Semantic-Security Capacity for Wiretap Channels of Type II

Ziv Goldfeld
Ben Gurion University
gziv@post.bgu.ac.il

Paul Cuff
Princeton University
cuff@princeton.edu

Haim H. Permuter
Ben Gurion University
haimp@bgu.ac.il

***Abstract*—The secrecy capacity of the type II wiretap channel (WTC II) with a noisy main channel is currently an open problem. Herein its secrecy-capacity is derived and shown to be equal to its semantic-security (SS) capacity. In this setting, the legitimate users communicate via a discrete-memoryless (DM) channel in the presence of an eavesdropper that has perfect access to a subset of its choosing of the transmitted symbols, constrained to a fixed fraction of the blocklength. The secrecy criterion is achieved simultaneously for all possible eavesdropper subset choices. On top of that, SS requires negligible mutual information between the message and the eavesdropper's observations even when maximized over all message distributions.**

A key tool for the achievability proof is a novel and stronger version of Wyner's soft covering lemma. Specifically, the lemma shows that a random codebook achieves the soft-covering phenomenon with high probability. The probability of failure is doubly-exponentially small in the blocklength. Since the combined number of messages and subsets grows only exponentially with the blocklength, SS for the WTC II is established by using the union bound and invoking the stronger soft-covering lemma. The direct proof shows that rates up to the weak-secrecy capacity of the classic WTC with a DM erasure channel (EC) to the eavesdropper are achievable. The converse follows by establishing the capacity of this DM wiretap EC as an upper bound for the WTC II.

I. Introduction

Information theoretic security has adopted the weak- and strong-secrecy metrics as a standard for measuring security. Respectively, weak- and strong-secrecy refer to the normalized and unnormalized mutual information between the secret message and the channel symbol string observed by the eavesdropper. However, recent work argues that, from a cryptographic point of view, both these metrics are insufficient to provide security of applications [1], [2]. Their main drawback lies in the assumption that the message is random and uniformly distributed, as real-life messages are neither (messages may be files, votes or any type of structured data, often with low entropy). Semantic-security (SS) [3], [4] is a cryptographic gold standard that was proposed in [2] as an adequate alternative and shown to be equivalent to a vanishing unnormalized mutual information for all message distributions. Adopting SS as our secrecy measure, we establish the SS-capacity of the wiretap channel of type II (WTC II) with a noisy main channel, for which even the weak-secrecy-capacity was an open problem until now. On top of that, the SS-capacity and the weak-secrecy-capacity are shown to coincide.

Secret communication over noisy channels dates back to Wyner who introduced the degraded wiretap channel (WTC) and derived its weak-secrecy-capacity [5]. Csiszár and Körner extended Wyner's result to the non-degraded WTC [6], which is henceforth referred to as the WTC I. A special instance of the WTC I is when the eavesdropper's observation is an outcome of a discrete-memoryless (DM) erasure channel (EC), which essentially means that he observes a subset of the transmitted symbols that is chosen at random by nature. The WTC II was proposed by Ozarow and Wyner [7] as a generalization of this instance, where a more powerful eavesdropper selects which subset to observe. In [7], the rate-equivocation region for the case where the main channel is *noiseless* was established by using a unique randomized coset coding scheme in the proof of achievability. The WTC II with a general (i.e., possibly *noisy*) DM main channel, however, remained an open problem ever since.

A recent endeavor at the optimal secrecy rate of the WTC II with a noisy main channel was presented in [8] (see also [9]–[12] for related work). Requiring a vanishing *average* error probability and security with respect to the *weak-secrecy* metric (namely, while assuming a uniformly distributed message and a normalized mutual information), the authors of [8] extended the coset coding scheme from [7] to obtain an inner bound on the rate-equivocation region. An outer bound was also established in [8] by assuming that the subset the eavesdropper chooses to observe is revealed to all parties (i.e., to the legitimate users). Specializing these bounds to the case where equivocation is maximal results in an inner and an outer bound on the weak-secrecy-capacity of a general WTC II; these bounds do not match.

In this work, we strengthen both the reliability and the security criteria, and derive the *SS-capacity* of the WTC II with a noisy main channel under a vanishing *maximal* error probability requirement. In the heart of the proof stands a stronger version of the soft-covering lemma. Wyners original soft-covering lemma [13, Theorem 6.3] is a valuable tool for achievability proofs of information theoretic security [14]–[16], resolvability [17], channel synthesis [18], and source coding [19] (see also references therein). The result herein sharpens the claim of soft-covering by moving away from an expected value analysis. Instead, we show that a random codebook achieves the soft-covering phenomenon with high probability. The probability of failure is doubly-exponentially small in the blocklength, enabling more powerful applications through the union bound.

Using the stronger soft-covering lemma, we move away

978-1-5090-1806-2/16/$31.00 ©2016 IEEE

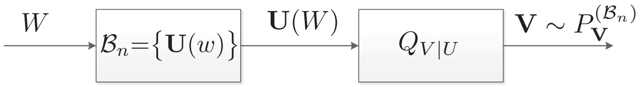
W
$\mathcal{B}_n=\{\mathbf{U}(w)\}$
$\mathbf{U}(W)$
$Q_{V|U}$
$\mathbf{V}\sim P_{\mathbf{V}}^{(\mathcal{B}_n)}$

Fig. 1: Coding problem with the goal of making $P_{\mathbf{V}}^{(\mathcal{B}_n)} \approx Q_V^n$.

from the coset coding schemes from [7], [8] and show that classic wiretap codes achieve the SS-capacity of the WTC II. SS requires a vanishing unnormalized mutual information (between the message and the eavesdropper's observation), when maximized over all message distributions and eavesdropper's subset choices. However, noting that their combined number grows only exponentially with the blocklength, the stronger soft-covering lemma is sharp enough to imply that the probability of an insecure random wiretap code is doubly-exponentially small. Using the expurgation method (e.g., cf. [20, Theorem 7.7.1]) we upgrade reliability to account for a maximal error probability criteria. The direct proof shows that any rate up to the weak-secrecy-capacity of the WTC I with a DM-EC[1] to the eavesdropper, is achievable. The converse follows by showing that the weak-secrecy-capacity of this WTC I upper bounds the SS-capacity of the WTC II. An important consequence of the WTC II SS-capacity proof is that Wyner's wiretap codes for the erasure WTC I, are optimal. The binary version of these codes is one of the few examples for which there are explicit constructions of practical secure encoders and decoders with optimal performance [21], [22].

II. The Stronger Soft-Covering Lemma

We use notation from [23, Section II]. Wyner's soft-covering lemma [13, Theorem 6.3] states that the distribution induced by selecting a u-sequence at random from an appropriately chosen set $\mathcal{B}_n$ and passing it through a memoryless channel $Q_{V|U}$, results in a good approximation of Q_V^n in the limit of large n, as long as the set is of size $|\mathcal{B}_n| = 2^{nR}$, where $R > I(U;V)$ (Fig. 1). In fact, the set $\mathcal{B}_n$ can be chosen quite carelessly - by random codebook construction, drawing each sequence independently from the distribution Q_U^n.

The soft-covering lemmas in the literature use a distance metric on distributions (commonly total variation or relative entropy) and claim that the distance between the induced distribution $P_{\mathbf{V}}^{(\mathcal{B}_n)}$ and the desired distribution Q_V^n vanishes exponentially fast in expectation over the random selection of the set[2]. Here we give a stronger claim. With high probability with respect to the set construction, the distance vanishes exponentially quickly with the blocklength n. The negligible probability of the random set not producing this desired result is doubly-exponentially small.

Let $\mathcal{W} = \big[1 : 2^{nR}\big]$ and $\mathbb{B}_n = \big\{\mathbf{U}(w)\big\}_{w\in\mathcal{W}}$ be a set of random vectors of length n that are independently and identically distributed (i.i.d.) according to Q_U^n. We refer to $\mathbb{B}_n$ as the random codebook. Let $\mathcal{B}_n = \big\{\mathbf{u}(w,\mathcal{B}_n)\big\}_{w\in\mathcal{W}}$ denote a realization of $\mathbb{B}_n$. For a fixed $\mathcal{B}_n$, the induced distribution is:

$$P^{(\mathcal{B}_n)}(\mathbf{v}) = 2^{-nR}\sum_{w\in\mathcal{W}} Q_{V|U}^n\big(\mathbf{v}\big|\mathbf{u}(w,\mathcal{B}_n)\big). \qquad (1)$$

Lemma 1 (Stronger Soft-Covering Lemma) *For any Q_U, $Q_{V|U}$, and $R > I(U;V)$, where $|\mathcal{V}| < \infty$, there exist $\gamma_1,\gamma_2 > 0$, such that for n large enough*

$$\mathbb{P}\bigg(D\Big(P_{\mathbf{V}}^{(\mathbb{B}_n)}\Big|\Big|Q_V^n\Big) > e^{-n\gamma_1}\bigg) \leq e^{-e^{n\gamma_2}}, \qquad (2)$$

where $D\big(P\big|\big|Q\big) = \sum_{x\in\mathrm{supp}(P)} P(x)\log\frac{P(x)}{Q(x)}$ is the relative entropy between two PMFs P and Q on a countable set $\mathcal{X}$.

The proof of Lemma 1 splits the analysis to separately consider typical and atypical input-output pairs. The double-exponential decay of the probability is derived via a Chernoff bound. The full details are omitted due to space limitations. The reader is referred to [23] for a full statement that includes exact exponents of decay and its proof (see also [24]).

III. Wiretap Channels of Type II

A. Problem Definition

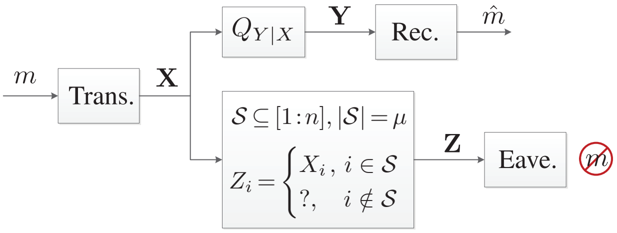
m
Trans.
$\mathbf{X}$
$Q_{Y|X}$
$\mathbf{Y}$
Rec.
$\hat{m}$
$\mathcal{S}\subseteq[1\!:\!n], |\mathcal{S}|=\mu$
$Z_i=\begin{cases}X_i, & i\in\mathcal{S}\\ ?, & i\notin\mathcal{S}\end{cases}$
$\mathbf{Z}$
Eave.
$\cancel{m}$

Fig. 2: The type II wiretap channel.

The WTC II is illustrated in Fig. 2. The sender chooses a message m from the set $\big[1 : 2^{nR}\big]$ and maps it into a sequence $\mathbf{x} \in \mathcal{X}^n$ (the mapping may be random). The sequence $\mathbf{x}$ is transmitted over a point-to-point DMC with transition probability $Q_{Y|X}$. Based on the received channel output sequence $\mathbf{y} \in \mathcal{Y}^n$, the receiver produces an estimate $\hat{m}$ of m. The eavesdropper noiselessly observes a subset of its choice of the n transmitted symbols. Namely, the eavesdropped chooses $\mathcal{S} \subseteq [1 : n]$, $|\mathcal{S}| = \mu \leq n$, and observes $\mathbf{z} \in \big(\mathcal{X}\cup\{?\}\big)^n$, where

$$z_i = \begin{cases} x_i\,, & i\in\mathcal{S} \\ ?, & i\notin\mathcal{S}\end{cases}\quad. \qquad (3)$$

Definition 1 (Code) *An (n,R) WTC II code $\mathcal{C}_n$ has: (i) A message set $\mathcal{M} = \big[1 : 2^{nR}\big]$; (ii) A stochastic encoder $f : \mathcal{M} \to \mathcal{P}(\mathcal{X}^n)$, where $\mathcal{P}(\mathcal{X}^n)$ denotes the set of all PMFs on $\mathcal{X}^n$; (iii) A decoding function $\phi : \mathcal{Y}^n \to \mathcal{M}$.*

Definition 2 (Maximal Error Probability) *The maximal error probability of an (n,R) WTC II code $\mathcal{C}_n$ is $P_e^\star(\mathcal{C}_n) = \max_{m\in\mathcal{M}} e_m(\mathcal{C}_n)$, where*

$$e_m(\mathcal{C}_n) = \sum_{\mathbf{x}\in\mathcal{X}^n} f(\mathbf{x}|m) \sum_{\substack{\mathbf{y}\in\mathcal{Y}^n:\\ \phi(\mathbf{y})\neq m}} Q_{Y|X}^n(\mathbf{y}|\mathbf{x}). \qquad (4)$$

[1] the erasure probability corresponds to the portion of symbols the eavesdropper in the WTC II does not intercept

[2] Many of the theorems only claim existence of a good codebook, but all of the proofs use expected value to establish existence.

Definition 3 (SS Metric) *The SS metric with respect to a (n,R) WTC II code $\mathcal{C}_n$ is* [3]

$$\mathrm{Sem}_\mu(\mathcal{C}_n) = \max_{\substack{P_M\in\mathcal{P}(\mathcal{M}),\\ \mathcal{S}\subseteq[1:n]:\ |\mathcal{S}|=\mu}} I_{\mathcal{C}_n}(M;\mathbf{Z}), \tag{5}$$

where $I_{\mathcal{C}_n}$ denotes a mutual information term that is calculated with respect to the joint PMF of M and $\mathbf{Z}$ induced by $\mathcal{C}_n$. Namely, for any $\mathcal{S}\subseteq[1:n]$ and $P_M\in\mathcal{P}(\mathcal{M})$, $P^{(\mathcal{C}_n,\mathcal{S})}_{M,\mathbf{Z}}$ is

$$P^{(\mathcal{C}_n,\mathcal{S})}_{M,\mathbf{Z}}(m,\mathbf{z}) = P(m)\sum_{\mathbf{x}\in\mathcal{X}^n} f(\mathbf{x}|m)\mathbb{1}_{\left\{z_i=x_i,\,i\in\mathcal{S}\right\}\cap\left\{z_i=?,\,i\notin\mathcal{S}\right\}}.$$

Remark 1 *SS requires that a single code works well for all message PMFs and all subset choices. Accordingly, the mutual information term in* (5) *is maximized over P_M and $\mathcal{S}$ when the code $\mathcal{C}_n$ is known. In other words, although not stated explicitly in our notation, $\mathcal{S}$ and P_M are functions of $\mathcal{C}_n$.*

Definition 4 (Semantically-Secure Codes) *Let $\alpha\in[0,1]$ and $\mu=\lfloor\alpha n\rfloor$, a sequence of (n,R) WTC II codes $\left\{\mathcal{C}_n\right\}_{n\geq1}$ is α-semantically-secure if there is a constants $\gamma>0$, such that $\mathrm{Sem}_\mu(\mathcal{C}_n)\leq e^{-n\gamma}$ for sufficiently large n.*

Remark 2 *By Definition 4, for a sequence of codes to be semantically-secure, the corresponding SS metrics must vanish exponentially fast. This is a standard requirement in the cryptography community, commonly referred to as strong-SS (see, e.g., [2, Section 3.2]).*

Definition 5 (SS-Achievability) *Let $\alpha\in[0,1]$ and $\mu=\lfloor\alpha n\rfloor$, a rate $R\in\mathbb{R}_+$ is α-SS-achievable if there is a sequence of (n,R) α-semantically-secure WTC II codes $\left\{\mathcal{C}_n\right\}_{n\geq1}$ with $P^\star_e(\mathcal{C}_n)\to0$ as $n\to\infty$.*

Definition 6 (SS-Capacity) *For any $\alpha\in[0,1]$, the α-SS-capacity of the WTC II $C_{\mathrm{Sem}}(\alpha)$ is the supremum of the set of α-SS-achievable rates.*

B. Capacity Results

The following proposition is subsequently used for the converse proof of the WTC II SS-capacity. The proposition states that the strong-secrecy-capacity of a WTC I with a DM-EC to the eavesdropper is an upper bound on the strong-secrecy-capacity of the WTC II. Strong-secrecy-capacity is defined with respect to the *average* error probability (instead of the maximal error probability from Definition 2) and the *strong-secrecy* metric (instead of the SS-metric from Definition 3). For example, the strong-secrecy metric of the WTC II is defined similarly to (5) while removing the maximization over $P_M\in\mathcal{P}(\mathcal{M})$ and letting M be a random variable uniformly distributed over $\mathcal{M}$. See [23] for the full definitions.

Proposition 2 (WTC I Upper Bounds WTC II) *Let $\alpha\in[0,1]$ and $C^{\mathrm{II}}_{\mathrm{S}}(\alpha)$ be the α-strong-secrecy-capacity of the WTC II with a main channel $Q^{(2)}_{Y|X}$. Furthermore, let $C^{\mathrm{I}}_{\mathrm{S}}(\alpha)$ be the strong-secrecy-capacity of the WTC I with transition probability $Q^{(1)}_{Y,Z|X}=Q^{(2)}_{Y|X}\mathcal{E}^{(\alpha)}_{Z|X}$, where $\mathcal{E}^{(\alpha)}_{Z|X}$ is a DM-EC with erasure probability $\bar{\alpha}=1-\alpha$, i.e.,*

$$\mathcal{E}^{(\alpha)}_{Z|X}(z|x) = \begin{cases}\alpha, & z=x\\ \bar{\alpha}, & z=?\end{cases}, \ \forall x\in\mathcal{X}. \tag{6}$$

Then

$$C^{\mathrm{II}}_S(\alpha)\leq C^{\mathrm{I}}_S(\alpha)=\max_{\substack{Q_{V,X}:\\ V-X-Y}}\Big[I(V;Y)-\alpha I(V;X)\Big]. \tag{7}$$

Due to space limitations, Section IV-A gives only an outline of the proof of Proposition 2 (see [23] for the full derivation). The proof leverages Sanov's Theorem and the continuity of mutual information.

Theorem 3 (WTC II SS-Capacity) *For any $\alpha\in[0,1]$,*

$$C_{\mathrm{Sem}}(\alpha)=\max_{\substack{Q_{V,X}:\\ V-X-Y}}\Big[I(V;Y)-\alpha I(V;X)\Big], \tag{8}$$

and one may restrict the cardinality of V to $|\mathcal{V}|<|\mathcal{X}|$.

The direct part of Theorem 3 is proven in Section IV-B. The stronger soft-covering lemma is key in the security analysis. The converse is a direct consequence of Proposition 2 by noting that the α-strong-secrecy-capacity of the WTC II upper bounds its α-SS-capacity.

Remark 3 *Theorem 3 recovers the achievability result from [8, Equation 7] by setting $V=X$ and taking X to be uniformly distributed over $\mathcal{X}$. Although we require security with respect to a stricter metric than used in [8] (SS versus weak-secrecy), we achieve higher rates than [8, Equation 7] and establish their optimality. Moreover, $C_{\mathrm{Sem}}(\alpha)$ is achieved via classic wiretap codes, making the (rather convoluted) coset coding scheme from [8] (inspired by [7]) no longer required.*

IV. Proofs

A. Proof Outline of Proposition 2

The equality in (7) follows by evaluating the strong-secrecy-capacity formula of a general WTC I, i.e.,

$$\max_{\substack{Q_{V,X}:\\ V-X-(Y,Z)}}\Big[I(V;Y)-I(V;Z)\Big], \tag{9}$$

under the transition probability matrix $Q^{(1)}_{Y,Z|X}$.

To prove the inequality in (7), we first show that for any $\alpha\in(0,1]$ and $\beta\in[0,\alpha)$, an α-strong-secrecy-achievable rate for the WTC II is also achievable for the WTC I with erasure probability $\bar{\beta}$. Having this, Proposition 2 follows by a continuity argument of the strong-secrecy-capacity of WTC I with respect to β.

For any α-strong-secrecy-achievable rate $R\in\mathbb{R}_+$ for the WTC II, let $\left\{\mathcal{C}^{\mathrm{II}}_n\right\}_{n\geq1}$ be the corresponding sequence of (n,R) codes. Since the channels to the legitimate receivers are the same for both versions of the WTC, $\left\{\mathcal{C}^{\mathrm{II}}_n\right\}_{n\geq1}$ also achieves reliability for the WTC I.

[3] $\mathrm{Sem}_\mu(\mathcal{C}_n)$ is actually the mutual-information-security (MIS) metric, which is equivalent to SS by [2]. We use the representation in (5) rather than the formal definition of SS (see, e.g., [2, Equation (4)]) out of analytical convenience.

Denote by $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ the eavesdroppers observation in the WTC I and the WTC II, respectively, that is induced by $\big\{\mathcal{C}_n^{\text{II}}\big\}_{n\geq 1}$. To establish (7), it suffices to show that $I_{\mathcal{C}_n^{\text{II}}}\big(M;\mathbf{Z}^{(1)}\big)$ can be made arbitrarily small with the blocklength n. To this end, for every $\mathbf{z}\in\mathcal{Z}^n$, where $\mathcal{Z}\triangleq\mathcal{X}\cup\{?\}$, define

$$\mathcal{A}(\mathbf{z})\triangleq\big\{i\in[1:n]\big|z_i=?\big\} \tag{10}$$

$$\Theta(\mathbf{Z})\triangleq\mathbb{1}_{\big\{|\mathcal{A}(\mathbf{Z})|\leq\lceil\bar{\alpha}n\rceil\big\}}. \tag{11}$$

Namely, Θ indicates if the number of erasures in a sequence $\mathbf{z}\in\mathcal{Z}^n$ is greater than or equal to $\lceil\bar{\alpha}n\rceil$ or not.

By conditioning the mutual information term $I_{\mathcal{C}_n^{\text{II}}}\big(M;\mathbf{Z}^{(1)}\big)$ on $\Theta(\mathbf{Z}^{(1)})$, we distinguish between the two cases of $\mathbf{Z}^{(1)}$ being better or worse than $\mathbf{Z}^{(2)}$. When $\Theta(\mathbf{Z}^{(1)})=0$, i.e., $\mathbf{Z}^{(1)}$ is worse that $\mathbf{Z}^{(2)}$, security for the WTC I is ensured since $\big\{\mathcal{C}_n^{\text{II}}\big\}_{n\geq 1}$ achieve security for the WTC II. Otherwise, for the case that $\Theta(\mathbf{Z}^{(1)})=1$ and $\mathbf{Z}^{(1)}$ is better than $\mathbf{Z}^{(2)}$, we use Sanov's theorem to show that the probability of such an event exponentially decreases with the blocklength n, while the mutual information grows linearly at most.

B. Direct Proof of Theorem 3

We show the achievability of (8) when $V=X$. Then, a standard channel prefixing argument extends the proof to any V with $V-X-Y$. Furthermore, we present here a construction of a sequence of semantically-secure codes with a vanishing *average* error probability. The expurgation technique [20, Theorem 7.7.1] allows upgrading reliability to achieve a vanishing *maximal* error probability, while preserving SS.

Fix $\alpha\in[0,1]$, $\epsilon>0$ and a PMF $Q_X\in\mathcal{P}(\mathcal{X})$, and let M and W be independent random variables uniformly distributed over $\mathcal{M}$ and $\mathcal{W}=\big[1:2^{n\tilde{R}}\big]$, respectively.

Codebook Construction: Let $\mathbb{B}_n$ be a *random codebook* given by a collection of i.i.d. random vectors $\mathbb{B}_n=\big\{\mathbf{X}(m,w)\big\}_{(m,w)\in\mathcal{M}\times\mathcal{W}}$, each distributed according to Q_X^n. A realization of $\mathbb{B}_n$ is denoted by $\mathcal{B}_n=\big\{\mathbf{x}(m,w,\mathcal{B}_n)\big\}_{(m,w)\in\mathcal{M}\times\mathcal{W}}$, with respect to which we construct a classic wiretap code.

Encoder f: To send $m\in\mathcal{M}$ the encoder randomly and uniformly chooses w from $\mathcal{W}$ and transmits $\mathbf{x}(m,w,\mathcal{B}_n)$.

Decoder ϕ: Upon observing $\mathbf{y}$, the receiver searches for a unique pair $(\hat{m},\hat{w})\in\mathcal{M}\times\mathcal{W}$ such that $\big(\mathbf{x}(\hat{m},\hat{w},\mathcal{B}_n),\mathbf{y}\big)\in\mathcal{T}_\epsilon^n(Q_{X,Y})$. If such a unique pair is found, then $\hat{m}$ is declared as the decoded message; otherwise, an error is declared.

The the triple $(\mathcal{M},f,\phi)$ defined with respect to the codebook $\mathcal{B}_n$ constitute an (n,R) code for the WTC II. When a random codebook $\mathbb{B}_n$ is used, we denote the corresponding random code by $\mathbb{C}_n$. Standard joint-typicality decoding arguments show that

$$\mathbb{E}_{\mathbb{C}_n}\frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}e_m(\mathbb{C}_n)\xrightarrow[n\to\infty]{}0, \tag{12}$$

(see (4)) provided that

$$R+\tilde{R}<I(X;Y). \tag{13}$$

Security Analysis: Fix $\mathcal{S}\subseteq[1:n]$ with $|\mathcal{S}|=\mu=\lfloor\alpha n\rfloor$ and define the following PMF on $\mathcal{Z}^n$,

$$\Gamma_{\mathbf{Z}}^{(\mathcal{S})}(\mathbf{z})=\prod_{j\in\mathcal{S}^c}\mathbb{1}_{\big\{z_j=?\big\}}\prod_{j\in\mathcal{S}}\mathcal{I}_Z(z_j),\quad\forall\mathbf{z}\in\mathcal{Z}^n, \tag{14}$$

where $\mathcal{I}_Z$ is the marginal output PMF of the identity DMC with input distribution Q_X:

$$\mathcal{I}_Z(z)=\sum_{x\in\mathcal{X}}Q_X(x)\mathbb{1}_{\{z=x\}}=Q_X(z)\mathbb{1}_{\{z\in\mathcal{X}\}}. \tag{15}$$

For any $\mathcal{C}_n$ (defined by fixing $\mathcal{B}_n$) and $P_M\in\mathcal{P}(\mathcal{M})$, the relative entropy chain rule implies

$$\begin{aligned}I_{\mathcal{C}_n}(M;\mathbf{Z})&=D\left(P_{\mathbf{Z}|M}^{(\mathcal{C}_n,\mathcal{S})}\middle|\middle|P_{\mathbf{Z}}^{(\mathcal{C}_n,\mathcal{S})}\middle|P_M\right)\\&=D\left(P_{\mathbf{Z}|M}^{(\mathcal{C}_n,\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}}^{(\mathcal{S})}\middle|P_M\right)-D\left(P_{\mathbf{Z}}^{(\mathcal{C}_n,\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}}^{(\mathcal{S})}\middle|P_M\right),\end{aligned} \tag{16}$$

and therefore

$$\begin{aligned}&\max_{P_M\in\mathcal{P}(\mathcal{M})}I_{\mathcal{C}_n}(M;\mathbf{Z})\leq\max_{P_M\in\mathcal{P}(\mathcal{M})}D\left(P_{\mathbf{Z}|M}^{(\mathcal{C}_n,\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}}^{(\mathcal{S})}\middle|P_M\right)\\&\leq\max_{P_M\in\mathcal{P}(\mathcal{M})}\sum_{m\in\mathcal{M}}P(m)\max_{\tilde{m}\in\mathcal{M}}D\left(P_{\mathbf{Z}|M=\tilde{m},\mathcal{C}_n}^{(\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}}^{(\mathcal{S})}\right)\\&=\max_{m\in\mathcal{M}}D\left(P_{\mathbf{Z}|M=m,\mathcal{C}_n}^{(\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}}^{(\mathcal{S})}\right).\end{aligned} \tag{17}$$

For any $\emptyset\neq\mathcal{A}\subseteq[1:n]$ and $\mathbf{z}\in\mathcal{Z}^n$, let $\mathbf{z}^{\mathcal{A}}\triangleq(z_i)_{i\in\mathcal{A}}$ be the sub-vector of $\mathbf{z}$ indexed by the elements of $\mathcal{A}$. The relative entropy chain rule further simplifies the RHS of (17) as follows. For any $m\in\mathcal{M}$, we have

$$\begin{aligned}D\left(P_{\mathbf{Z}|M=m,\mathcal{C}_n}^{(\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}}^{(\mathcal{S})}\right)&=D\left(P_{\mathbf{Z}^{\mathcal{S}},\mathbf{Z}^{\mathcal{S}^c}|M=m,\mathcal{C}_n}^{(\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}^{\mathcal{S}},\mathbf{Z}^{\mathcal{S}^c}}^{(\mathcal{S})}\right)\\&\overset{(a)}{=}D\left(P_{\mathbf{Z}^{\mathcal{S}}|M=m,\mathcal{C}_n}^{(\mathcal{S})}\middle|\middle|\Gamma_{\mathbf{Z}^{\mathcal{S}}}^{(\mathcal{S})}\right)\\&\overset{(b)}{=}D\left(P_{\mathbf{Z}^{\mathcal{S}}|M=m,\mathcal{C}_n}^{(\mathcal{S})}\middle|\middle|\mathcal{I}_Z^{\mu}\right),\end{aligned} \tag{18}$$

where (a) is because $P_{\mathbf{Z}^{\mathcal{S}^c}|M=m,\mathbf{Z}^{\mathcal{S}}=\mathbf{z}^{\mathcal{S}}}^{(\mathcal{C}_n,\mathcal{S})}=\mathbb{1}_{\big\{Z_i=?,\ i\in\mathcal{S}^c\big\}}=\Gamma_{\mathbf{Z}^{\mathcal{S}^c}}^{(\mathcal{S})}$, for every $\mathbf{z}^{\mathcal{S}}\in\mathcal{Z}^{|\mathcal{S}|}$, and (b) follows from (14).

Combining (16)-(18) while maximizing over all subsets $\mathcal{S}$, for every $\mathcal{C}_n$ we have

$$\text{Sem}_\mu(\mathcal{C}_n)\leq\max_{\substack{m\in\mathcal{M},\\ \mathcal{S}\subseteq[1:n]:\ |\mathcal{S}|=\mu}}D\left(P_{\mathbf{Z}^{\mathcal{S}}|M=m,\mathcal{C}_n}^{(\mathcal{S})}\middle|\middle|\mathcal{I}_Z^{\mu}\right). \tag{19}$$

Now, for an arbitrary $\tilde{\gamma}>0$ (to be determined later) consider

$$\begin{aligned}&\mathbb{P}\Big(\Big\{\text{Sem}_\mu(\mathbb{C}_n)\leq e^{-n\tilde{\gamma}}\Big\}^c\Big)\\&\overset{(a)}{\leq}\mathbb{P}\left(\max_{\substack{m\in\mathcal{M},\\ \mathcal{S}\subseteq[1:n]:\ |\mathcal{S}|=\mu}}D\left(P_{\mathbf{Z}^{\mathcal{S}}|M=m}^{(\mathbb{C}_n,\mathcal{S})}\middle|\middle|\mathcal{I}_Z^{\mu}\right)>e^{-n\tilde{\gamma}}\right)\\&=\mathbb{P}\left(\bigcup_{\substack{m\in\mathcal{M},\\ \mathcal{S}\subseteq[1:n]:\ |\mathcal{S}|=\mu}}\left\{D\left(P_{\mathbf{Z}^{\mathcal{S}}|M=m}^{(\mathbb{C}_n,\mathcal{S})}\middle|\middle|\mathcal{I}_Z^{\mu}\right)>e^{-n\tilde{\gamma}}\right\}\right)\\&\overset{(b)}{\leq}\sum_{\substack{m\in\mathcal{M},\\ \mathcal{S}\subseteq[1:n]:\ |\mathcal{S}|=\mu}}\mathbb{P}\Big(D\left(P_{\mathbf{Z}^{\mathcal{S}}|M=m}^{(\mathbb{C}_n,\mathcal{S})}\middle|\middle|\mathcal{I}_Z^{\mu}\right)>e^{-n\tilde{\gamma}}\Big),\end{aligned} \tag{20}$$

where (a) uses (19), and (b) is the union bound.

Each term in the sum on the RHS of (20) falls into the framework of the stronger soft-covering lemma, with respect to a blocklength of μ and the identity channel. Noting that $|\mathcal{W}| = 2^{n\tilde{R}} = 2^{\mu\frac{n\tilde{R}}{\mu}}$, we have that as long as

$$\frac{n\tilde{R}}{\mu} > H(X), \tag{21}$$

there exist $\gamma_1, \gamma_2 > 0$ that for sufficiently large n satisfy

$$\mathbb{P}\left(D\left(P^{(\mathbb{C}_n,\mathcal{S})}_{\mathbf{Z}^{\mathcal{S}}|M=m} \middle|\middle| \mathcal{I}^{\mu}_{Z} \right) > e^{-n\gamma_1} \right) \leq e^{-e^{n\gamma_2}}. \tag{22}$$

Since $\mu = \lfloor \alpha n \rfloor \leq \alpha n$, taking $\tilde{R} > \alpha H(X)$ satisfies (21). Setting $\tilde{\gamma} = \gamma_1$ into (20) and using (22), gives

$$\mathbb{P}\Big(\mathrm{Sem}_{\mu}(\mathbb{C}_n) \geq e^{-n\gamma_1} \Big) \leq 2^n \cdot 2^{nR} \cdot e^{-e^{n\gamma_2}} \xrightarrow[n\to\infty]{} 0. \tag{23}$$

Having (12) and (23), the selection lemma [23, Lemma 5] implies that there exists a sequence of (n, R) α-semantically-secure codes $\big\{\mathbb{C}_n\big\}_{n\geq 1}$ with a vanishing average error probability. The expurgation technique (namely, discarding the worst half of the codewords in each $\mathcal{C}_n$, while inflicting a negligible rate-loss) lets one upgrade $\big\{\mathcal{C}_n\big\}_{n\geq 1}$ to be reliable with respect to the maximal error probability while preserving SS. Combining both rate bounds shows the achievability of

$$R < \max_{Q_X} \Big[I(X;Y) - \alpha H(X) \Big]. \tag{24}$$

Finally, we prefix a DMC $Q_{X|V}$ to the original WTC II to obtain a new main channel $Q_{Y|V}$, given by

$$Q^n_{Y|V}(\mathbf{y}|\mathbf{v}) = \sum_{\mathbf{x}\in\mathcal{X}^n} Q^n_{X|V}(\mathbf{x}|\mathbf{v}) Q^n_{Y|X}(\mathbf{y}|\mathbf{x}). \tag{25}$$

Furthermore, $\Gamma^{(\mathcal{S})}_{\mathbf{Z}}$ from (14) is redefined as

$$\Gamma^{(\mathcal{S})}_{\mathbf{Z}}(\mathbf{z}) = \prod_{j\in\mathcal{S}^c} \mathbb{1}_{\{z_j=?\}} \prod_{j\in\mathcal{S}} Q_Z(z_j), \quad \forall \mathbf{z}\in\mathcal{Z}^n, \tag{26}$$

where Q_Z is given by

$$Q_Z(z) = \begin{cases} \sum_{v\in\mathcal{V}} Q_V(v) Q_{X|V}(z|v), & z\in\mathcal{X} \\ 0, & z = ? \end{cases}. \tag{27}$$

Repeating a similar analysis as above shows that reliability is achieved if

$$R + \tilde{R} < I(V;Y), \tag{28}$$

while the rate needed for the stronger soft-covering lemma is

$$\tilde{R} > \alpha I(V;X). \tag{29}$$

Putting (28)-(29) together yields that any rate $R \in \mathbb{R}_+$ satisfying

$$R < \max_{\substack{Q_{V,X}: \\ V-X-Y}} \Big[I(V;Y) - \alpha I(V;X) \Big], \tag{30}$$

is strongly α-SS-achievable and established Theorem 3.

ACKNOWLEDGMENT

The work of Z. Goldfeld and H. H. Permuter was supported by the Israel Science Foundation (grant no. 684/11), an ERC starting grant and the Cyber Security Research Grant at Ben-Gurion University of the Negev. The work of P. Cuff was supported by the National Science Foundation (grant CCF-1350595) and the Air Force Office of Scientific Research (grant FA9550-15-1-0180).

REFERENCES

[1] M. Bellare and S. Tessaro. Polynomial-time, semantically-secure encryption achieving the secrecy capacity. *Cryptology ePrint Archive*, Report 2012/022, 2012. Available at http://eprint.iacr.org/.

[2] M. Bellare, S. Tessaro, and A. Vardy. A cryptographic treatment of the wiretap channe. In *Proc. Adv. Crypto. (CRYPTO 2012)*, Santa Barbara, CA, USA, Aug. 2012.

[3] S. Goldwasser and S. Micali. Probabilistic encryption. *Journal of Comp. and Sys. Sci.*, 28(2):270–299, Apr. 1984.

[4] M. Bellare, A. Desai, E. Jokipii, and P. Rogaway. A concrete security treatment of symmetric encryption. In *Proc. IEEE 38th Symp. Foundations of Comp. Sci.*, pages 394–403, Miami, Florida, US, Oct. 1997.

[5] A. D. Wyner. The wire-tap channel. *Bell Sys. Techn.*, 54(8):1355–1387, Oct. 1975.

[6] I. Csiszár and J. Körner. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory*, 24(3):339–348, May 1978.

[7] L. Ozarow and A. D. Wyner. Wire-tap channel II. *Bell Sys. Techn. Journal*, 63(10):2135–2157, Dec. 1984.

[8] M. Nafea and A. Yener. Wiretap channel II with a noisy main channel. In *Proc. Int. Symp. Inf. Theory (ISIT-2015)*, Hong-Kong, Jun. 2015.

[9] M. J. Mihaljević. On message protection in cryptosystems modeled as the generalized wire-tap channel II. In *Lecture Notes in Computer Science*, pages 13–24, Berlin, Germany, 1994. Springer-Verlag.

[10] Y. Luo, C. Mitrpant, and A. J. H. Vinck. Some new characters on the wire-tap channel of type II. *IEEE Trans. Inf. Theory*, 51(3):1222–, Mar. 2005.

[11] R. Liu, Y. Liang, and H. V. Poor. Secure nested codes for type II wiretap channels. In *Proc. Inf. Theory Workshop (ITW-2007)*, Lake Tahoe, California, US, Sep. 2007.

[12] V. Aggarwal, L. Lai A. R. Calderbank, and H. V. Poor. Wiretap channel type II with an active eavesdropper. In *Proc. Int. Symp. Inf. Theory (ISIT-2009)*, Seoul, Korea, Jun.-Jul. 2009.

[13] A. D. Wyner. The common information of two dependent random variables. *IEEE Trans. Inf. Theory*, 21(2):163–179, Mar. 1975.

[14] M. Bloch and N. Laneman. Strong secrecy from channel resolvability. *IEEE Trans. Inf. Theory*, 59(12):8077–8098, Dec. 2013.

[15] C. Schieler and P. Cuff. The henchman problem: Measuring secrecy by the minimum distortion in a list. *submitted to IEEE Trans. Inf. Theory*, 2014.

[16] C. Schieler and P. Cuff. Rate-distortion theory for secrecy systems. *IEEE Trans. Inf. Theory*, 66(12):7584–7605, Dec. 2014.

[17] T. Han and S.Verdú. Approximation theory of output statistics. *IEEE Trans. Inf. Theory*, 39(3):752–772, May 1993.

[18] P. Cuff. Distributed channel synthesis. *IEEE. Trans. Inf. Theory*, 59(11):7071–7096, Nov. 2013.

[19] E. Song, P. Cuff, and V. Poor. The likelihood encoder for lossy compression. *submitted to IEEE Trans. Inf. Theory*, Aug. 2014.

[20] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley-Interscience, New York, NY, USA, 1991.

[21] V. K. Wei. Generalized Hamming weights for linear codes. *IEEE Trans. Inf. Theory*, 37(5):1412–1418, Sep. 1991.

[22] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J-M. Merolla. Applications of LDPC codes to the wiretap channel. *IEEE Trans. Inf. Theory*, 53(8):2933–2945, Aug. 2007.

[23] Z. Goldfeld, P. Cuff, and H. H. Permuter. Semantic-security capacity for wiretap channels of type II. *Submitted to IEEE Trans. Inf. Theory*, 2015. Available on ArXiv at http://arxiv.org/abs/1509.03619.

[24] P. Cuff. Soft covering with high probability. *Accepted to IEEE Int. Symp. Inf. Theory (ISIT-2016)*, Jul. 2016.