Semantically-Secured Message-Key Trade-Off over Wiretap Channels with Random Parameters

Invited Paper

Alexander Bunin, Ziv Goldfeld, Haim H. Permuter, Shlomo Shamai (Shitz), Paul Cuff and Pablo Piantanida

Abstract We study the trade-off between secret message (SM) and secret key (SK) rates simultaneously achievable over a state-dependent (SD) wiretap channel (WTC) with non-causal channel state information (CSI) at the encoder. This model subsumes all other instances of CSI availability as special cases, and calls for an efficient utilization of the state sequence both for reliability and security purposes. An inner bound on the semantic-security (SS) SM-SK capacity region is derived based on a novel superposition coding scheme. Our inner bound improves upon the previously best known SM-SK trade-off result by Prabhakaran et al., and to the best of our knowledge, upon all other existing lower bounds for either SM or SK for this setup. The results are derived under the strict semantic-security metric that requires negligible information leakage for all message-key distributions. The achievability proof uses the strong soft-covering lemma for superposition codes.

A. Bunin · S. Shamai (Shitz) (✉)
Technion—Israel Institute of Technology, Haifa, Israel
e-mail: sshlomo@ee.technion.ac.il

A. Bunin
e-mail: albun@tx.technion.ac.il

Z. Goldfeld · H.H. Permuter
Ben-Gurion University of the Negev, Beersheba, Israel
e-mail: gziv@post.bgu.ac.il

H.H. Permuter
e-mail: haimp@bgu.ac.il

P. Cuff
Princeton University, Princeton, US
e-mail: cuff@princeton.edu

P. Piantanida
CentraleSupélec-CNRS-Université, Paris-Sud, France
e-mail: pablo.piantanida@centralesupelec.fr

M. Baldi et al. (eds.), *Proceedings of the 2nd Workshop on Communication Security*, Lecture Notes in Electrical Engineering 447,
DOI 10.1007/978-3-319-59265-7_3

1 Introduction

Modern communication systems usually present an architectural separation between error correction and data encryption. The former is typically realized at the physical layer by transforming the noisy communication channel into a reliable "bit pipe". The data encryption is implemented on top of that by applying cryptographic principles. The cryptographic approach relies on restricting the computational power of the eavesdropper. The looming prospect of quantum computers (QCs) (some companies have recently reported a working prototype of a QC with over than 1000 qbits [15, 16]), however, would boost computational abilities, rendering some critical cryptosystems insecure and weakening others.[1] Post-QC cryptography offers partial solutions that rely on larger keys, but even now considerable efforts are made to save this expensive resource.

Physical layer security (PLS) [5, 18, 28], rooted in information-theoretic (IT) principles, is an alternative approach to provably secure communication that dates back to Wyner's celebrated 1975 paper on the wiretap channel (WTC) [26]. By harnessing randomness from the noisy communication channel and combining it with proper physical layer coding, PLS guarantees protection against computationally-unlimited eavesdroppers with no requirement that the legitimate parties share a secret key (SK) in advance. The eavesdropper's computational abilities are of no consequence here since the signal he/she observes from the channel carries only negligible information about the secret data.

1.1 Background

Two fundamental questions in PLS are those of the best achievable transmission rate of a secret message (SM) over a noisy channel, and the highest attainable SK rate that distributed parties can agree upon.

1.1.1 Secret-Message Transmission

The base model for SM transmission is Wyner's WTC [26], where two legitimate parties communicate over a noisy channel in the presence of an untrusted eavesdropper. A full characterization of the secrecy capacity of WTCs that are degraded in favor

[1] More specifically, asymmetric ciphers that rely on the hardness of integer factorization or discrete logarithms can be completely broken using QCs via Shor's algorithm (or a variant thereof) [4, 22]. Symmetric encryption, on the other hand, would be weakened by QC attacks but could regain its strength by increasing the size of the key [20]. This essentially follows since a QC can search through a space of size 2^n in time $2^{\frac{n}{2}}$, so by doubling the size of the key a symmetric cryptosystem would offer the same protection versus a QC attack, as the original system did versus a classic attack.

of the legitimate parties was derived in [26]. The solution was extended to the not necessarily degraded case by Csiszár and Körner [7].

A common method used in IT security proofs that dates back to the early days of Wyner, Csiszàr and Körner, relies on evaluating rather complicated equivocation terms. Recently, however, distribution approximation arguments emerged as a tool of choice for proving security. The core result on which this approach relies is called the *soft-covering lemma* (SCL), which originated from another 1975 paper by Wyner [25]. Interestingly, while both the WTC and the SCL appear in two works by Wyner from the same year, he did not seem to make a connection between the two results (although he must have been aware of a relation).

The SCL states the distribution induced by randomly selecting a codeword from an appropriately chosen codebook and passing it through a discrete memoryless channel (DMC) will be asymptotically indistinguishable from the distribution of random noise. Wyner's original result was sharpened throughout the years to hold under stricter proximity measure between distributions [10, 11, 13, 14]. Based on these sharper versions, one can make the channel output observed by the eavesdropper in the WTC look like noise and, in particular, be approximately independent of the confidential data. More specifically, a wiretap code assigns a sub-codebook that satisfies the soft-covering phenomenon to each confidential message. To transmit a certain secret message, a codeword from its associated sub-codebook is randomly and uniformly chosen and is fed into the WTC. Consequently, the distribution induced on the output sequence observed by the eavesdropper given each confidential message is indistinguishable from the distribution of random noise. This, in particular, implies that the eavesdropper's observation is asymptotically independent of the confidential data, which implies security. The notion of soft-covering is key for deriving the results of this work.

1.1.2 Secret-Key Agreement

The study of SK agreement was pioneered by Maurer [19], and independently by Ahlswede and Csiszár [1], who studied the achievable SK rates based on correlated observations at the terminals who may communicate via a noiseless and rate unlimited public link. A characterization of the SK capacity was found in [1] for the case where only one-way public communication is allowed. If the eavesdropper does not observe a correlated source, thus having access only to the public communication, the optimal SK agreement protocol uses Slepian-Wolf coding [23] for lossless reconstruction with side information. When the eavesdropper also observes a correlated source, a superposition coding scheme combined with Wyner-Ziv coding [27] is needed to achieve optimality. The inner layer of the code carries no secret information. It is designed to glut the eavesdropper with redundant information, thereby wasting his/hers resources. The confidential data is encoded in the outer layer of the superposition code and is protected by virtue of random binning. A generalization to the case where the public link is of finite capacity is due to Csiszár and Narayan [8]. If the encoder controls its source (rather than just observing it), this source becomes

a channel input and the setup evolves to a WTC. This is a special case of the so called SK channel-type model that was also studied in [1].

1.2 Model and Contributions

A more general framework to consider is a state-dependent (SD) WTC with non-causal encoder channel state information (CSI) (sometimes referred to as the Gelfand and Pinsker (GP) WTC, due to the study of the corresponding point-to-point scenario by the aforementioned authors [9]). The dependence of the channel on the state accounts for the possible availability of correlated sources observations at the terminals.

The similarity between the SM transmission and the SK agreement tasks makes their integration in a single model only natural. Adhering to the most general framework, we study the trade-off between the SM-SK rates that are simultaneously achievable over a SD-WTC with non-causal encoder CSI. The scenario where there is only a SM was considered in [6], where an achievable SM rate formula was established. This result was recently improved upon in [12] based on a novel superposition coding scheme. SK agreement over the GP-WTC was the focus of [17], and more recently was also studied in [2] (see also references therein). The combined model was considered by Prabhakaran et al. [21], who derived a benchmark inner bound on the SK-SM capacity region. The result from [21] was shown to be optimal for various special cases. We propose a novel superposition coding scheme for the combined model that not only subsumes [21] as a special case, but also captures [2, 6, 12, 17] and, to the best of our knowledge, all other existing achievability results for SM transmission, SK agreement or both.

Our coding scheme uses an over-populated superposition codebook that encodes the entire confidential message in its outer layer. Using the redundancies in the inner and outer layers, the transmission is correlated with the state sequence by means of the likelihood encoder [24]. Although the redundancy indices are chosen as part of the encoding process (rather than by the user), via the strong soft-covering lemma (SCL) for superposing codes [12, Lemma 1], we show that their true distribution is well approximated by a uniform distribution. Consequently, as long as a certain redundancy index is kept secret (along with the confidential message) from the eavesdropper, it may be declared as a SK. The security analysis is based on constructing the inner codebook such that it is better observable by the eavesdropper, making the inner layer index decodable by him. This enhances the secrecy resources that the legitimate parties can extract from the outer layer, which they use to secure the SM and part of the redundancy index of the outer layer. The encoder and decoder then declare the secured redundancy index as the SK. The agreed SK may be used to further boost the SM rate by encrypting part of the message using a one-time pad and transmitting it over the inner (unsecured) layer.

Our results are derived under the strict metric of semantic-security (SS). The SS criterion is a cryptographic gold standard that was adapted to the information-theoretic framework (of computationally unbounded adversaries) in [3]. As was shown in [3], SS is equivalent to a negligible mutual information (MI) between the confidential information (in our case, the SM-SK pair) and the eavesdropper's observations for all message-key distributions. The proof of SS relies on the strong SCL for superposition [12, Lemma 1] and the heterogeneous SCL [10, Lemma 1]. Since most of the past secrecy results mentioned above were derived under the weak-secrecy metric (i.e., a vanishing *normalized* MI with respect to a *uniformly distributed* message-key pair), our achievability outperforms the schemes from [2, 6, 17, 21] for the SD-WTC with non-causal encoder CSI not only in terms of the achievable rate pairs, but also in the upgraded sense of security it provides.

1.3 Organization

This paper is organized as follows. Section 2 establishes notations and preliminary definitions. Section 3 describes the SD-WTC setting and states an inner bound on SM-SK optimal trade-off region. In Sect. 4 we discuss past results that are captured within our framework. An outline of the proof of our main result is the content of Sect. 5. Finally, Sect. 6 summarizes the main achievements and insights of this work.

2 Preliminaries

We use the following notations. As customary $\mathbb{N}$ is the set of natural numbers (which does not include 0), while $\mathbb{R}$ are the reals. We further define $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$. Given two real numbers a, b, we denote by $[a : b]$ the set of integers $\left\{n \in \mathbb{N} \middle| \lceil a \rceil \leq n \leq \lfloor b \rfloor\right\}$. Calligraphic letters denote sets, e.g., $\mathcal{X}$, while $|\mathcal{X}|$ stands for its cardinality. $\mathcal{X}^n$ denotes the n-fold Cartesian product of $\mathcal{X}$. An element of $\mathcal{X}^n$ is denoted by $x^n = (x_1, x_2, \ldots, x_n)$; whenever the dimension n is clear from the context, vectors (or sequences) are denoted by boldface letters, e.g., $\mathbf{x}$.

Let $\left(\mathcal{X}, \mathcal{F}, \mathbb{P}\right)$ be a probability space, where $\mathcal{X}$ is the sample space, $\mathcal{F}$ is the σ-algebra and $\mathbb{P}$ is the probability measure. Random variables over $\left(\mathcal{X}, \mathcal{F}, \mathbb{P}\right)$ are denoted by uppercase letters, e.g., X, with conventions for random vectors similar to those for deterministic sequences. The probability of an event $\mathcal{A} \in \mathcal{F}$ is denoted by $\mathbb{P}(\mathcal{A})$, while $\mathbb{P}(\mathcal{A}|\mathcal{B})$ denotes conditional probability of $\mathcal{A}$ given $\mathcal{B}$. We use $\mathbb{1}_{\mathcal{A}}$ to denote the indicator function of $\mathcal{A} \in \mathcal{F}$. The set of all probability mass functions (PMFs) on a finite set $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$. PMFs are denoted by the letters such as p or q, with a subscript that identifies the random variable and its possible conditioning. For example, for a two discrete correlated random variables X and Y over the same probability space, we use p_X, $p_{X,Y}$ and $p_{X|Y}$ to denote, respectively, the marginal PMF of X, the joint PMF of (X, Y) and the conditional PMF of X

given Y. In particular, $p_{X|Y}$ represents the stochastic matrix whose elements are given by $p_{X|Y}(x|y) = \mathbb{P}\big(X = x|Y = y\big)$. Expressions such as $p_{X,Y} = p_X p_{Y|X}$ are to be understood to hold pointwise, i.e., $p_{X,Y}(x, y) = p_X(x)p_{Y|X}(y|x)$, for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Accordingly, when three random variables X, Y and Z satisfy $p_{X|Y,Z} = p_{X|Y}$, they form a Markov chain, which we denote by $X - Y - Z$. We omit subscripts if the arguments of a PMF are lowercase versions of the random variables.

For a sequence of random variable X^n, if the entries of X^n are drawn in an identically and independently distributed (i.i.d.) manner according to p_X, then for every $\mathbf{x} \in \mathcal{X}^n$ we have $p_{X^n}(\mathbf{x}) = \prod_{i=1}^n p_X(x_i)$ and we write $p_{X^n}(\mathbf{x}) = p_X^n(\mathbf{x})$. Similarly, if for every $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ we have $p_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n p_{Y|X}(y_i|x_i)$, then we write $p_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) = p_{Y|X}^n(\mathbf{y}|\mathbf{x})$. The conditional product PMF $p_{Y|X}^n$ given a specific sequence $\mathbf{x} \in \mathcal{X}^n$ is denoted by $p_{Y|X=\mathbf{x}}^n$.

The empirical PMF $\nu_{\mathbf{x}}$ of a sequence $\mathbf{x} \in \mathcal{X}^n$ is $\nu_{\mathbf{x}}(x) \triangleq \frac{N(x|\mathbf{x})}{n}$, where $N(x|\mathbf{x}) = \sum_{i=1}^n \mathbb{1}_{\{x_i=x\}}$. We use $\mathcal{T}_\varepsilon^n(p_X)$ to denote the set of letter-typical sequences of length n with respect to the PMF p_X and the non-negative number ε, i.e., we have

$$\mathcal{T}_\varepsilon^n(p_X) = \Big\{\mathbf{x} \in \mathcal{X}^n \Big| \big|\nu_{\mathbf{x}}(x) - p_X(x)\big| \le \varepsilon p_X(x),\ \forall x \in \mathcal{X}\Big\}. \tag{1}$$

Definition 1 *(Total Variation)* Let $(\mathcal{X}, \mathcal{F})$ be a measurable space and p and q be two probability measures on $\mathcal{F}$. The total variation between p and q is $||p - q||_{\mathsf{TV}} = \sup_{\mathcal{A}\in\mathcal{F}} \big|p(\mathcal{A}) - q(\mathcal{A})\big|$. If the sample space $\mathcal{X}$ is countable, the total variation reduces to $||p - q||_{\mathsf{TV}} = \frac{1}{2}\sum_{x\in\mathcal{X}} \big|p(\{x\}) - q(\{x\})\big|$.

3 SM-SK Trade-Off over Wiretap Channels with Non-Causal Encoder CSI

We study the SD-WTC with non-causal encoder CSI, for which we establish a novel achievable region of semantically-secured message-key pairs that subsumes the previously best known coding schemes for this scenario.

3.1 Problem Setup

Let $\mathcal{S}$, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ be finite sets. The $\big(\mathcal{S}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}, W_S, W_{Y,Z|X,S}\big)$ discrete and memoryless SD-WTC with non-causal encoder CSI is illustrated in Fig. 1. A state sequence $\mathbf{s} \in \mathcal{S}^n$ is generated in an i.i.d. manner according to W_S and is revealed in a non-causal fashion to the sender, who chooses a message m from the set $\big[1 : 2^{nR_M}\big]$. The sender then maps the observed state sequence $\mathbf{s}$ and the chosen message m into a channel input sequence $\mathbf{x} \in \mathcal{X}^n$ and a key index $k \in \big[1 : 2^{nR_K}\big]$ (the mapping may be random). The sequence $\mathbf{x}$ is transmitted over the SD-WTC with transition probability $W_{Y,Z|X,S}$. The output sequences $\mathbf{y} \in \mathcal{Y}^n$ and $\mathbf{z} \in \mathcal{Z}^n$ are observed by the receiver and

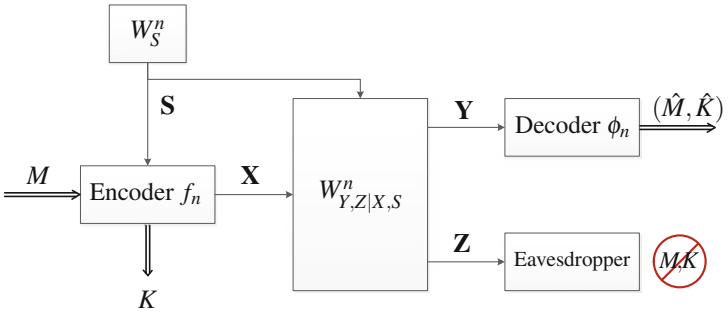

Fig. 1 The state-dependent wiretap channel with non-casual encoder channel state information

the eavesdropper, respectively. Based on $\mathbf{y}$, the receiver produces the estimates pair $(\hat{m}, \hat{k})$ of (m, k). The eavesdropper tries to glean whatever it can about the message and the generated key from $\mathbf{z}$.

Remark 1 (Most General Model) Before rigorously defining the setup and stating the result, we note that the considered model is the most general instance of a SD-WTC with non-causal CSI known at some or all of the terminals. The broadest model one may consider is when the SD-WTC $W_{\tilde{Y},\tilde{Z}|X,S_1,S_2,S_3}$ is driven by a triple of correlated state random variables $(S_1, S_2, S_3) \sim W_{S_1,S_2,S_3}$, where S_1 is known to the transmitter, S_2 is known to the receiver and S_3 is available at the eavesdropper's site. However, setting $S = S_1$, $Y = (\tilde{Y}, S_2)$, $Z = (\tilde{Z}, S_3)$ in SD-WTC with non-causal encoder CSI and defining the channel's transition probability as

$$W_{Y,Z|X,S} = W_{(\tilde{Y},S_2),(\tilde{Z},S_3)|X,S_1} = W_{S_2,S_3|S_1} W_{\tilde{Y},\tilde{Z}|X,S_1,S_2,S_3},$$

one clearly recovers this (prima facie) general SD-WTC from the model with non-causal encoder CSI only.

Definition 2 *(Code)* An (n, R_M, R_K)-code c_n for the SD-WTC with non-causal encoder CSI has a message set $\mathcal{M}_n \triangleq \left[1 : 2^{nR_M}\right]$, a key set $\mathcal{K}_n \triangleq \left[1 : 2^{nR_K}\right]$, a stochastic encoder $f_n : \mathcal{M}_n \times \mathcal{S}^n \to \mathcal{P}(\mathcal{K}_n \times \mathcal{X}^n)$ and a decoder $\phi_n : \mathcal{Y}^n \to \mathcal{M}_n \times \mathcal{K}_n$.

For any message distribution $P_M \in \mathcal{P}(\mathcal{M}_n)$ and any (n, R_M, R_K)-code c_n, the induced joint PMF is:

$$p^{(c_n)}(\mathbf{s}, m, k, \mathbf{x}, \mathbf{y}, \mathbf{z}, \hat{m}, \hat{k}) = W_S^n(\mathbf{s}) P_M(m) f_n(k, \mathbf{x}|m, \mathbf{s}) W^n_{Y,Z|X,S}(\mathbf{y}, \mathbf{z}|\mathbf{x}, \mathbf{s}) \times \mathbb{1}_{\left\{(\hat{m},\hat{k})=\phi_n(\mathbf{y})\right\}}. \quad (2)$$

The performance of c_n is evaluated in terms of its rate pair (R_M, R_K), its maximal decoding error probability, the maximal distance of the distribution of K from being uniform and independent of M, and the SS-metric.

Definition 3 *(Maximal Error Probability)* The maximal error probability of an (n, R_M, R_K)-code c_n is $e(c_n) = \max_{m \in \mathcal{M}_n} e_m(c_n)$, where:

$$e_m(c_n) = \sum_{\substack{(\mathbf{s},k,\mathbf{x}) \\ \in \mathcal{S}^n \times \mathcal{K}_n \times \mathcal{X}^n}} W_S^n(\mathbf{s})\, f_n(k, \mathbf{x}|m, \mathbf{s}) \sum_{\substack{\mathbf{y} \in \mathcal{Y}^n: \\ \phi_n(\mathbf{y}) \neq (m,k)}} W_{Y|X,S}^n(\mathbf{y}|\mathbf{x}, \mathbf{s})$$

Definition 4 *(Maximal Distance to Key Uniformity)* The maximal distance to key uniformity and independence of the message of an (n, R_M, R_K)-code c_n is $\delta(c_n) = \max_{m \in \mathcal{M}_n} \delta_m(c_n)$, where $\delta_m(c_n) = \left|\left| p_{K|M=m}^{(c_n)} - p_{\mathcal{K}_n}^{(U)} \right|\right|_{\mathsf{TV}}$ and $p_{\mathcal{K}_n}^{(U)}$ is the uniform PMF over $\mathcal{K}_n$.

Definition 5 *(Information Leakage and SS Metric)* The information leakage to the eavesdropper under the (n, R_M, R_K)-code c_n and the message-key PMF $p_M \in \mathcal{P}(\mathcal{M}_n)$ is $\ell(p_M, c_n) = I_{c_n}(M, K; \mathbf{Z})$, where I_{c_n} denotes that the MI is taken with respect to the marginal $p_{M,K,\mathbf{Z}}^{(c_n)}$ of (2). The SS metric with respect to c_n is[2] $\ell_{\mathsf{Sem}}(c_n) = \max_{p_M \in \mathcal{P}(\mathcal{M}_n)} \ell(p_M, c_n)$.

Definition 6 *(Achievability)* A pair $(R_M, R_K) \in \mathbb{R}_+^2$ is called an achievable SS message-key pair for the SD-WTC with non-causal encoder CSI, if for every $\varepsilon > 0$ and sufficiently large n, there exists a (n, R_M, R_K)-code c_n with $e(c_n) \leq \varepsilon, \delta(c_n) \leq \varepsilon$ and $\ell_{\mathsf{Sem}}(c_n) \leq \varepsilon$.

Definition 7 *(SS-Capacity)* The SS message-key capacity region $\mathcal{C}_{\mathsf{Sem}}$ of the SD-WTC with non-causal encoder CSI is the closure of the set of achievable rate pairs.

3.2 Main Results

The main result of this work is a novel inner bound on the SS message-key capacity region of the SD-WTC with non-causal encoder CSI. Our achievable region is at least as good as the best known achievability results for the considered problem. To state our main result, let $\mathcal{U}$ and $\mathcal{V}$ be finite alphabets and for any $q_{U,V,X|S} : \mathcal{S} \to \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{X})$ define

$$\mathcal{R}_{\mathsf{A}}\left(q_{U,V,X|S}\right) \triangleq \left\{ (R_M, R_K) \in \mathbb{R}_+^2 \,\middle|\, \begin{array}{l} R_M \leq I(U,V;Y) - I(U,V;S) \\ R_M + R_K \leq I(V;Y|U) - I(V;Z|U), \\ R_M + R_K \leq I(U,V;Y) - I(V;Z|U) - I(U;S) \end{array} \right\}, \quad (3)$$

[2] $\ell_{\mathsf{Sem}}(c_n)$ is actually the mutual-information-security (MIS) metric, which is equivalent to SS by [3]. We use this representation rather than the formal definition of SS (see, e.g., [3, Eq. (4)]) out of analytical convenience.

where the MI terms are calculated with respect to the joint PMF $W_S q_{U,V,X|S} \times W_{Y,Z|X,S}$, i.e., where $(U, V) - (X, S) - (Y, Z)$ forms a Markov chain.

Theorem 1 (Semantic-Security SM-SK Capacity Inner Bound) *The following inclusion holds:*

$$\mathcal{C}_{\mathsf{Sem}} \supseteq \mathcal{R}_{\mathsf{A}} \triangleq \bigcup_{q_{U,V,X|S}} \mathcal{R}_{\mathsf{A}}\left(q_{U,V,X|S}\right). \tag{4}$$

An extended outline of the proof of Theorem 1 is given in Sect. 5, and is based on a secured superposition coding scheme. An over-populated two-layered superposition codebook is constructed (independently of the state sequence), in which the entire secret message is encoded in the *outer layer*, meaning no information is carried by the inner layer. The likelihood encoder [24] uses the redundancies in the inner and outer codebooks to correlate the transmitted codewords with the observed state sequence. Upon doing so, part of the correlation index from the outer layer is declared by the encoder as the key. The inner layer is designed to utilize the part of the channel which is better observable by the eavesdropper. This saturates the eavesdropper with redundant information and leaves him/her with insufficient resources to gather any information on the SM-SK pair from the outer layer. The legitimate decoder, on the other hand, decodes both layers of the codebook and declares the appropriate indices as the decoded message-key pair.

Remark 2 (Interpretation of Theorem 1) To get some intuition on the result of Theorem 1, we examine $\mathcal{R}_{\mathsf{A}}(q_{U,V,X|S})$ from two different perspectives: when the joint PMF $W_S q_{U,V,X|S} W_{Y,Z|X,S}$ satisfies $I(U;Y) \geq I(U;S)$, or when the opposite inequality holds.

If $I(U;Y) \geq I(U;S)$, the third rate bound in $\mathcal{R}_{\mathsf{A}}(q_{U,V,X|S})$ becomes redundant and the dominating bounds are

$$R_M \leq I(U,V;Y) - I(U,V;S) \tag{5a}$$
$$R_M + R_K \leq I(V;Y|U) - I(V;Z|U). \tag{5b}$$

The right-hand side (RHS) of (5a) is the total rate of reliable (secured and unsecured) communication that our superposition codebook supports. This clearly bounds the rate of the SM that may be transmitted. For (5b), the MI difference on the RHS is the total rate of secrecy resources that are produced by the outer layer of the codebook. Since the security of our SM-SK pair all comes from that outer layer, this MI difference is an upper bound on the sum of rates.

For the opposite case when $I(U;Y) < I(U;S)$, the second inequality in $\mathcal{R}_{\mathsf{A}}$ becomes redundant and we are left with

$$R_M \leq I(U,V;Y) - I(U,V;S) \tag{6a}$$
$$R_M + R_K \leq I(V;Y|U) - I(V;Z|U) - \big[I(U;S) - I(U;Y)\big]. \tag{6b}$$

While the interpretation of (6a) remains as before, to understand (6b), consider the following. Since $I(U;S)$ is approximately the rate of the inner codebook, $I(U;Y) < I(U;S)$ means that looking solely on the inner layer, the decoder is lacking the resolution to decode it. Yet, the success of our communication protocol relies on the decoder reliably decoding both layers. Therefore, in this case, some of the rate from the outer layer is allocated to convey the inner layer index. Recalling that our security analysis is based on revealing the inner layer to the eavesdropper, this rate allocation effectively results in a loss of $I(U;S) - I(U;Y)$ in the secrecy resources of the outer layer, giving rise to the rate bound from (6b).

4 Past Results as Special Cases

4.1 Prabhakarn's SM-SK Trade-Off Region

The result of Theorem 1 recovers the previously best known achievable SM-SK trade-off region over the SD-WTC with non-causal encoder CSI from [21]. In [21, Theorem 1] the following region was established as an inner bound on the SM-SK trade-off capacity region:

$$\mathcal{R}_{\mathsf{PER}} \triangleq \bigcup_{q_U \times q_{V,X|U,S}} \mathcal{R}_{\mathsf{PER}}\left(q_U \times q_{V,X|U,S}\right), \tag{7a}$$

where for any $q_U \in \mathcal{P}(\mathcal{U})$ and $q_{V,X|U,S} : \mathcal{U} \times \mathcal{S} \to \mathcal{P}(\mathcal{V} \times \mathcal{X})$,

$$\mathcal{R}_{\mathsf{PER}}\left(q_U \times q_{V,X|U,S}\right) \triangleq \left\{ (R_M, R_K) \in \mathbb{R}_+^2 \,\middle|\, \begin{array}{l} R_M \le I(U,V;Y) - I(U,V;S) \\ R_M + R_K \le I(V;Y|U) - I(V;Z|U) \end{array} \right\}, \tag{7b}$$

and the MI terms are taken with respect to $W_S q_U q_{V,X|U,S} W_{Y,Z|X,S}$, i.e., U and S are independent and $(U,V) - (X,S) - (Y,Z)$ forms a Markov chain.

First note that Theorem 1 recovers $\mathcal{R}_{\mathsf{PER}}$ by restricting U to be independent of S in $\mathcal{R}_{\mathsf{A}}$. This is since for an independent pair (U,S), we have $I(U;S) = 0$, while $I(U,V;Y) \ge I(V;Y|U)$ always holds. This makes the third rate bound in $\mathcal{R}_{\mathsf{A}}$ redundant and $\mathcal{R}_{\mathsf{PER}}$ is recovered.

The result from [21] was derived under the weak-secrecy metric (i.e., a vanishing *normalized* MI between the SM-SK pair and the eavesdropper's observation sequence $\frac{1}{n} I(M,K;\mathbf{Z})$ where the message is assumed to be uniformly distributed). Our achievability, on the other hand, ensures performance with respect to the stringent SS-metric. Since Theorem 1 captures [21, Theorem 1] as a special case, it also upgrades its result to SS.

4.2 SM Transmission over SD-WTCs

In [12, Theorem 1] a lower bound on the SS-capacity of a SM transmission over the considered SD-WTC was established. The model considered in [12] is recovered from the one considered here by removing the SK ($R_K = 0$). The SS-capacity of a SM transmission was shown to be lower bounded as

$$C_{\mathsf{SM-Sem}} \geq R_{\mathsf{GCP}} \triangleq \max_{q_{U,V,X|S}} R_{\mathsf{GCP}}\left(q_{U,V,X|S}\right), \tag{8a}$$

where for any $q_{U,V,X|S} : \mathcal{S} \to \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{X})$,

$$R_{\mathsf{GCP}}\left(q_{U,V,X|S}\right) \triangleq \min \left\{ \begin{array}{l} I(U,V;Y) - I(U,V;S), \\ I(V;Y|U) - I(V;Z|U), \\ I(U,V;Y) - I(V;Z|U) - I(U;S) \end{array} \right\}, \tag{8b}$$

and the MI terms are taken with respect to $W_S q_{U,V,X|S} W_{Y,Z|X,S}$, i.e., $(U, V) - (X, S) - (Y, Z)$ forms a Markov chain.

R_{GCP} is the projection in the (R_M, R_K)-plane of $\mathcal{R}_{\mathsf{A}}$ from Theorem 1 to the R_M axis when $R_K = 0$. Then main difference between the coding scheme from [12] and our superposition code is the introduction of the additional index $k \in \mathcal{K}_n$ in the outer layer of the codebook (that also encodes the SM $m \in \mathcal{M}_n$). Along with the other redundancy indices, k is used to correlate the transmission with the observed state sequence via the likelihood encoder [24]. Based on distribution approximation arguments we show that K is approximately independent of the message M and approximately uniform. The pair (M, K) is known to the transmitter (who chooses them) and is reliably decoded by the receiver. Finally, by securing K along with M in our analysis, it is established as a SK.

The intuition behind the SK construction is that, unlike the message, the key does not have to be independent of the state sequence nor it is chosen by the user. Therefore, the padding that ensures the correlation with the state sequence is a valid key, as long as it is protected in the security analysis.

4.3 SK Agreement over SD-WTCs

In [2] two achievable schemes were proposed for SK agreement over a wiretap channel when the terminals have access to correlated sources. The results from [2] do not imply one another and differ in one scheme being based on source and channel separation [2, Theorem 2], while in the other the coding is done jointly [2, Theorem 3].

The setup in [2] consists of three correlated sources S_x, S_y and S_z that are observed by the encoder, decoder and eavesdropper, respectively, and a SD-WTC in which the triple (S_x, S_y, S_z) plays the role of the state. Our general framework is defined through

the state distribution W_S and the SD-WTC $W_{\tilde{Y},\tilde{Z}|X,S}$. Setting $S = S_x$, $\tilde{Y} = (S_y, Y)$ and $\tilde{Z} = (S_z, Z)$ recovers the model from [2] (see Remark 1).

The first scheme from [2, Theorem 2] operates under the assumption that the SD-WTC decomposes as $W_{(S_y,Y),(S_z,Z)|X,S_x} = W_{S_y,S_z|S_x} W_{Y,Z|X}$ into a product of two WTCs, one being independent of the state, while the other one depends only on it. Thus, the legitimate receiver (respectively, the eavesdropper) observes not only the output $\mathbf{Y}$ (respectively, $\mathbf{Z}$) of the WTC $W_{Y,Z|X}$, but also $\mathbf{S}_y$ (respectively, $\mathbf{S}_z$) - a noisy version of the state sequence drawn according to the marginal of $W_{S_y,S_z|S}$. This scheme shows that the SK capacity C_{SK} is lower bounded as

$$C_{\mathsf{SK}} \geq R_{\mathsf{BPS}}^{(\mathsf{Separate})} \triangleq \max \Big[I(T;Y|Q) - I(T;Z|Q) + I(\tilde{V};S_y|\tilde{U}) - I(\tilde{V};S_z|\tilde{U}) \Big] \tag{9}$$

where the maximization is over all $q_{\tilde{V}|S_x} q_{\tilde{U}|\tilde{V}} : \mathcal{S}_x \to \mathcal{P}(\tilde{\mathcal{V}} \times \tilde{\mathcal{U}})$ and $q_{Q,T} q_{X|T} \in \mathcal{P}(\mathcal{Q} \times \mathcal{T} \times \mathcal{X})$ that give rise to a joint PMF $W_{S_x,S_y,S_z} q_{\tilde{V}|S_x} q_{\tilde{U}|\tilde{V}} \times q_{Q,T} q_{X|T} W_{Y,Z|X}$ satisfying $I(\tilde{U};S_x|S_y) \leq I(Q;Y)$ and $I(\tilde{V};S_x|S_y) \leq I(T;Y)$. With respect to this distribution $(S_y, S_z) - S_x - V - U$ and $Q - T - X - (Y, Z)$ form Markov chains and (S_y, S_z, S_x, V, U) are independent of (Q, T, X, Y, Z). This independence is the essence of separation that uses the channel for two purposes: carrying communication for SK agreement based on the sources, and securing part of this communication using wiretap coding.

Setting $R_M = 0$, $U = (Q, \tilde{U})$, $V = (T, \tilde{V})$ in Theorem 1, and limiting ourselves to joint PMFs that satisfy $I(U;S_Y,Y) \geq I(U;S_x)$, while keeping the above distribution X, recovers (9).

The joint coding scheme from [2, Theorem 3] does not require sources and channel independence. i.e., no factorization property of $W_{(S_y,Y),(S_z,Z)|X,S_x}$ is assumed. It lower bounds C_{SK} as

$$C_{\mathsf{SK}} \geq R_{\mathsf{BPS}}^{(\mathsf{Joint})} \triangleq \max \Big[I(\tilde{V};S_y,Y|\tilde{U}) - I(\tilde{V};S_z,Z|\tilde{U}) \Big] \tag{10}$$

where the maximization is over all $q_{\tilde{V},X|S_x} q_{\tilde{U}|\tilde{V}} : \mathcal{S}_x \to \mathcal{P}(\tilde{\mathcal{V}} \times \mathcal{X} \times \tilde{\mathcal{U}})$ that give rise to a joint PMF $W_{S_x} q_{\tilde{V},X|S_x} q_{\tilde{U}|\tilde{V}} W_{(S_y,Y),(S_z,Z)|S_x,X}$ satisfying $I(\tilde{U};S_x) \leq I(\tilde{U};S_y,Y)$ and $I(\tilde{V};S_x|\tilde{U}) \leq I(\tilde{V};S_y,Y|\tilde{U})$. Inserting into Theorem 1 $R_M = 0$ and $(U, V) = (\tilde{U}, \tilde{V})$, where $(\tilde{U}, \tilde{V})$ is a valid auxiliary pair in $R_{\mathsf{BPS}}^{(\mathsf{Joint})}$, recovers (10). Consequently, Theorem 1 unifies the schemes from [2], and since the results from [2] are under the weak-secrecy metric, Theorem 1 also upgrades them to SS (see the discussion from Sect. 4.1).

5 Outline of Proof of Theorem 1

We give a detailed description of the codebook construction and of the encoding and decoding processes. Due to space limitation, the analysis of reliability and SS is omitted and only the required rate bounds accompanied by broad explenations are provided. Fix a conditional PMF $q_{U,V,X|S}$.

Codebook $\mathcal{C}_n$: We use a superposition codebook where the outer layer carries both the SM and the SK. The codebook is constructed independently of $\mathbf{S}$, but has sufficient redundancy to correlate the transmission with $\mathbf{S}$.

Define the index sets $\mathcal{I}_n \triangleq \left[1 : 2^{nR_1}\right]$ and $\mathcal{J}_n \triangleq \left[1 : 2^{nR_2}\right]$, and let $\mathcal{B}_U^{(n)} \triangleq \left\{\mathbf{u}(i)\right\}_{i\in\mathcal{I}_n}$ be an inner layer codebook generated as i.i.d. samples of q_U^n. For every $i \in \mathcal{I}_n$, let $\mathcal{B}_V^{(n)}(i) \triangleq \left\{\mathbf{v}(i, j, k, m)\right\}_{(j,k,m)\in\mathcal{J}_n\times\mathcal{K}_n\times\mathcal{M}_n}$ be a collection of $|\mathcal{J}_n||\mathcal{K}_n||\mathcal{M}_n|$ vectors of length n drawn according to the distribution $q_{V|U=\mathbf{u}(i)}^n$. We use $\mathcal{C}_n$ to denote our superposition codebook, i.e., the collection of the inner and all the outer layer codebooks. The encoder and decoder are described next for a fixed superposition codebook $\mathcal{C}_n$.

Encoder $f_n^{(\mathcal{C}_n)}$: The encoding phase is based on the likelihood-encoder [24], which, in turn, allows us to approximate the (rather cumbersome) induced joint distribution by a much simpler distribution which we use for the analysis. Given $m \in \mathcal{M}_n$ and $\mathbf{s} \in \mathcal{S}^n$, the encoder randomly chooses $(i, j, k) \in \mathcal{I}_n \times \mathcal{J}_n \times \mathcal{K}_n$ according to

$$p_{\mathsf{LE}}^{(\mathcal{C}_n)}(i, j, k|m, \mathbf{s}) = \frac{q_{S|U,V}^n\left(\mathbf{s}\middle|\mathbf{u}(i), \mathbf{v}(i, j, k, m)\right)}{\sum\limits_{\substack{(i',j',k')\\ \in\mathcal{I}_n\times\mathcal{J}_n\times\mathcal{K}_n}} q_{S|U,V}^n\left(\mathbf{s}\middle|\mathbf{u}(i'), \mathbf{v}(i', j', k', m)\right)} \tag{11}$$

where $q_{S|U,V}$ is the conditional marginal of $q_{S,U,V}$ defined by $q_{S,U,V}(s, u, v) = \sum_{x\in\mathcal{X}} W_S(s)q_{U,V,X|S}(u, v, x|s)$, for every $(s, u, v) \in \mathcal{S} \times \mathcal{U} \times \mathcal{V}$. The encoder declares the index $k \in \mathcal{K}_n$ chosen by the by $p_{\mathsf{LE}}^{(\mathcal{C}_n)}$ as the key. Furthermore, the channel input sequence is generated by feeding the chosen u- and v-codewords along with the state sequence into the DMC $q_{X|U,V,S}^n$.

Decoder $\phi_n^{(\mathcal{C}_n)}$: Upon observing $\mathbf{y} \in \mathcal{Y}^n$, the decoder searches for a unique tuple $(\hat{i}, \hat{j}, \hat{k}, \hat{m}) \in \mathcal{I}_n \times \mathcal{J}_n \times \mathcal{K}_n \times \mathcal{M}_n$ such that $\left(\mathbf{u}(\hat{i}), \mathbf{v}(\hat{i}, \hat{j}, \hat{k}, \hat{m}), \mathbf{y}\right) \in \mathcal{T}_\varepsilon^n(q_{U,V,Y})$. If such a unique quadruple is found, then set $\phi_n^{(\mathcal{C}_n)}(\mathbf{y}) = \left(\hat{m}, \hat{k}\right)$; otherwise, $\phi_n^{(\mathcal{B}_n)}(\mathbf{y}) = (1, 1)$.

The quadruple $(\mathcal{M}_n, \mathcal{K}_n, f_n^{(\mathcal{C}_n)}, \phi_n^{(\mathcal{C}_n)})$ defined with respect to the codebook $\mathcal{C}_n$ constitutes an (n, R_M, R_K)-code c_n.

Main ideas for the analysis: The key step is to approximate (in total variation) the joint PMF induced by the above encoding and decoding scheme, say $p^{(\mathcal{C}_n)}$, by a new distribution $\Gamma^{(\mathcal{C}_n)}$, which lands itself easier for the reliability and security analyses. For any $p_M \in \mathcal{P}(\mathcal{M}_n)$, $\Gamma^{(\mathcal{C}_n)}$ is

$$\Gamma^{(\mathcal{C}_n)}(m,i,j,k,\mathbf{u},\mathbf{v},\mathbf{s},\mathbf{x},\mathbf{y},\mathbf{z},\hat{m}) = p_M(m)\frac{1}{|\mathcal{I}_n||\mathcal{J}_n||\mathcal{K}_n|}\mathbb{1}_{\left\{\mathbf{u}=\mathbf{u}(i),\mathbf{v}=\mathbf{v}(i,j,k,m)\right\}} \times q^n_{S|U,V}(\mathbf{s}|\mathbf{u},\mathbf{v})q^n_{X|U,V,S}(\mathbf{x}|\mathbf{u},\mathbf{v},\mathbf{s})W^n_{Y,Z|X,S}(\mathbf{y},\mathbf{z}|\mathbf{x},\mathbf{s})\mathbb{1}_{\left\{\phi_n^{(\mathcal{C}_n)}(\mathbf{y})=(\hat{m},\hat{k})\right\}}. \quad (12)$$

Namely, with respect to $\Gamma^{(\mathcal{C}_n)}$, the indices $(i,j,k) \in \mathcal{I}_n \times \mathcal{J}_n \times \mathcal{K}_n$ are uniformly drawn from their respective ranges. Then, the sequence $\mathbf{s}$ is generated by feeding the corresponding u- and v-codewords into the DMC $q^n_{S|U,V}$. Based on [12, Lemma 1], it can be shown that with respect to a random superposition codebook C_n, $p^{(\mathsf{C}_n)}$ and $\Gamma^{(\mathsf{C}_n)}$ are close in total variation in several senses (both in expectation and with high probability), if

$$R_1 > I(U;S) \quad (13a)$$

$$R_1 + R_2 + R_K > I(U,V;S). \quad (13b)$$

Having this, standard properties of total variation imply that K is indeed approximately uniform and independent of M. Furthermore, based on the approximation of $p^{(\mathsf{C}_n)}$ with $\Gamma^{(\mathsf{C}_n)}$, both the reliability and the security analysis are executed with respect to $\Gamma^{(\mathsf{C}_n)}$ rather than $p^{(\mathsf{C}_n)}$. Standard joint-typicality decoding arguments for superposition codes show that reliability follows provided that

$$R_2 + R_K + R_M < I(V;Y|U), \quad (14a)$$

$$R_1 + R_2 + R_K + R_M < I(U,V;Y). \quad (14b)$$

With the help of the heterogeneous strong SCL from [10, Lemma 1], SS is ensured if

$$R_2 > I(V;Z|U). \quad (15)$$

The rate bound in (15) ensures that the distribution of the eavesdropper's observation given the inner layer codeword and each SM-SK pair is asymptotically indistinguishable form random noise. This asymptotic independence, in turn, implies semantic security.

Finally, applying the Fourier-Motzkin Elimination on (13), (14) and (15) to remove R_1 and R_2, shows that $\mathcal{R}_{\mathsf{A}}\left(q_{U,V,X|S}\right)$ is achievable.

6 Summary and Concluding Remarks

We studied the trade-off between SM and SK rates simultaneously achievable over a SD-WTC with non-causal CSI at the encoder. This model subsumes all other instances of CSI availability as special cases. An inner bound on the semantic-security SM-SK capacity region was derived based on a novel superposition coding scheme, the likelihood encoder and soft-converging arguments. We showed that our inner

bound recovers the previously best known SM-SK trade-off region by Prabhakaran et al. [21]. Furthermore, our result recovers the best lower bounds that we are aware of for either SM or SK rates achievable in this setup [2, 12]. Unlike most of the previous results that were derived under the weak secrecy metric, our derivations ensure semantic-security. It would be interesting to demonstrate a strict improvement of the scheme presented here over the results in [21].

Acknowledgements The work of Alexander Bunin and Shlomo Shamai was supported by the European Union's Horizon 2020 Research And Innovation Programme, grant agreement No. 694630. 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 Paul 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. Ahlswede R, Csiszár I (1993) Common randomness in information theory and cryptography. part i: secret sharing. IEEE Trans Inf Theory 39(4):1121–1132
2. Bassi G, Piantanida P, Shamai (Shitz) S (2016) Secret key generation over noisy channels with common randomness. ArXiv preprint arXiv.org/abs/1609.08330
3. Bellare M, Tessaro S, Vardy A (2012) A cryptographic treatment of the wiretap channel. In: Proceedings of the advances in cryptology (CRYPTO 2012), Santa Barbara, CA, USA
4. Bernstein DJ (2009) Introduction to post-quantum cryptography. In: Post-quantum cryptography. Springer, Berlin, pp 1–14
5. Bloch M, Barros J (2011) Physical-layer security: from information theory to security engineering. Cambridge University Press, Cambridge, UK
6. Chen Y, Vinck AJH (2008) Wiretap channel with side information. IEEE Trans Inf Theory 54(1):395–402
7. Csiszár I, Körner J (1978) Broadcast channels with confidential messages. IEEE Trans Inf Theory 24(3):339–348
8. Csiszár I, Narayan P (2000) Common randomness and secret key generation with a helper. IEEE Trans Inf Theory 46(2):344–366
9. Gelfand SI, Pinsker MS (1980) Coding for channel with random parameters. Problemy Pered Inform (Probl Inf Trans) 9(1): 19–31
10. Goldfeld Z, Cuff P, Permuter HH (2016) Arbitrarily varying wiretap channels with type constrained states. IEEE Trans Inf Theory 62(12):7216–7244
11. Goldfeld Z, Cuff P, Permuter HH (2016) Semantic-security capacity for wiretap channels of type II. IEEE Trans Inf Theory 62(7):3863–3879
12. Goldfeld Z, Cuff P, Permuter HH (2016) Wiretap channel with random states non-causally available at the encoder. Submitted to IEEE Trans Inf Theory
13. Han T, Verdú S (1993) Approximation theory of output statistics. IEEE Trans Inf Theory 39(3):752–772
14. Hou J, Kramer G (2013) Informational divergence approximations to product distributions. In: Proceedings of the 13th Canadian Workshop Information Theory (CWIT), Toronto, Ontario, Canada
15. Johnson MW et al (2011) Quantum annealing with manufactured spins. Nature 473(7346):194–198
16. Jones N (2013) Google and NASA snap up quantum computer D-Wave two. http://www.scientificamerican.com/article.cfm?id=google-nasa-snap-up-quantum-computer-dwave-two

17. Khisti A, Diggavi SN, Wornell GW (2011) Secret-key agreement with channel state information at the transmitter. IEEE Trans Inf Forensics Secur 6(3):672–681
18. Liu Y, Chen HH, Wang L (First quarter 2017) Physical layer security for next generation wireless networks: theories, technologies, and challenges. IEEE Commun Surv Tut 19(1): 347–376
19. Maurer UM (1993) Secret key agreement by public discussion from common information. IEEE Trans Inf Theory 39(3):733–742
20. Perlner RA, Cooper DA (2009) Quantum resistant public key cryptography: a survey. In: Proceedings of symposium on identity and trust on the internet (IDtrust). pp. 85–93. ACM, Gaithersburg, Maryland
21. Prabhakaran V, Eswaran K, Ramchandran K (2012) Secrecy via sources and channels. IEEE Trans Inf Theory 85(11):6747–6765
22. Shor PW (1999) Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM review 41(2):303–332
23. Slepian D, Wolf J (1973) Noiseless coding of correlated information sources. IEEE Trans Inf Theory 19(4):471–480
24. Song E, Cuff P, Poor V (2016) The likelihood encoder for lossy compression. IEEE Trans Inf Theory 62(4):1836–1849
25. Wyner AD (1975) The common information of two dependent random variables. IEEE Trans Inf Theory 21(2):163–179
26. Wyner AD (1975) The wire-tap channel. Bell Sys. Techn. 54(8):1355–1387
27. Wyner AD, Ziv J (1976) The rate-distortion function for source coding with side information at the decoder. IEEE Trans Inf Theory 1:1–10
28. Zeng K (2015) Physical layer key generation in wireless networks: challenges and opportunities. IEEE Commun Mag 53(6):33–39