\dots,n\},\ \ell \in \{1,2\}, \tag{267}$$

and the derived positive quantities

$$\rho_{n,i,\ell} \triangleq \frac{\alpha_{n,i,\ell}}{\sum_{i=1}^n \frac{\alpha_{n,i,1}+\alpha_{n,i,2}}{n}}, \quad i \in \{1,\dots,n\},\ \ell \in \{1,2\}. \tag{268}$$

We observe that by the uniform law of T:

$$\mathbb{E}_{P_T}\left[\rho_{n,T,1} + \rho_{n,T,2}\right] = \frac{1}{n}\sum_{i=1}^n n \frac{\alpha_{n,i,1}+\alpha_{n,i,2}}{\sum_{i=1}^n \alpha_{n,i,1}+\alpha_{n,i,2}} = 1. \tag{269}$$

A. Auxiliary Lemmas:

We will make use of the following lemma, which is an extension of [4, Lemma 1]. Recall the definitions in (30).

Lemma 4. *Let X_1, X_2 be binary over $\{0,1\}$ and T, X_3, Y over arbitrary finite alphabets $\mathcal{T}$, $\mathcal{X}_3$, and $\mathcal{Y}$, with joint pmf of the form $P_T P_{X_1|T} P_{X_3|T} \Gamma_{Y|X_1,X_2,X_3}$. Then, for any $x_2 \in \{0,1\}$ and $t \in \mathcal{T}$:*

$$\begin{aligned}\mathbb{I}(X_1;Y \mid X_2 = x_2, X_3, T = t) = P_{X_1|T=t}(1)\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}(\Gamma_{Y|X_1X_2X_3}(\cdot \mid 1,x_2,X_3)||\Gamma_{Y|X_1X_2X_3}(\cdot|0,0,X_3))\right]\\ -\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}(\Gamma_{Y|X_2X_3}(\cdot \mid x_2,X_3)||\Gamma_{Y|X_1X_2X_3}(\cdot|0,0,X_3))\right].\end{aligned} \tag{270}$$

Similarly, for any $x_1 \in \{0,1\}$ and $t \in \mathcal{T}$:

$$\begin{aligned}\mathbb{I}(X_2;Y \mid X_1 = x_1, X_3, T = t) = P_{X_2|T=t}(1)\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}(\Gamma_{Y|X_1X_2X_3}(\cdot \mid x_1,1,X_3)||\Gamma_{Y|X_1X_2X_3}(\cdot|0,0,X_3))\right]\\ -\mathbb{E}_{P_{X_3|T=t}}[\mathbb{D}(\Gamma_{Y|X_1X_3}(\cdot \mid x_1,X_3)||\Gamma_{Y|X_1X_2X_3}(\cdot|x_1,0,X_3))].\end{aligned} \tag{271}$$

Proof: Follows by simple rewriting. Details omitted. ■

The following lemma is a direct consequence of Lemma 4 and the nonnegativity of Kullback-Leibler divergence.

Lemma 5. *Let (T, X_1, X_2, X_3, Y) be as in Lemma 4 where in addition we assume that for any $t \in \mathcal{T}$:*

$$\lim_{n\to\infty} P_{X_1|T=t}(1) = \lim_{n\to\infty} P_{X_2|T=t}(1) = 0. \tag{272}$$

Then:

$$\mathbb{I}(X_1;Y \mid X_2, X_3, T = t) \le P_{X_1|T=t}(1)\left(\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}_Y^{(1)}(X_3)\right] + o(1)\right) \tag{273}$$

$$\mathbb{I}(X_2;Y \mid X_1, X_3, T = t) \le P_{X_2|T=t}(1)\left(\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}_Y^{(2)}(X_3)\right] + o(1)\right) \tag{274}$$

Proof: We can write the mutual information term as follows:

$$\begin{aligned}&\mathbb{I}(X_1;Y \mid X_2, X_3, T = t)\\ &= P_{X_2|T=t}(0)\mathbb{I}(X_1;Y \mid X_2 = 0, X_3, T = t) + P_{X_2|T=t}(1)\mathbb{I}(X_1;Y \mid X_2 = 1, X_3, T = t) &(275)\\ &\le \mathbb{I}(X_1;Y \mid X_2 = 0, X_3, T = t) + P_{X_2|T=t}(1)\mathbb{I}(X_1;Y \mid X_2 = 1, X_3, T = t) &(276)\\ &\overset{(a)}{\le} P_{X_1|T=t}(1)\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}_Y^{(1)}(X_3)\right] + P_{X_2|T=t}(1)\cdot P_{X_1|T=t}(1)\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}(\Gamma_{Y|X_1X_2X_3}(\cdot \mid 1,1,X_3)||\Gamma_{Y|X_1X_2X_3}(\cdot|0,0,X_3))\right]\\ &= P_{X_1|T=t}(1)\left(\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}_Y^{(1)}(X_3)\right] + o(1)\right), &(277)\end{aligned}$$

where (a) holds by Lemma 4.

Similarly, one can show that

$$\mathbb{I}(X_2; Y \mid X_1, X_3, T=t) \leq P_{X_2|T=t}(1)\left(\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}_Y^{(2)}(X_3)\right] + o(1)\right). \tag{278}$$

■

B. Lower bound on $\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\delta_{n,w_3}$*:*

Recalling also the definition of $\widehat{Q}^n_{\mathcal{C},w_3}(z^n)$ in (22), we obtain for a specific code $\mathcal{C}$:

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

$$= \frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\sum_{z^n}\widehat{Q}^n_{\mathcal{C},w_3}(z^n)\log\left(\frac{\widehat{Q}^n_{\mathcal{C},w_3}(z^n)}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|0^n,0^n,x_3^n(w_3))}\right) \tag{279}$$

$$\stackrel{(a)}{\geq} \frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\sum_{i=1}^{n}\sum_{z_i}\widehat{Q}^{(i)}_{\mathcal{C},w_3}(z_i)\log\left(\frac{\widehat{Q}^{(i)}_{\mathcal{C},w_3}(z_i)}{\Gamma_{Z|X_1X_2X_3}(z_i|0,0,x_{3i}(w_3))}\right) \tag{280}$$

$$= \frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\sum_{i=1}^{n}\mathbb{D}\left(\widehat{Q}^{(i)}_{\mathcal{C},w_3}\|\Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_{3i}(w_3))\right) \tag{281}$$

$$\begin{aligned}\stackrel{(b)}{=} \frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\sum_{i=1}^{n}\mathbb{D}\Big(&\alpha_{n,i,1}\alpha_{n,i,2}\Gamma_{Z|X_1X_2X_3}(\cdot\mid 1,1,x_{3i}(w_3)) + \alpha_{n,i,1}(1-\alpha_{n,i,2})\Gamma_{Z|X_1X_2X_3}(\cdot\mid 1,0,x_{3i}(w_3))\\ &+(1-\alpha_{n,i,1})\alpha_{n,i,2}\Gamma_{Z|X_1X_2X_3}(\cdot\mid 0,1,x_{3i}(w_3)) + (1-\alpha_{n,i,1})(1-\alpha_{n,i,2})\Gamma_{Z|X_1X_2X_3}(\cdot\mid 0,0,x_{3i}(w_3))\\ &\|\Gamma_{Z|X_1X_2X_3}(\cdot|0,0,x_{3i}(w_3))\Big),\end{aligned} \tag{282}$$

where (a) holds by the memoryless nature of the channel and by defining $\widehat{Q}^{(i)}_{\mathcal{C},w_3}(z_i)$ as the probability of the event $Z_i = z_i$ conditioned on $W_3 = w_3$, and by writing out the expectations over the independent random variables $X_{1,i}$ and $X_{2,i}$.

Notice that since for all $w_3 \in \mathcal{M}_3$ we have that $\lim_{n\to\infty}\delta_{n,w_3} = 0$, by (282) we can conclude that

$$\lim_{n\to\infty}\alpha_{n,i,\ell} = 0, \quad \forall i \in \{1,\ldots,n\},\ \ell \in \{1,2\}. \tag{283}$$

Combining (282), (283) and Lemma 3, we can conclude that

$$\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\delta_{n,w_3} \geq n\mathbb{E}_{P_{TX_3}}\left[(1+o(1))\frac{(\alpha_{n,T,1}+\alpha_{n,T,2})^2}{2}\chi^2_n(\rho_{n,T,1},\rho_{n,T,2},X_3)\right], \tag{284}$$

where we define the time random variable T to be uniform over $\{1,\ldots,n\}$ and independent of all other random variables.

C. Upper bound on $\log(\mathsf{M}_1)$*:*

Since the message W_1 is uniform over $\{1,\ldots,\mathsf{M}_1\}$ and independent of the local randomness C_1, C_2, we have

$$\begin{aligned}\log(\mathsf{M}_1) &= \mathbb{H}(W_1) &(285)\\ &= \mathbb{H}(W_1 \mid W_2, S_1, S_2, W_3) &(286)\\ &= \mathbb{I}(W_1; Y^n \mid W_2, S_1, S_2, W_3) + \mathbb{H}(W_1 \mid Y^n, W_2, S_1, S_2, C_1, C_2, W_3) &(287)\\ &\stackrel{(a)}{\leq} \mathbb{I}(W_1; Y^n \mid W_2, S_1, S_2, C_1, C_2, X_2^n, X_3^n) + \mathbb{H}_b(P_{e,1}) + P_{e,1}\log(\mathsf{M}_1) &(288)\\ &= \frac{1}{1-P_{e,1}}(\mathbb{I}(W_1; Y^n \mid W_2, S_1, S_2, C_1, C_2, X_2^n, X_3^n) + \mathbb{H}_b(P_{e,1})) &(289)\\ &\stackrel{(b)}{=} \frac{1}{1-P_{e,1}}\Bigg(\sum_{i=1}^{n}\mathbb{H}(Y_i \mid Y^{i-1}, W_2, S_1, S_2, C_1, C_2, X_2^n, X_3^n)\\ &\qquad - \mathbb{H}(Y_i \mid Y^{i-1}, W_1, W_2, S_1, S_2, C_1, C_2, X_1^n, X_2^n, X_3^n) + \mathbb{H}_b(P_{e,1})\Bigg) &(290)\\ &\stackrel{(c)}{\leq} \frac{1}{1-P_{e,1}}\left(\sum_{i=1}^{n}\mathbb{H}(Y_i \mid X_{2i}, X_{3i}) - \mathbb{H}(Y_i \mid X_{1i}, X_{2i}, X_{3i}) + \mathbb{H}_b(P_{e,1})\right) &(291)\end{aligned}$$

$$= \frac{1}{1-P_{e,1}} \left(n \sum_{i=1}^{n} \frac{1}{n} \mathbb{I}(X_{1i}; Y_i \mid X_{2i}, X_{3i}) + \mathbb{H}_b(P_{e,1}) \right) \tag{292}$$

$$\leq \frac{1}{1-P_{e,1}} (n\mathbb{I}(X_{1,T}; Y_T \mid X_{2,T}, X_{3,T}) + 1) \tag{293}$$

$$\overset{(d)}{\leq} \frac{1}{1-P_{e,1}} n \left(\mathbb{E}_{P_{TX_{3,T}}} \left[\alpha_{n,T,1} \left(\mathbb{D}_Y^{(1)}(X_{3,T}) + o(1) \right) \right] + \frac{1}{n} \right). \tag{294}$$

Above sequence of (in)equalities are justified as follows:
(a) holds by Fano's inequality and because $X_2^n = \varphi_2^{(n)}(W_2, S_2, C_2)$ as well as $X_3^n = \varphi_3^{(n)}(W_3)$;
(b) holds by the chain rule of entropy and because $X_1^n = \varphi_1^{(n)}(W_1, S_1, C_1)$;
(c) holds respectively because conditioning reduces entropy and because conditioned on $X_{1,i}, X_{2,i}, X_{3,i}$ the output Y_i is independent of all messages, keys, and randomness;
(d) is obtained by applying Lemma 5 for each realization of T and by upper-bounding the binary entropy by 1.

D. *Upper Bound on* $\log(\mathsf{M}_2)$

The desired upper bound can be derived using similar steps to those for $\log(\mathsf{M}_1)$.

$$\log(\mathsf{M}_2) \leq \frac{1}{1-P_{e,1}} n \left(\mathbb{E}_{P_{TX_3}} \left[\alpha_{n,T,2} \left(\mathbb{D}_Y^{(2)}(X_3) + o(1) \right) \right] + \frac{1}{n} \right) \tag{295}$$

E. *Upper Bound on* $\log(\mathsf{M}_3)/n$*:*

Using standard steps, one can find the upper bound

$$\frac{1}{n} \log(\mathsf{M}_3) \leq \frac{1}{1-P_{e,1}} \mathbb{I}(X_{3,T}; Y_T | X_{1,T} = 0, X_{2,T} = 0, T). \tag{296}$$

F. *Upper bound on* $\log(\mathsf{M}_1)/\sqrt{n \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3}}$*:*

By (294) and (284), we obtain the bound

$$\frac{\log(\mathsf{M}_1)}{\sqrt{n \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3}}} \leq \frac{\sqrt{2}}{1-P_{e,1}} \frac{n \left(\mathbb{E}_{P_{TX_3}} \left[\alpha_{n,T,1} \left(\mathbb{D}_Y^{(1)}(X_3) + o(1) \right) \right] + \frac{1}{n} \right)}{\sqrt{n \left(n(1+o(1)) \mathbb{E}_{P_{TX_3}} \left[(\alpha_{n,T,1} + \alpha_{n,T,2})^2 \chi_n^2(\rho_{n,T,1}, \rho_{n,T,2}, X_3) \right] \right)}} \tag{297}$$

$$\overset{(a)}{=} \frac{\sqrt{2}}{1-P_{e,1}} \frac{\mathbb{E}_{P_{TX_3}} \left[\frac{\alpha_{n,T,1}}{\mathbb{E}_{P_T}[\alpha_{n,T,1}+\alpha_{n,T,2}]} \left(\mathbb{D}_Y^{(1)}(X_3) + o(1) \right) \right] + \frac{1}{n}}{\sqrt{(1+o(1)) \mathbb{E}_{P_{TX_3}} \left[\left(\frac{\alpha_{n,T,1}+\alpha_{n,T,2}}{\mathbb{E}_{P_T}[\alpha_{n,T,1}+\alpha_{n,T,2}]} \right)^2 \chi_n^2(\rho_{n,T,1}, \rho_{n,T,2}, X_3) \right]}} \tag{298}$$

$$\overset{(b)}{=} \frac{\sqrt{2}}{1-P_{e,1}} \frac{\mathbb{E}_{P_{TX_3}} \left[\rho_{n,T,1} \left(\mathbb{D}_Y^{(1)}(X_3) + o(1) \right) \right] + \frac{1}{n}}{\sqrt{(1+o(1)) \mathbb{E}_{P_{TX_3}} \left[(\rho_{n,T,1} + \rho_{n,T,2})^2 \chi_n^2(\rho_{n,T,1}, \rho_{n,T,2}, X_3) \right]}} \tag{299}$$

where (a) follows by normalization by $\mathbb{E}_{P_T}[\alpha_{n,T,1} + \alpha_{n,T,2}]$, and (b) by the definition of $\rho_{n,i,\ell}$ in (268) for all $(i, \ell) \in \{1, \ldots, n\} \times \{1, 2\}$.

Similarly, one can show that

$$\frac{\log(\mathsf{M}_2)}{\sqrt{n \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3}}} \leq \frac{\sqrt{2}}{1-P_{e,1}} \frac{\mathbb{E}_{P_{TX_3}} \left[\rho_{n,T,2} \left(\mathbb{D}_Y^{(2)}(X_3) + o(1) \right) \right] + \frac{1}{n}}{\sqrt{(1+o(1)) \mathbb{E}_{P_{TX_3}} \left[(\rho_{n,T,1} + \rho_{n,T,2})^2 \chi_n^2(\rho_{n,T,1}, \rho_{n,T,2}, X_3) \right]}} \tag{300}$$

G. *Upper bound on* $\frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3}$*:*

Notice that the new parameters $\rho_{n,T,\ell}$ are well defined because $\mathbb{E}_T[\alpha_{n,T,1} + \alpha_{n,T,2}]$ characterizes the sum of the fractions of 1-entries in the codebooks $\{x_1^n(W_1, S_1, C_1)\}$ and $\{x_2^n(W_2, S_2, C_2)\}$, and is thus non-zero because otherwise no communication is going on. Moreover, by Jensen's Inequality, $\mathbb{E}_T[(\rho_{n,T,1} + \rho_{n,T,2})^2] \geq (\mathbb{E}_T[\rho_{n,T,1} + \rho_{n,T,2}])^2 = 1$.

It then follows by Assumptions (27), that the right-hand sides of (296), (299) and (300) lie in bounded intervals. Combining the inequalities in (296), (299) and (300) with trivial positivity considerations, one can conclude that also the left-hand sides of these inequalities must lie in bounded intervals. Consequently there exists subsequence of blocklengths so that both the left- and ride-hand sides of (296), (299) and (300) all converge. We shall restrict to such a subsequence of blocklengths $\{n_i\}_{i=1}^{\infty}$.

Let then β_1 and β_2 be the two numbers in $[0,1]$ that satisfy

$$\lim_{i\to\infty} \frac{\log(\mathsf{M}_\ell)}{\sqrt{n_i \frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\delta_{n_i,w_3}}} = \sqrt{2}\beta_\ell \lim_{i\to\infty} \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{n_i,T,\ell}\mathbb{D}_Y^{(\ell)}(X_{3,T})\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{n_i,T,1}+\rho_{n_i,T,2})^2\chi_n^2(\rho_{n_i,T,1},\rho_{n_i,T,2},X_{3,T})\right]}}, \quad \ell\in\{1,2\} \tag{301}$$

where notice that the limit on the right-hand side coincides with the limits on the right-hand sides in (299) or (300), respectively.

Assume for the moment that β_1,β_2 are strictly larger than 0, and thus we can divide by it. Then, (301) combined with (294) and (295) implies that for all blocklengths n_i:

$$\sqrt{n_i\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\delta_{n_i,w_3}} \leq \frac{n_i}{\beta_\ell\sqrt{2}}\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{n_i,T,1}+\rho_{n_i,T,2})^2\chi_n^2(\rho_{n_i,T,1},\rho_{n_i,T,2},X_{3,T})\right]}+o(1), \qquad \ell\in\{1,2\}. \tag{302}$$

H. *Lower bound on* $\log(\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2)$*:*

We start with the lower bound

$$\begin{aligned}
\log(\mathsf{M}_1\mathsf{K}_1) &= \mathbb{H}(W_1,S_1|C_1,C_2) && (303)\\
&= \mathbb{H}(W_1,S_1 \mid X_3^n,C_1,C_2) && (304)\\
&\geq \mathbb{I}(W_1,S_1;Z^n \mid X_3^n,C_1,C_2) && (305)\\
&\overset{(a)}{\geq} \mathbb{I}(X_1^n;Z^n \mid X_3^n,C_1,C_2) && (306)\\
&= \mathbb{I}(X_1^n,X_2^n;Z^n \mid X_3^n,C_1,C_2) - \mathbb{I}(X_2^n;Z^n \mid X_1^n,X_3^n,C_1,C_2) && (307)\\
&\overset{(b)}{\geq} \mathbb{I}(X_1^n,X_2^n;Z^n \mid X_3^n,C_1,C_2) - \mathbb{I}(X_2^n;Z^n \mid X_1^n,X_3^n) && (308)
\end{aligned}$$

where (a) holds because $X_1^n = x_1^n(W_1,S_1,C_1)$ is a function of W_1, S_1, and C_1 and (b) holds because of the Markov chain $(C_1,C_2)\to(X_1^n,X_2^n,X_3^n)\to Z^n$ and because conditioning reduces entropy.

Likewise,

$$\log(\mathsf{M}_2\mathsf{K}_2) \geq \mathbb{I}(X_1^n,X_2^n;Z^n \mid X_3^n,C_1,C_2) - \mathbb{I}(X_1^n;Z^n \mid X_1^n,X_3^n), \tag{309}$$

We next focus on the first mutual-information term that is common to the RHS of (308) and (309). To this end, we define for each pair $(c_1,c_2)\in\mathcal{G}_1\times\mathcal{G}_2$ the warden's average output distribution conditioned on the local randomness c_1 and c_2 and on a non-covert message w_3, the distribution

$$\widehat{Q}^n_{\mathcal{C},(c_1,c_2,w_3)}(z^n) \triangleq \frac{1}{\mathsf{M}_1\mathsf{M}_2\mathsf{K}_1\mathsf{K}_2}\sum_{(w_1,s_1)}\sum_{(w_2,s_2)}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n|x_1^n(w_1,s_1,c_1),x_2^n(w_2,s_2,c_2),x_3^n(w_3)) \tag{310}$$

and the divergence

$$\delta_{n,(c_1,c_2,w_3)} \triangleq \mathbb{D}\left(\widehat{Q}^n_{\mathcal{C},(c_1,c_2,w_3)} \parallel \Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot|0^n,0^n,x_3^n(w_3))\right). \tag{311}$$

With this definition, we can write:

$$\begin{aligned}
&\mathbb{I}(X_1^n,X_2^n;Z^n \mid X_3^n,C_1,C_2) && (312)\\
&= \mathbb{E}_{X_1^n,X_2^n,X_3^n,C_1,C_2}\left[\mathbb{D}\left(\Gamma^{\otimes n}_{Z|X_1X_2X_3}(\cdot \mid X_1^n,X_2^n,X_3^n)\|\widehat{Q}^n_{\mathcal{C},(C_1,C_2,W_3)}\right)\right] && (313)\\
&= \mathbb{E}_{X_1^n,X_2^n,X_3^n,C_1,C_2}\Bigg[\sum_{z^n}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid X_1^n,X_2^n,X_3^n)\log\left(\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid X_1^n,X_2^n,X_3^n)}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid 0^n,0^n,X_3^n)}\right)\\
&\qquad -\sum_{z^n}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid X_1^n,X_2^n,X_3^n)\log\left(\frac{\widehat{Q}^n_{\mathcal{C},(C_1,C_2,W_3)}(z^n)}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid 0^n,0^n,X_3^n)}\right)\Bigg] && (314)\\
&\overset{(a)}{=} \mathbb{E}_{X_1^n,X_2^n,X_3^n}\left[\sum_{z^n}\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid X_1^n,X_2^n,X_3^n)\log\left(\frac{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid X_1^n,X_2^n,X_3^n)}{\Gamma^{\otimes n}_{Z|X_1X_2X_3}(z^n \mid 0^n,0^n,X_3^n)}\right)\right]\\
&\quad -\mathbb{E}_{C_1,C_2,W_3}\left[\delta_{n,(C_1,C_2,W_3)}\right] && (315)
\end{aligned}$$

$$\stackrel{(b)}{\geq} \sum_{i=1}^{n} \mathbb{E}_{X_{1,i},X_{2,i},X_{3,i}} \left[\sum_{z_i} \Gamma_{Z|X_1X_2X_3}(z_i \mid X_{1,i}, X_{2,i}, X_{3,i}) \log \left(\frac{\Gamma_{Z|X_1X_2X_3}(z_i \mid X_{1i}, X_{2i}, X_{3i})}{\Gamma_{Z|X_1X_2X_3}(z_i \mid 0, 0, X_{3i})} \right) \right] - \mathbb{E}_{W_3}\left[\delta_{n,W_3}\right] \tag{316}$$

$$\stackrel{(c)}{=} \sum_{i=1}^{n} \mathbb{E}_{X_{3i}} \left[\left(\alpha_{n,i,1} \mathbb{D}_Z^{(1)}(X_{3i}) + \alpha_{n,i,2} \mathbb{D}_Z^{(2)}(X_{3,i}) \right) (1 + o(1)) \right] - \mathbb{E}_{W_3}\left[\delta_{n,W_3}\right] \tag{317}$$

$$\stackrel{(d)}{=} n \mathbb{E}_{P_{TX_{3T}}} \left[\left(\alpha_{n,T,1} \mathbb{D}_Z^{(1)}(X_{3,T}) + \alpha_{n,T,2} \mathbb{D}_Z^{(2)}(X_{3,T}) \right) (1 + o(1)) \right] - \mathbb{E}_{W_3}\left[\delta_{n,W_3}\right], \tag{318}$$

where (a) holds by the definition of $\delta_{n,(c_1,c_2,w_3)}$ and by replacing the average over X_3^n by the average over W_3; (b) holds by by convexity of the divergence; (c) by writing out the expectations over the independent random variables $X_{1,i}$ and $X_{2,i}$ and by noting that for $X_{1,i} = X_{2,i} = 0$ the term in the expectation evaluates to 0; (d) holds because T is uniform over $\{1, \dots, n\}$.

For the second mutual-information term on the RHS of (308), we have:

$$\begin{aligned}
\mathbb{I}(X_1^n; Z^n \mid X_2^n, X_3^n) &= \mathbb{H}(Z^n \mid X_2^n, X_3^n) - \mathbb{H}(Z^n \mid X_1^n, X_2^n, X_3^n) &(319)\\
&= \sum_{i=1}^{n} \mathbb{H}(Z_i \mid Z^{i-1}, X_2^n, X_3^n) - \mathbb{H}(Z_i \mid Z^{i-1}, X_1^n, X_2^n, X_3^n) &(320)\\
&\stackrel{(a)}{=} \sum_{i=1}^{n} \mathbb{H}(Z_i \mid X_2^n, X_3^n) - \mathbb{H}(Z_i \mid X_{1,i}, X_{2,i}, X_{3,i}) &(321)\\
&\stackrel{(b)}{\leq} \sum_{i=1}^{n} \mathbb{H}(Z_i \mid X_{2i}, X_{3i}) - \mathbb{H}(Z_i \mid X_{1i}, X_{2i}, X_{3i}) &(322)\\
&= n \sum_{i=1}^{n} \frac{1}{n} \mathbb{I}(X_{1i}; Z_i \mid X_{2i}, X_{3i}) &(323)\\
&= n \mathbb{I}(X_{1,T}; Z_T \mid X_{2,T}, X_{3,T}, T) &(324)\\
&\stackrel{(c)}{\leq} n \mathbb{E}_{P_{TX_{3,T}}} \left[\alpha_{n,T,1} \mathbb{D}_Z^{(1)}(X_{3,T})(1 + o(1)) \right], &(325)
\end{aligned}$$

where (a) holds by the memoryless nature of the channel; (b) because conditioning reduces entropy; and (c) by applying Lemma 5 to output Z_T instead of Y_T.

In an analogous way, one can show that

$$\mathbb{I}(X_2^n; Z^n \mid X_1^n, X_3^n) \leq n \mathbb{E}_{P_{TX_{3,T}}} \left[\alpha_{n,T,2} \mathbb{D}_Z^{(2)}(X_{3,T})(1 + o(1)) \right]. \tag{326}$$

Combining (318), (325) and (326) with (308) we can conclude that

$$\log(\mathsf{M}_\ell \mathsf{K}_\ell) \geq n \left(\mathbb{E}_{P_{TX_3}} \left[\alpha_{n,T,\ell} \mathbb{D}_Z^{(\ell)}(X_3)(1 + o(1)) \right] - \frac{\mathbb{E}_{W_3}[\delta_{n,W_3}]}{n} \right), \quad \ell \in \{1, 2\}. \tag{327}$$

I. Lower bound on $\frac{\log(\mathsf{M}_2\mathsf{K}_2)}{\sqrt{n\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\delta_{n,w_3}}}$, $\frac{\log(\mathsf{M}_2\mathsf{K}_2)}{\sqrt{n\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\delta_{n,w_3}}}$ *and Discussion*

By combining (302) with (327), for $n \in \{n_i\}$ we obtain the bound for $\ell \in \{1, 2\}$:

$$\frac{\log(\mathsf{M}_\ell \mathsf{K}_\ell)}{\sqrt{n \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3}}} \geq \beta_\ell \left(\frac{\mathbb{E}_{P_{TX_3}} \left[\rho_{n,T,\ell} \mathbb{D}_Z^{(\ell)}(X_{3T}) \right]}{\sqrt{\mathbb{E}_{P_{TX_3}} \left[(\rho_{n,T,1} + \rho_{n,T,2})^2 \chi_n^2(\rho_{n,T,1}, \rho_{n,T,2}, X_3) \right]}} - \frac{\frac{\mathbb{E}_{W_3}[\delta_{n,W_3}]}{n\mathbb{E}_{P_T}[\alpha_{n,T,1} + \alpha_{n,T,2}]}}{\sqrt{\mathbb{E}_{P_{TX_3}} \left[(\rho_{n,T,1} + \rho_{n,T,2})^2 \chi_n^2(\rho_{n,T,1}, \rho_{n,T,2}, X_3) \right]}} \right). \tag{328}$$

Notice that the second term on the right-hand side of (328) vanishes whenever $\log(\mathsf{M}_1) \to \infty$ or $\log(\mathsf{M}_2) \to \infty$. In fact, we have

$$\begin{aligned}
\lim_{n \to \infty} \mathbb{E}_{W_3}[\delta_{n,W_3}] &= 0, &(329)\\
\liminf_{n \to \infty} n \cdot \mathbb{E}_T[\alpha_{n,T,1} + \alpha_{n,T,2}] &> 0, &(330)\\
\liminf_{n \to \infty} \mathbb{E}_{P_{TX_3}} \left[(\rho_{n,T,1} + \rho_{n,T,2})^2 \chi_n^2(\rho_{n,T,1}, \rho_{n,T,2}, X_3) \right] &> 0. &(331)
\end{aligned}$$

Limit (329) holds because $\delta_{n,w_3} \to 0$ for any $w_3 \in \mathcal{M}_3$;
Limit (330) holds because when the left-hand side tends to 0 then the number of 1-symbols in all the codewords tends to zero, in which case no messages can be transmitted reliably;
Limit (331) holds because by definition (31) and assumptions (27) the random variable in the expectation is 0 only when $\rho_{n,T,1} = \rho_{n,T,2} = 0$, which cannot happen with probability 1 over T as this would contradict (269).

J. Limits and Cardinality Bound

Notice that so far we have assumed that $\beta_1, \beta_2 > 0$. For $\beta_\ell \geq 0$, we simply note that:

$$\liminf_{n\to\infty} \frac{\log(\mathsf{M}_\ell \mathsf{K}_\ell)}{\sqrt{n \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n,w_3}}} \geq 0. \tag{332}$$

In view of this, and collecting all results in the previous subsections, we can conclude that there exists an increasing subsequences of blocklengths $\{n_i\}$ so that for any $\ell \in \{1,2\}$ and some $\beta_\ell \in [0,1]$:

$$\lim_{n_i\to\infty} \frac{\log(\mathsf{M}_\ell)}{\sqrt{n_i \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n_i,w_3}}} = \lim_{n_i\to\infty} \sqrt{2}\beta_\ell \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{n_i,T,\ell}\mathbb{D}_Y^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{n_i,1,T}+\rho_{n_i,2,T})^2 \chi^2_{n_i}(\rho_{n_i,T,1},\rho_{n_i,T,2},X_3)\right]}}, \tag{333a}$$

$$\limsup_{n_i\to\infty} \frac{1}{n_i}\log \mathsf{M}_3 \leq \lim_{n_i\to\infty} \mathbb{I}(X_{3,T};Y_T|X_{1,T}=0,X_{2,T}=0,T), \tag{333b}$$

$$\liminf_{n_i\to\infty} \frac{\log(\mathsf{M}_\ell \mathsf{K}_\ell)}{\sqrt{n_i \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n_i,w_3}}} \geq \lim_{n_i\to\infty} \sqrt{2}\beta_\ell \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{n_i,T,\ell}\mathbb{D}_Z^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{n_i,1,T}+\rho_{n_i,2,T})^2 \chi^2_{n_i}(\rho_{n_i,T,1},\rho_{n_i,T,2},X_3)\right]}}. \tag{333c}$$

Applying the Fenchel-Eggleston-Carathéodory theorem to vectors of the form

$$\mathbf{v} = \begin{pmatrix} \rho_{n_i,t,1}\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}_Y^{(1)}(X_3)\right] \\ \rho_{n_i,t,2}\mathbb{E}_{P_{X_3|T=t}}\left[\mathbb{D}_Y^{(2)}(X_3)\right] \\ \mathbb{I}(X_3;Y \mid X_1=0,X_2=0,T=t) \\ (\rho_{n_i,t,1}+\rho_{n_i,t,2})^2\,\mathbb{E}_{P_{X_3|T=t}}\left[\chi^2_{n_i}(\rho_{n_i,T,1},\rho_{n_i,T,2},X_3)\right] \\ \mathbb{E}_{P_{X_3|T=t}}\left[\rho_{n_i,t,1}\mathbb{D}_Z^{(1)}(X_3)\right] \\ \mathbb{E}_{P_{X_3|T=t}}\left[\rho_{n_i,t,2}\mathbb{D}_Z^{(2)}(X_3)\right] \end{pmatrix}, \tag{334}$$

we conclude that for any blocklength n_i there exists a modified distribution $\tilde{P}_T$ over an alphabet of size $|\mathcal{T}|= 6$ so that the bounds (333) hold also if P_T is replaced by this new pmf $\tilde{P}_T$. In the rest of the proof, we can thus restrict to these modified distributions $\tilde{P}_T$ over $|\mathcal{T}|= 6$.

K. The Limiting Distribution

To conclude the proof, we notice that by the Bolzano-Weierstrass theorem there exists an increasing subsequence $\{n_{i_k}\}$ of $\{n_i\}$ so that $\{P_{X_3|T}(\cdot|t)\}$ and $\{P_T(\cdot)\}$ converge on this subsequence. If also $\rho_{n_{i_k},t,1}$ and $\rho_{n_{i_k},t,2}$ converge for each value of $t \in \mathcal{T} \triangleq \{1,\ldots,6\}$, then by the continuity of the expressions, we obtain a converse result by considering the convergence points of $\{P_{X_3|T}(\cdot|t)\}$, $\{P_T(\cdot)\}$, and $\{(\rho_{1,t},\rho_{2,t})\}_{t\in\mathcal{T}}$. In fact, we can conclude that

$$\lim_{k\to\infty} \frac{\log(\mathsf{M}_\ell)}{\sqrt{n_{i_k} \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n_{i_k},w_3}}} = \sqrt{2}\beta_\ell \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T}\mathbb{D}_Y^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T}+\rho_{2,T})^2 \chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]}}, \quad \ell \in \{1,2\} \tag{335}$$

$$\limsup_{k\to\infty} \frac{1}{n_{i_k}}\log \mathsf{M}_3 \leq \mathbb{I}(X_3;Y|X_1=0,X_2=0,T), \tag{336}$$

$$\liminf_{k\to\infty} \frac{\log(\mathsf{M}_\ell \mathsf{K}_\ell)}{\sqrt{n_{i_k} \frac{1}{\mathsf{M}_3} \sum_{w_3=1}^{\mathsf{M}_3} \delta_{n_{i_k},w_3}}} \geq \sqrt{2}\beta_\ell \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T}\mathbb{D}_Z^{(\ell)}(X_3)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T}+\rho_{2,T})^2 \chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]}}, \quad \ell \in \{1,2\}. \tag{337}$$

for some pmf P_{X_3T} over $\mathcal{X}_3 \times \mathcal{T}$ and some set of positive pairs $\{(\rho_{1,t},\rho_{2,t})\}_{t\in\mathcal{T}}$.

If instead for some $t \in \mathcal{T}$ the sequence $\rho_{n_{i_k},t,\ell}$ diverges to ∞ we proceed as follows. We first notice that for each of these t-values the probability $P_T(t) \to 0$ as $n \to \infty$, because otherwise the expectation (269) is violated, and one of the following three cases applies:

1.) $P_T(t)\rho_{n_{i_k},t,\ell} \to 0$ and $P_T(t)\rho^2_{n_{i_k},t,\ell} \to 0$;
2.) $P_T(t)\rho_{n_{i_k},t,\ell} \to 0$ and $\lim_{n_{i_k}\to\infty} P_T(t)\rho^2_{n_{i_k},t,\ell} = c$ for some $c \in (0,\infty)$;
3.) $P_T(t)\rho_{n_{i_k},t,\ell} \in [0,1]$ and $P_T(t)\rho^2_{n_{i_k},t,\ell} \to \infty$.

All t-values satisfying case 1.) can simply be ignored since they do not change the bounds. Whenever there exists a t-value in case 3.), then bounds (333a) and (333c) are 0 for both $\ell \in \{1,2\}$ and the result is trivial. In case 2.) we can modify the probabilities $P_T(t)$ and the parameters $\rho_{n_{i_k},t,\ell}$ to values in a bounded interval $[a,b]$ for $b > a > 0$, while still approximating the bounds (333) arbitrarily closely. We then fall back to the case where all sequences $\rho_{n_{i_k},t,\ell}$ converge, which we discussed above.

L. Recovering the Bound on the Secret-Key Size

Above converse result remains valid if we add the additional constraint that results when taking the difference between (337) and (335):

$$\liminf_{k\to\infty} \frac{\log(\mathsf{K}_\ell)}{\sqrt{n_{i_k}\frac{1}{\mathsf{M}_3}\sum_{w_3=1}^{\mathsf{M}_3}\delta_{n_{i_k},w_3}}} \geq \sqrt{2}\beta_\ell \frac{\mathbb{E}_{P_{TX_3}}\left[\rho_{\ell,T}\left(\mathbb{D}_Z^{(\ell)}(X_3) - \mathbb{D}_Y^{(\ell)}(X_3)\right)\right]}{\sqrt{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T}+\rho_{2,T})^2\,\chi^2(\rho_{1,T},\rho_{2,T},X_3)\right]}}, \quad \ell \in \{1,2\}. \tag{338}$$

Relaxing the sum-constraint (337) completely and further relaxing the equality in (335) into an $\leq$-inequality establishes finally the desired converse proof to Theorem 2.

APPENDIX E
PROOF OF LEMMA 2

Fix two pmfs P_{TX_3} and Q_{TX_3} as well as the tuples $(\rho_{1,t},\rho_{2,t})$ and $(\rho'_{1,t},\rho'_{2,t})$ in $\mathbb{R}_0^{+^2}$ for all $t \in \mathcal{T}$. Define the tuples

$$\boldsymbol{\rho}_t \triangleq (\rho_{1,t},\rho_{2,t}), \tag{339}$$

$$\boldsymbol{\rho}'_t \triangleq (\rho'_{1,t},\rho'_{2,t}), \tag{340}$$

$$\boldsymbol{\mu} \triangleq (\boldsymbol{\rho}_1,\ldots,\boldsymbol{\rho}_6,\boldsymbol{\rho}'_1,\ldots,\boldsymbol{\rho}'_6), \tag{341}$$

for all $t \in \mathcal{T}$.

Let (r_1,r_2,r_3,k), (r'_1,r'_2,r'_3,k'), and $(\tilde{r}_1,\tilde{r}_2,\tilde{r}_3,\tilde{k})$ be the tuples of messages and key rates given by the right-hand sides of (59)–(61) when evaluated for P_{TX_3} and $(\boldsymbol{\rho}_1,\ldots,\boldsymbol{\rho}_6)$, for Q_{TX_3} and $(\boldsymbol{\rho}'_1,\ldots,\boldsymbol{\rho}'_6)$, and for R_{TX_3} and μ. We shall show that

$$\lambda\begin{pmatrix} r_1\\ r_2\\ r_3\\ k\end{pmatrix} + (1-\lambda)\begin{pmatrix} r'_1\\ r'_2\\ r'_3\\ k'\end{pmatrix} = \begin{pmatrix} \tilde{r}_1\\ \tilde{r}_2\\ \tilde{r}_3\\ \tilde{k}\end{pmatrix}, \quad \forall\lambda \in [0,1]. \tag{342}$$

The desired equality for the $\tilde{r}_3$-component is directly obtained by the linearity of conditional mutual information and because it does not depend on the $\boldsymbol{\rho}$-, $\boldsymbol{\rho}'$-, and $\boldsymbol{\mu}$-tuples. To see the equality for the other three components, fix $\lambda \in [0,1]$, and set $\nu > 0$ so that

$$\nu^2 \triangleq \frac{\mathbb{E}_{P_{TX_3}}\left[(\rho_{1,T}+\rho_{2,T})^2\,\chi^2(\boldsymbol{\rho}_T,X_3)\right]}{\mathbb{E}_{Q_{TX_3}}\left[\left(\rho'_{1,T}+\rho'_{2,T}\right)^2\,\chi^2(\boldsymbol{\rho}'_T,X_3)\right]}. \tag{343}$$

For all $\ell \in \{1,2\}$, upon forming the new pmf R_{TX_3} by choosing

$$R_T(t) = \begin{cases} \lambda \cdot P_T(t) & t \in \{1,\ldots,6\}\\ (1-\lambda)\cdot Q_T(t-6) & t \in \{7,\ldots,12\}\end{cases} \tag{344}$$

and

$$R_{X_3|T}(x_3|t) = \begin{cases} P_{X_3|T}(x_3|t) & t \in \{1,\ldots,6\}\\ Q_{X_3|T}(x_3|t-6) & t \in \{7,\ldots,12\},\end{cases} \tag{345}$$

and by defining the following tuples and constants

$$\tilde{\boldsymbol{\rho}}_t \triangleq \begin{cases} \boldsymbol{\rho}_t, & t \in \{1,\ldots,6\}\\ \boldsymbol{\rho}'_{t-6}, & t \in \{7,\ldots,12\}\end{cases} \tag{346}$$

$$\alpha_{\ell,t} \triangleq \begin{cases} \rho_{\ell,t}, & t \in \{1,\ldots,6\} \\ \nu \cdot \rho'_{\ell,t-6}, & t \in \{7,\ldots,12\} \end{cases} \tag{347}$$

$$\beta_t \triangleq \begin{cases} \rho_{1,t} + \rho_{2,t}, & t \in \{1,\ldots,6\} \\ \nu \cdot \left(\rho'_{1,t-6} + \rho'_{2,t-6}\right), & t \in \{7,\ldots,12\} \end{cases} \tag{348}$$

One can notice that for any function $f\colon \mathcal{X}_3 \to \mathbb{R}$ we have

$$\begin{aligned}
&\lambda \frac{\mathbb{E}_{P_{X_3 T}}\left[\rho_{\ell,T} f(X_3)\right]}{\sqrt{\mathbb{E}_{P_{X_3 T}}\left[\left(\rho_{1,T}+\rho_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}_T, X_3)\right]}} + (1-\lambda)\frac{\mathbb{E}_{Q_{X_3 T}}\left[\rho'_{\ell,T} f(X_3)\right]}{\sqrt{\mathbb{E}_{Q_{X_3 T}}\left[\left(\rho'_{1,T}+\rho'_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}'_T, X_3)\right]}} \\
&= \lambda \frac{\mathbb{E}_{P_{X_3 T}}\left[\rho_{\ell,T} f(X_3)\right]}{\sqrt{\mathbb{E}_{P_{X_3 T}}\left[\left(\rho_{1,T}+\rho_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}_T, X_3)\right]}} + (1-\lambda)\frac{\mathbb{E}_{Q_{X_3 T}}\left[\nu\rho'_{\ell,T} f(X_3)\right]}{\sqrt{\mathbb{E}_{Q_{X_3 T}}\left[\nu^2\left(\rho'_{1,T}+\rho'_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}'_T, X_3)\right]}} && (349) \\
&\overset{(a)}{=} \frac{\lambda\mathbb{E}_{P_{X_3 T}}\left[\rho_{\ell,T} f(X_3)\right] + (1-\lambda)\ \mathbb{E}_{Q_{X_3 T}}\left[\nu\rho'_{\ell,T} f(X_3)\right]}{\sqrt{\lambda\mathbb{E}_{P_{X_3 T}}\left[\left(\rho_{1,T}+\rho_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}_T, X_3)\right] + (1-\lambda)\mathbb{E}_{Q_{X_3 T}}\left[\nu^2\left(\rho'_{1,T}+\rho'_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}'_T, X_3)\right]}} && (350) \\
&= \frac{\mathbb{E}_{R_{X_3 T}}\left[\alpha_{\ell,T} f(X_3)\right]}{\sqrt{\mathbb{E}_{R_{X_3 T}}\left[\beta_T^2 \chi^2(\tilde{\boldsymbol{\rho}}_T, X_3)\right]}} && (351)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_{Q_{X_3 T}}\left[\nu^2\left(\rho'_{1,T}+\rho'_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}'_T, X_3)\right] \\
&\overset{(b)}{=} \mathbb{E}_{P_{X_3 T}}\left[\left(\rho_{1,T}+\rho_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}_T, X_3)\right] && (352) \\
&= \lambda\mathbb{E}_{P_{X_3 T}}\left[\left(\rho_{1,T}+\rho_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}_T, X_3)\right] + (1-\lambda)\mathbb{E}_{Q_{X_3 T}}\left[\nu^2\left(\rho'_{1,T}+\rho'_{2,T}\right)^2 \chi^2(\boldsymbol{\rho}'_T, X_3)\right], && (353)
\end{aligned}$$

where in (b) we used the definition of ν. This concludes the proof of the Lemma.

References

[1] Abdelaziz Bounhar, Mireille Sarkiss, and Michele Wigger. "Covert Multi-Access Communication with a Non-Covert User". In: *IEEE International Conference on Communications*. 2024, pp. 1473–1478. DOI: 10.1109/ICC51166.2024.10622567.

[2] Boulat A. Bash, Dennis Goeckel, and Don Towsley. "Limits of Reliable Communication with Low Probability of Detection on AWGN Channels". In: *IEEE Journal on Selected Areas in Communications* 31.9 (2013), pp. 1921–1930. DOI: 10.1109/JSAC.2013.130923.

[3] Boulat A. Bash et al. "Hiding information in noise: fundamental limits of covert wireless communication". In: *IEEE Communications Magazine* 53.12 (2015), pp. 26–31. DOI: 10.1109/MCOM.2015.7355562.

[4] Matthieu R. Bloch. "Covert Communication Over Noisy Channels: A Resolvability Perspective". In: *IEEE Transactions on Information Theory* 62.5 (2016), pp. 2334–2354. DOI: 10.1109/TIT.2016.2530089.

[5] Ligong Wang, Gregory W. Wornell, and Lizhong Zheng. "Fundamental Limits of Communication With Low Probability of Detection". In: *IEEE Transactions on Information Theory* 62.6 (2016), pp. 3493–3503. DOI: 10.1109/TIT.2016.2548471.

[6] Pak Hou Che et al. "Reliable deniable communication with channel uncertainty". In: *IEEE Information Theory Workshop*. 2014, pp. 30–34. DOI: 10.1109/ITW.2014.6970786.

[7] Si-Hyeon Lee et al. "Covert Communication With Channel-State Information at the Transmitter". In: *IEEE Transactions on Information Forensics and Security* 13.9 (2018), pp. 2310–2319. DOI: 10.1109/TIFS.2018.2818650.

[8] Hassan ZivariFard, Matthieu Bloch, and Aria Nosratinia. "Keyless Covert Communication in the Presence of Non-causal Channel State Information". In: *IEEE Information Theory Workshop*. 2019, pp. 1–5. DOI: 10.1109/ITW44776.2019.8989317.

[9] Hassan Zivarifard, Matthieu R. Bloch, and Aria Nosratinia. "Keyless Covert Communication via Channel State Information". In: *IEEE Transactions on Information Theory* 68.8 (2022), pp. 5440–5474. DOI: 10.1109/TIT.2021.3135291.

[10] Tamara V. Sobers et al. "Covert Communication in the Presence of an Uninformed Jammer". In: *IEEE Transactions on Wireless Communications* 16.9 (2017), pp. 6193–6206. DOI: 10.1109/TWC.2017.2720736.

[11] Ori Shmuel, Asaf Cohen, and Omer Gurewitz. "Multi-Antenna Jamming in Covert Communication". In: *IEEE Transactions on Communications* 69.7 (2021), pp. 4644–4658. DOI: 10.1109/TCOMM.2021.3067386.

[12] Hassan ZivariFard, Matthieu R. Bloch, and Aria Nosratinia. "Covert Communication via Non-Causal Cribbing from a Cooperative Jammer". In: *IEEE International Symposium on Information Theory*. 2021, pp. 202–207. DOI: 10.1109/ISIT45174.2021.9517883.

[13] Keerthi Suria Kumar Arumugam and Matthieu R. Bloch. "Embedding Covert Information in Broadcast Communications". In: *IEEE Transactions on Information Forensics and Security* 14.10 (2019), pp. 2787–2801. DOI: 10.1109/TIFS.2019.2907190.

[14] David Kibloff, Samir M. Perlaza, and Ligong Wang. "Embedding Covert Information on a Given Broadcast Code". In: *IEEE International Symposium on Information Theory*. 2019, pp. 2169–2173. DOI: 10.1109/ISIT.2019.8849435.

[15] Sang Wu Kim and Hien Q. Ta. "Covert Communications Over Multiple Overt Channels". In: *IEEE Transactions on Communications* 70.2 (2022), pp. 1112–1124. DOI: 10.1109/TCOMM.2021.3127532.

[16] Hien Quang Ta et al. "Covert Communications Over Non-Orthogonal Multiple Overt Channels". In: *IEEE Access* 10 (2022), pp. 122361–122375. DOI: 10.1109/ACCESS.2022.3223357.

[17] Keerthi Suria Kumar Arumugam and Matthieu R. Bloch. "Covert Communication Over a K -User Multiple-Access Channel". In: *IEEE Transactions on Information Theory* 65.11 (2019), pp. 7020–7044. DOI: 10.1109/TIT.2019.2930484.

[18] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory, 2nd Edition*. Wiley, 2006.

[19] Kang-Hee Cho and Si-Hyeon Lee. "Treating Interference as Noise Is Optimal for Covert Communication Over Interference Channels". In: *IEEE Transactions on Information Forensics and Security* 16 (2021), pp. 322–332. DOI: 10.1109/TIFS.2020.3013213.

[20] Erich L. Lehmann and Joseph P. Romano. *Testing Statistical Hypotheses*. New York, NY, USA: Springer-Verlag, 2005.

[21] Keerthi Suria Kumar Arumugam and Matthieu R. Bloch. "Covert Communication over a K-User Multiple Access Channel". In: *CoRR* abs/1803.06007 (2018). arXiv: 1803.06007. URL: http://arxiv.org/abs/1803.06007.

[22] Abdelaziz Bounhar, Mireille Sarkiss, and Michèle Wigger. "Mixing a Covert and a Non-Covert User". In: *IEEE International Symposium on Information Theory*. 2023, pp. 2577–2582. DOI: 10.1109/ISIT54713.2023.10206954.

[23] Jie Hou. "Coding for relay networks and effective secrecy for wire-tap channels". PhD dissertation. Technischen Universitat München, 2014.