Strong Secrecy From Channel Resolvability

Matthieu R. Bloch, *Member, IEEE*, and J. Nicholas Laneman, *Senior Member, IEEE*

***Abstract*—We analyze physical-layer security based on the premise that the coding mechanism for secrecy over noisy channels is tied to the notion of channel resolvability. Instead of considering *capacity-based* constructions, which associate to each message a subcode that operates just below the capacity of the eavesdropper's channel, we consider *channel-resolvability-based* constructions, which associate to each message a subcode that operates just above the resolvability of the eavesdropper's channel. Building upon the work of Csiszár and Hayashi, we provide further evidence that channel resolvability is a powerful and versatile coding mechanism for secrecy by developing results that hold for strong secrecy metrics and arbitrary channels. Specifically, we show that at least for symmetric wiretap channels, random capacity-based constructions fail to achieve the strong secrecy capacity, while channel-resolvability-based constructions achieve it. We then leverage channel resolvability to establish the secrecy-capacity region of arbitrary broadcast channels with confidential messages and a cost constraint for strong secrecy metrics. Finally, we specialize our results to study the secrecy capacity of wireless channels with perfect channel state information (CSI), mixed channels, and compound channels with receiver CSI, as well as the secret-key capacity of source models for secret-key agreement. By tying secrecy to channel resolvability, we obtain achievable rates for strong secrecy metrics with simple proofs.**

***Index Terms*—Channel resolvability, information-spectrum, information-theoretic security, secret-key agreement, wireless channels, wiretap channel.**

I. Introduction

In virtually every communication system, the problems of reliability and secrecy are handled in fundamentally different ways. Typically, error-correcting schemes in the physical-layer guarantee reliable communications, while encryption algorithms and key-exchange protocols in the upper layers[1] ensure data secrecy. Physical-layer security puts forward an alternative role for the physical layer, whereby reliability and secrecy can be handled jointly by means of appropriate coding schemes. The idea is to recognize the presence of noise in every communication channel, including the channel of a potential adversary who eavesdrops on transmitted signals, and to exploit knowledge of noise statistics to prevent eavesdroppers from retrieving information. Unlike most existing security schemes, physical-layer security can guarantee information-theoretic security, by which secrecy is measured quantitatively in terms of the statistical dependence between the messages transmitted and the observations of eavesdroppers.

The theoretical foundation of physical-layer security is the early works of Wyner [1] and Csiszár and Körner [2], which prove the existence of coding schemes ensuring reliability and secrecy for the wiretap channel; however, the recent surge of information-theoretic results regarding the wiretap channel has fostered few practical engineering solutions. This state of affairs is partly due to the fact that most works extend the coding schemes of [1] and [2], in which the coding mechanism that guarantees secrecy is tied to channel capacity. This mechanism will be precisely defined in Section III; at this point, suffice to say that the codes in [1] and [2] are a union of subcodes that operate just below the capacity of the eavesdropper's channel as the blocklength grows large. Although such coding schemes have been successfully used to study many multiuser information-theoretic secrecy problems [3], [4], deriving secrecy from channel capacity leaves open a few lingering issues:

1) wiretap channel models that incorporate the limitations of modern communication systems, such as the presence of memory, are difficult to analyze;
2) the results obtained by tying secrecy to channel capacity are deemed too weak for cryptographic applications.

This paper builds upon an original observation of Csiszár [5] and the work of Hayashi [6] to explore an alternative approach to physical-layer security that addresses the aforementioned issues; the premise of the approach is to relate the coding mechanism for secrecy to the notion of channel resolvability [7], [8] and not to channel capacity.

A. Motivating Examples

To motivate the usefulness of channel resolvability, we start with two intuitive examples that shed light on the mechanisms one could exploit to ensure information-theoretic security.

Example 1 (One-Time Pad): Consider a binary message $W \in \{0,1\}$ that is encoded into a codeword Z as $Z = W \oplus K$, where $K \sim \mathcal{B}\left(\frac{1}{2}\right)$ is a secret key and $\oplus$ denotes modulo-two addition. The crypto lemma [9] shows that the output distributions $p_{Z|W=0}$ and $p_{Z|W=1}$ are identical and equal to the uniform distribution on $\{0,1\}$; hence, messages are statistically indistinguishable for an eavesdropper only observing Z. From an operational perspective, note that the encoder exploits the key K to ensure that all messages induce the same output distribution.

Example 2 (Transmission Over a Noisy Gaussian Channel): Consider an uncoded message W uniformly distributed in the

Manuscript received June 13, 2012; revised April 05, 2013; accepted June 17, 2013. Date of publication September 26, 2013; date of current version November 19, 2013. This work was supported in part by the NSF under Award CCF05-44618. This paper was presented in part at the 46th Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, 2011, in part at the 1st International ICST Workshop on Secure Wireless Networks, Cachan, France, 2011, and in part at the 2011 IEEE International Symposium on Information Theory.

M. R. Bloch is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA, and also with the GT-CNRS UMI 2958, Metz, France (e-mail: matthieu.bloch@ece.gatech.edu).

J. N. Laneman is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: jnl@nd.edu).

Communicated by T. Uyematsu, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2013.2283722

[1]Specific cryptographic schemes are implemented at all upper layers of the protocol stack, including MAC, transport, network, and application layers.

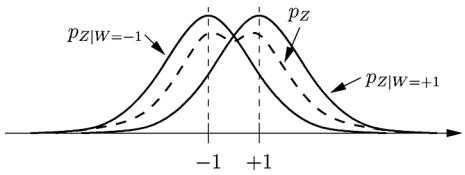

Fig. 1. Distributions of channel outputs for uncoded transmission of $\{-1,+1\}$ over an AWGN channel.

set $\{-1,+1\}$ and observed by an eavesdropper at the output of a real additive white Gaussian noise channel as $Z = W + N$, where $N \sim \mathcal{N}(0,\sigma^2)$. As illustrated in Fig. 1, the output distributions $p_{Z|W=-1}$ and $p_{Z|W=+1}$ become indistinguishable from the average distribution p_Z as the noise variance increases. Specifically, as σ goes to infinity, one can show that, for each $m \in \{-1,+1\}$, the variational distance between $p_{Z|W=m}$ and p_Z is at most $\mathcal{O}(\sigma^{-\frac{1}{2}})$. In other words, if the noise introduces enough randomness, then the channel itself ensures that all messages approximately induce the same output distribution.

In each example, statistical indistinguishability is obtained because there exists a source of randomness (key or channel noise) and a coding mechanism by which all messages induce the same distribution for the eavesdropper's observations; this mechanism is reminiscent of the codes analyzed in [7], [8], and [10] to study the notion of *channel resolvability*. At this point, the connection between secrecy and channel resolvability may seem contrived but, nevertheless, it suggests the possibility of ensuring secrecy by means that are different from those based on channel capacity and used in [1] and [2]; this was already observed in [5] and more formally explored in [6]. In the remainder of this paper, we further expand upon this idea and we not only highlight the benefits of explicitly connecting secrecy to channel resolvability but also show the limitations of an approach based on channel capacity.

B. Related Work

Most communication architectures providing information-theoretic security are based on two models of communication. The *wiretap channel*, introduced by Wyner [1] and generalized by Csiszár and Körner [2], models an architecture in which a transmitter encodes messages W into codewords $\mathbf{X}$ of n symbols for transmission to a receiver, in the presence of an eavesdropper that obtains noisy observations $\mathbf{Z}$ of $\mathbf{X}$. In the case of discrete memoryless channels, Wyner [1] and Csiszár and Körner [2] have shown the existence of coding schemes simultaneously ensuring reliable transmission to the receiver and secrecy with respect to the eavesdropper. In particular, it is possible to characterize the *secrecy capacity* of a wiretap channel, defined as the supremum of all reliable and secure rates. The extension of this result to Gaussian [11] and wireless channels (see, for instance, [12] and references therein) suggests the potential of such coding schemes to secure communication networks at the physical layer. An alternative to the wiretap channel is the source model for secret-key agreement introduced by Maurer [13] and Ahlswede and Csiszár [14], which considers an architecture in which two legitimate parties attempt to distill secret keys from a noisy source by communicating over a public channel. The resulting keys have to be secure with respect to an eavesdropper who obtains correlated observations from the source and observes all messages exchanged over the public channel. This architecture differs from the wiretap channel by exclusively focusing on the rate of the secret key that can be distilled from the source and by ignoring the cost of public communication. The counterpart of secrecy capacity is the *secret-key capacity*, defined as the supremum of the secret key rates that can be distilled. Although the aforementioned architectures model fundamentally different communication scenarios, they are related in that a coding scheme for the wiretap channel can be used to design a coding scheme for secret-key agreement and vice versa.

The early information-theoretic security results obtained for the wiretap channel and source model for secret-key agreement are criticized in some circles for measuring statistical dependence in terms of the average information rate leaked to the eavesdropper $\frac{1}{n}\mathbb{I}(W;\mathbf{Z})$. The weakness of this metric from a cryptographic standpoint has been highlighted in multiple works [4], [15], [16], which have instead advocated using the average information leaked $\mathbb{I}(W;\mathbf{Z})$. The analysis of secure communication architectures under this more stringent secrecy metric has been performed with different methods, such as graph-coloring techniques [5], privacy amplification [16], [17], and channel resolvability [6], [18]. The results presented in this paper further clarify the relation between secrecy and channel resolvability and highlight the potential of channel resolvability for solving secure communication problems.

The connection between secrecy and channel resolvability is better illustrated by studying secure communication architectures beyond the traditional memoryless setting; in particular, the distinction between the coding mechanisms for reliability and secrecy becomes apparent in the expressions of the results themselves. In this context, the information-spectrum methods pioneered by Han and Verdú turn out to be convenient mathematical tools, as they allow us to analyze general channels by focusing on the properties of mutual information as a random variable. We note that these tools have already been used to study some information-theoretic security problems and our results provide extensions of [6], [19]–[21].

C. Summary of Results

In this section, we highlight the results presented in this paper, preliminary versions of which have been reported in [22] and [23].

1) We clarify the relation between information-theoretic security and statistical independence by investigating alternatives to the average mutual information rate $\frac{1}{n}\mathbb{I}(W;\mathbf{Z})$, which is used as the *de facto* metric in most earlier works. The average mutual information rate is actually a normalized Kullback–Leibler divergence between the joint distribution $p_{W\mathbf{Z}}$ and the product distribution $p_W p_{\mathbf{Z}}$; the closeness of these two distributions can be measured by other means, such as the variational distance or even the cumulative distribution function (CDF) of the random variable $I(W;\mathbf{Z})$. By establishing relations among different metrics (see Proposition 1), we highlight the importance of choosing a measure of statistical dependence that is not only simple enough to be analytically tractable but also

strong enough to be cryptographically relevant. This discussion also provides the basis for elegant converse proofs.

2) We provide evidence that channel resolvability is a convenient mechanism for secure communication by formalizing the ideas introduced in Examples 1 and 2. Specifically, we connect secrecy to channel resolvability to analyze the fundamental limits of Shannon's cipher system (see Theorem 1) and of the broadcast channel with confidential messages (see Theorem 2). In the latter case, we show that at least for some specific wiretap channels, deriving secrecy from channel resolvability is more powerful than deriving secrecy from channel capacity (see Proposition 2); we also derive the secrecy-capacity region for general broadcast channels with a cost constraint and for strong secrecy metrics (see Theorems 2 and 3).
3) We further leverage the connection between secrecy and channel resolvability to revisit various models of secure communication. We first provide a simple proof of the strong secrecy capacity of ergodic-fading wireless channels with full channel state information (see Proposition 3). We then show that known achievable rates for mixed channels and compound channels with receiver CSI can be obtained with conceptually simple proofs, and that these results hold under stronger secrecy metrics than was previously established (see Propositions 4 and 5).
4) We finally exploit the general characterization of secrecy capacity to bound the secret-key capacity of a general discrete source model for secret-key agreement (see Proposition 7). The form of the result, which involves conditional entropy instead of mutual information, suggests that the mechanism behind secret-key agreement is not channel resolvability but rather channel intrinsic randomness [5], [24].

D. Outline

The remainder of the paper is organized as follows. Section II sets the notation used throughout the paper. Section III introduces and compares several secrecy metrics that can be used to measure information-theoretic security. Section IV analyzes the fundamental limits of secure communication for Shannon's cipher system. Section V, which forms the core of the paper, proves the impossibility of achieving strong secrecy capacity with random codes deriving secrecy from channel capacity for some wiretap channels and establishes the secrecy-capacity region of general broadcast channels with confidential messages. Section VI presents applications of the general results to wireless channels, mixed channels and compound channels, and secret-key agreement, which may be of independent interest. Section VII offers some concluding remarks. The technical details of the proofs are organized into a series of lemmas, whose proofs are relegated to the appendixes to streamline the presentation.

II. Notation

To fix notation for the sequel, consider three random variables X, Y, and Z with sample values x, y, and z taking values in alphabets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively. The joint probability distribution is denoted p_{XYZ}, and the marginal probability distributions are denoted by p_X, p_Y, and p_Z. Unless mentioned otherwise, alphabets are assumed to be abstract alphabets, including countably infinite or continuous alphabets. If the alphabets are finite, then the probability distributions correspond to probability mass functions; if the alphabets are uncountable, then the probability distributions correspond to probability densities, which we assume exist.[2] The *mutual information* between X and Y is the random variable[3]

$$I(X;Y) \triangleq \log \frac{p_{XY}(X,Y)}{p_X(X)p_Y(Y)}.$$

The average of this random variable is the usual average mutual information, which we denote by $\mathbb{I}(X;Y)$. For discrete random variables, $\mathbb{I}(X;Y)$ has the familiar expression

$$\begin{aligned}\mathbb{I}(X;Y) &\triangleq \mathbb{E}_{XY}[I(X;Y)] \\ &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{XY}(x,y)\log\frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}.\end{aligned}$$

The conditional mutual information between X and Y given Z and the average conditional mutual information are accordingly defined as

$$\begin{aligned} I(X;Y|Z) &\triangleq \log \frac{p_{XY|Z}(X,Y|Z)}{p_{X|Z}(X|Z)p_{Y|Z}(Y|Z)} \\ \text{and}\quad \mathbb{I}(X;Y|Z) &\triangleq \mathbb{E}_{XYZ}[I(X;Y|Z)],\end{aligned}$$

respectively. Similarly, the *entropy* and *average entropy* of X are

$$H(X) \triangleq \log\frac{1}{p_X(X)} \quad \text{and} \quad \mathbb{H}(X) \triangleq \mathbb{E}_X[H(X)],$$

and the conditional entropy and average conditional entropy of X given Y are

$$\begin{aligned} H(X|Y) &\triangleq \log\frac{1}{p_{X|Y}(X|Y)} \\ \text{and}\quad \mathbb{H}(X|Y) &\triangleq \mathbb{E}_{XY}[H(X|Y)].\end{aligned}$$

All the usual relations between average mutual information and average entropy that result from basic properties of joint, marginal, or conditional probability distributions can be shown to hold with probability one for the mutual information and entropy random variables. In particular, the chain rules of mutual information and entropy hold with probability 1.

The average mutual information $\mathbb{I}(X;X')$ between two random variables $X \in \mathcal{X}$ and $X' \in \mathcal{X}$ is a Kullback–Leibler divergence, which measures the closeness of the distributions $p_X p_{X'}$ and $p_{XX'}$. We will often use an alternative measure in terms of the *variational distance* between the distributions, defined as[4]

$$\mathbb{V}(p_X, p_{X'}) \triangleq 2\sup_{\mathcal{A}\subseteq\mathcal{X}} |\mathbb{P}_X[\mathcal{A}] - \mathbb{P}_{X'}[\mathcal{A}]|.$$

[2]We note that more general situations can be treated with the approach of Pinsker [25].

[3]Unless indicated otherwise, logarithms and exponentials in the paper are taken to base 2.

[4]This general definition of variational distance reduces to $\sum_{x\in\mathcal{X}} |p_X(x) - p_{X'}(x)|$ if $\mathcal{X}$ is countable.

The variational distance is not as convenient to manipulate as the average mutual information, but we provide simple rules for variational distance calculus in Appendix A.

Given two real numbers a, b, we define $[\![a, b]\!]$ as the set of integers $\{n \in \mathbb{N} : \lfloor a \rfloor \leqslant n \leqslant \lceil b \rceil\}$. To simplify notation, all vectors of length n are denoted by boldface letters; for instance, $\mathbf{x}$ denotes the vector of sample values $(x_1, \ldots, x_n)$ while $\mathbf{X}$ denotes the random vector $(X_1, \ldots, X_n)$. Given two random vectors $\mathbf{X}$ and $\mathbf{Y}$, characterized by a joint probability distribution $p_{\mathbf{XY}}$, the probability distribution of $\frac{1}{n} I(\mathbf{X}; \mathbf{Y})$ is referred to as the *mutual information rate spectrum*. In addition, the *spectral-inf mutual information rate* is defined as [10]

$$\operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}) \triangleq \sup\left\{\beta : \lim_{n\to\infty} \mathbb{P}\left[\frac{1}{n} I(\mathbf{X}; \mathbf{Y}) < \beta\right] = 0\right\},$$

and the spectral-sup mutual information rate is defined as

$$\operatorname*{p\text{-limsup}}_{n\to\infty} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}) \triangleq \inf\left\{\alpha : \lim_{n\to\infty} \mathbb{P}\left[\frac{1}{n} I(\mathbf{X}; \mathbf{Y}) > \alpha\right] = 0\right\}.$$

Operationally, the spectral-inf mutual information rate relates to channel capacity [26], whereas the spectral-sup mutual information rate relates to the channel resolvability [7]. Similarly, given an arbitrary sequence $\mathbf{X}$, the entropy rate spectrum is the distribution of the random variable $\frac{1}{n} H(\mathbf{X})$, and the spectral-inf entropy rate is defined as

$$\operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} H(\mathbf{X}) \triangleq \sup\left\{\beta : \lim_{n\to\infty} \mathbb{P}\left[\frac{1}{n} H(\mathbf{X}) < \beta\right] = 0\right\},$$

while the spectral-sup entropy rate is

$$\operatorname*{p\text{-limsup}}_{n\to\infty} \frac{1}{n} H(\mathbf{X}) \triangleq \inf\left\{\alpha : \lim_{n\to\infty} \mathbb{P}\left[\frac{1}{n} H(\mathbf{X}) > \alpha\right] = 0\right\}.$$

The spectral-sup and spectral-inf mutual information and entropy rates play a fundamental role in the analysis of reliable communication and randomness generation [7], [26], [27]. They also play a role in the analysis of secure communications, and our results combine these quantities in various ways.

III. Preliminaries: Secrecy Metrics

Let $n \in \mathbb{N}^*$ and $R > 0$. Let $W \in [\![1, 2^{nR}]\!]$ be a random variable that represents a message in a communication scheme. Assume that an eavesdropper has some knowledge about W represented by another random variable $\mathbf{Z} \in \mathcal{Z}^n$, characterized by the joint probability distribution $p_{W\mathbf{Z}}$. As mentioned in Section I, message W is information-theoretically secure if it is statistically independent of $\mathbf{Z}$; however, exact statistical independence between W and $\mathbf{Z}$ is extremely stringent and, for tractability, it is convenient to use a slightly weaker measure of secrecy, by which we only require W and $\mathbf{Z}$ to be *asymptotically* independent as the parameter n tends to infinity. Note that there is some leeway in the definition of asymptotic independence because one can choose how to measure the dependence between W and $\mathbf{Z}$. For instance, given any distance d for the space of joint probability distributions on $[\![1, 2^{nR}]\!] \times \mathcal{Z}^n$, the quantity $d(p_{W\mathbf{Z}}; p_W p_{\mathbf{Z}})$ could be used as a metric, and asymptotic statistical independence then amounts to the condition

$$\lim_{n\to\infty} d(p_{W\mathbf{Z}}; p_W p_{\mathbf{Z}}) = 0.$$

In the following, we specify six reasonable choices for secrecy metrics. The first metric measures statistical dependence using the Kullback–Leibler divergence:

$$\mathbb{S}_1(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) \triangleq \mathbb{D}(p_{W\mathbf{Z}} \| p_W p_{\mathbf{Z}}) = \mathbb{I}(W; \mathbf{Z}).$$

The secrecy condition $\lim_{n\to\infty} \mathbb{S}_1(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0$ corresponds to the well-known *strong secrecy* [15]. A second metric that we will find particularly useful is based on the variational distance:

$$\mathbb{S}_2(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) \triangleq \mathbb{V}(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}).$$

For any $\epsilon > 0$, the asymptotic independence of W and $\mathbf{Z}$ can also be measured in terms of the CDF of $I(W; \mathbf{Z})$:

$$\mathbb{S}_3(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) \triangleq \mathbb{P}[I(W; \mathbf{Z}) > \epsilon],$$

in which case the secrecy condition

$$\forall \epsilon > 0 \ \lim_{n\to\infty} \mathbb{S}_3(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0$$

means that the random variable $I(W; \mathbf{Z})$ converges in probability to zero. Finally, one could also weaken the metrics above by introducing a normalization by a factor of n as

$$\begin{aligned} &\mathbb{S}_4(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) \triangleq \tfrac{1}{n} \mathbb{D}(p_{W\mathbf{Z}} \| p_W p_{\mathbf{Z}}) \\ &\mathbb{S}_5(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) \triangleq \tfrac{1}{n} \mathbb{V}(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}), \\ \text{for } \epsilon > 0 \quad &\mathbb{S}_6(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) \triangleq \mathbb{P}\left[\tfrac{1}{n} I(W; \mathbf{Z}) > \epsilon\right]. \end{aligned}$$

The secrecy condition $\lim_{n\to\infty} \mathbb{S}_4(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0$ corresponds to the *weak secrecy* initially introduced by Wyner [1]. The secrecy conditions[5] $\lim_{n\to\infty} \mathbb{S}_i(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0$ may not be equivalent for all $i \in [\![1, 6]\!]$; by establishing an ordering among these metrics, we formalize what it means for a metric to be "stronger" than another. For $i, j \in [\![1, 6]\!]$, we say that $\mathbb{S}_i$ is *stronger* than $\mathbb{S}_j$ (or equivalently that $\mathbb{S}_j$ is *weaker* than $\mathbb{S}_i$), and we write $\mathbb{S}_i \succeq \mathbb{S}_j$ if and only if

$$\lim_{n\to\infty} \mathbb{S}_i(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0 \Rightarrow \lim_{n\to\infty} \mathbb{S}_j(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0.$$

By construction, it is clear that $\mathbb{S}_1 \succeq \mathbb{S}_4$, $\mathbb{S}_2 \succeq \mathbb{S}_5$ and $\mathbb{S}_3 \succeq \mathbb{S}_6$; however, we establish a more precise result.

Proposition 1: The secrecy metrics $\mathbb{S}_i$ for $i \in [\![1, 6]\!]$ are ordered as follows:

$$\mathbb{S}_1 \succeq \mathbb{S}_2 \succeq \mathbb{S}_3 \succeq \mathbb{S}_4 \succeq \mathbb{S}_5 \succeq \mathbb{S}_6.$$

Proof: The relations $\mathbb{S}_1 \succeq \mathbb{S}_2$ and $\mathbb{S}_4 \succeq \mathbb{S}_5$ directly follow from Pinsker's inequality [25, Corollary p. 16]. Similarly, the

[5]The limit should be understood for any $\epsilon > 0$ in the case of metrics $\mathbb{S}_3$ and $\mathbb{S}_6$.

relations $\mathbb{S}_2 \succeq \mathbb{S}_3$ and $\mathbb{S}_5 \succeq \mathbb{S}_6$ follow from [25, Corollary p. 18]; hence, we only need to prove that $\mathbb{S}_3 \succeq \mathbb{S}_4$.

Let $\epsilon, \gamma > 0$. Assume that $\lim_{n\to\infty} \mathbb{S}_3(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0$, so that $\lim_{n\to\infty} \mathbb{P}[I(W;\mathbf{Z}) > \epsilon] = 0$. Note that metric $\mathbb{S}_4(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}})$ can be written as

$$\begin{aligned}
&\mathbb{S}_4(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) \\
&\quad = \tfrac{1}{n}\mathbb{I}(W;\mathbf{Z}) \\
&\quad = \mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z})\right], \\
&\quad = \mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{I(W;\mathbf{Z}) \leqslant -\epsilon\}\right] \\
&\qquad + \mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{-\epsilon < I(W;\mathbf{Z}) \leqslant \epsilon\}\right] \\
&\qquad + \mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{\epsilon < I(W;\mathbf{Z}) \leqslant n(R+\gamma)\}\right] \\
&\qquad + \mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{I(W;\mathbf{Z}) > n(R+\gamma)\}\right].
\end{aligned}$$

Clearly, we have $\mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{I(W;\mathbf{Z}) \leqslant -\epsilon\}\right] < 0$, $\mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{-\epsilon < I(W;\mathbf{Z}) \leqslant \epsilon\}\right] \leqslant \tfrac{\epsilon}{n}$, and

$$\begin{aligned}
&\mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{\epsilon < I(W;\mathbf{Z}) \leqslant n(R+\gamma)\}\right] \\
&\qquad\qquad \leqslant (R+\gamma)\, \mathbb{P}[I(W;\mathbf{Z}) > \epsilon].
\end{aligned}$$

Following [10, p. 223], we can also prove that

$$\lim_{n\to\infty} \mathbb{E}\left[\tfrac{1}{n} I(W;\mathbf{Z}) \mathbb{1}\{I(W;\mathbf{Z}) > n(R+\gamma)\}\right] = 0.$$

Therefore, $\lim_{n\to\infty} \mathbb{S}_4(p_{W\mathbf{Z}}, p_W p_{\mathbf{Z}}) = 0$ and $\mathbb{S}_3 \succeq \mathbb{S}_4$. ■

A direct consequence of Proposition 1 is that any secure communication scheme satisfying the secrecy condition with the strongest secrecy metric $\mathbb{S}_1$ automatically satisfies it with the secrecy metrics $\mathbb{S}_i$ for $i \in [\![2,6]\!]$. Conversely, any secure communication scheme that does not satisfy the secrecy condition with the weakest metric $\mathbb{S}_6$ cannot satisfy it with any of the metrics $\mathbb{S}_i$ for $i \in [\![1,5]\!]$. Therefore, to establish a coding theorem for a secure communication scheme, we can prove achievability with metric $\mathbb{S}_1$ and a converse with metric $\mathbb{S}_6$.

Although the ordering in Proposition 1 follows strictly from mathematical properties, the idea that some metrics are stronger than others is also meaningful from a cryptographic perspective. One can construct examples of communication schemes that present obvious security loopholes while still satisfying a secrecy condition with metric $\mathbb{S}_4$ (see, for instance, the examples in [4], [28], and [29]). It is now accepted that information-theoretic secrecy conditions[6] should hold at least with metrics $\mathbb{S}_1$ or $\mathbb{S}_2$.

IV. Secrecy From Channel Resolvability for Shannon's Cipher System

As a first illustration of the connection between secrecy and channel resolvability, we elaborate on Example 1 and revisit Shannon's cipher system. We consider the model illustrated in Fig. 2, in which a message W uniformly distributed in $[\![1, 2^{nR}]\!]$ is to be communicated reliably from a transmitter (Alice) to a legitimate receiver (Bob) in the presence of an eavesdropper (Eve). Alice and Bob have access to a common discrete source of randomness $(\mathcal{K}, \{p_{\mathbf{K}}\}_{n\geqslant 1})$, characterized by an alphabet $\mathcal{K}$ and a sequence of symbol probabilities $\{p_{\mathbf{K}}\}_{n\geqslant 1}$, which is used to encode W into a codeword $Z \in \mathcal{Z}$. Bob's estimate of the message using Z and the source of randomness $\mathbf{K}$ is denoted by $\hat{W}$.

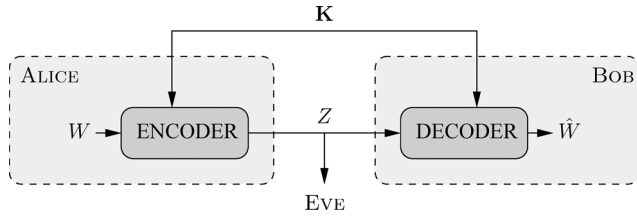

Fig. 2. Shannon's cipher system for a general common source of randomness.

Definition 1: A $(2^{nR}, n)$ cipher $\mathcal{E}_n$ consists of
1) an encoding function $f_n : [\![1, 2^{nR}]\!] \times \mathcal{K}^n \to \mathcal{Z}$;
2) a decoding function $g_n : \mathcal{Z} \times \mathcal{K}^n \to [\![1, 2^{nR}]\!]$.

The reliability performance of a cipher $\mathcal{E}_n$ is measured in terms of the probability of error $\mathbb{P}_e(\mathcal{E}_n) \triangleq \mathbb{P}\left[\hat{W} \neq W | \mathcal{E}_n\right]$ while its secrecy performance is measured in terms of the secrecy metric[7] $\mathbb{S}_i(\mathcal{E}_n) \triangleq \mathbb{S}_i(p_{WZ|\mathcal{E}_n}, p_W p_{Z|\mathcal{E}_n})$.

Definition 2: A rate R is achievable for secrecy metric $\mathbb{S}_i$ for Shannon's cipher system if there exists a sequence of $\left(2^{nR}, n\right)$ ciphers $\{\mathcal{E}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \mathbb{P}_e(\mathcal{E}_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{S}_i(\mathcal{E}_n) = 0.$$

The secrecy capacity $C_{SC}^{(i)}$ of Shannon's cipher system is the supremum of achievable rates for secrecy metric $\mathbb{S}_i$.

Theorem 1: The secrecy capacity of Shannon's cipher system is the same for all metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$ and is given by

$$C_{SC} = \operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} H(\mathbf{K}). \tag{1}$$

If the source $(\mathcal{K}, \{p_{\mathbf{K}}\}_{n\geqslant 1})$ is memoryless, then the secrecy capacity is also the same for metric $\mathbb{S}_1$.

Proof: We first show that all rates below $\operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} H(\mathbf{K})$ are achievable for secrecy metric $\mathbb{S}_2$. Let $\epsilon, \gamma > 0$ and $R \triangleq \operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} H(\mathbf{K}) - \gamma$. Let U_R be the random variable with uniform distribution on $[\![1, 2^{nR}]\!]$. By [27, Lemma 3], there exists an encoding function $f_n : \mathcal{K}^n \to [\![1, 2^{nR}]\!]$ such that $\mathbb{V}\left(p_{f_n(\mathbf{K})}, p_{U_R}\right) \leqslant \epsilon_n$ with $\lim_{n\to\infty} \epsilon_n = 0$. A message W is then encoded as $Z = f_n(\mathbf{K}) \oplus W$, where $\oplus$ represents the addition modulo $\lceil 2^{nR} \rceil$. By construction, Bob retrieves W without error since $W = Z \oplus f_n(\mathbf{K})$. We have

$$\begin{aligned}
\mathbb{S}_2(\mathcal{E}_n) &= \mathbb{V}(p_{WZ}, p_W p_Z) \\
&= \mathbb{E}_W\left[\mathbb{V}\left(p_{Z|W}, p_Z\right)\right] \\
&\leqslant \mathbb{E}_W\left[\mathbb{V}\left(p_{Z|W}, p_{U_R}\right)\right] + \mathbb{V}(p_{U_R}, p_Z) \\
&\leqslant 2\mathbb{E}_W\left[\mathbb{V}\left(p_{Z|W}, p_{U_R}\right)\right] \\
&= 2\mathbb{E}_W\left[\mathbb{V}\left(p_{f_n(\mathbf{K})}, p_{U_R}\right)\right] \\
&\leqslant 2\epsilon_n,
\end{aligned}$$

[6]The conditions could be further strengthened by imposing an exponential convergence with n; however, except in the case of exponentially information stable channels [5], such as memoryless channels, we were unable to prove general results with this additional constraint.

[7]We will drop the conditioning on $\mathcal{E}_n$ when this is clear from the context.

where we have used Lemma 7, the definition of Z, and the independence of $f_n(\mathbf{K})$ and W. Therefore, the rate R is achievable and, since γ can be chosen arbitrarily small, we conclude that

$$C_{SC}^{(2)} \geqslant \text{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{K}). \tag{2}$$

If the source $(\mathcal{K}, p_K)$ is i.i.d., one can modify the proof of [27, Lemma 3] to show that, if $R \triangleq \mathbb{H}(K) - \gamma$, there exists a function $f_n : \mathcal{K}^n \to [\![1, 2^{nR}]\!]$ and $\alpha_\gamma > 0$, such that $\mathbb{V}\big(p_{f_n(\mathbf{K})}, p_{U_R}\big) \leqslant 2^{-\alpha_\gamma n}$. Following the same steps as above, we then obtain that $\mathbb{S}_2(\mathcal{E}_n) \leqslant 2 \cdot 2^{-\alpha_\gamma n}$. Finally, [5, Lemma 1] shows that there exists $\beta_\gamma > 0$ such that, for n large enough $\mathbb{S}_1(\mathcal{E}_n) \leqslant 2^{-\beta_\gamma n}$.

We now prove the converse part of the result. Let R be an achievable rate for secrecy metric $\mathbb{S}_6$. There exists a sequence of $(2^{nR}, n)$ ciphers $\{\mathcal{E}_n\}_{n\geqslant 1}$ such that $\lim_{n\to\infty} \mathbb{P}_e(\mathcal{E}_n) = 0$ and $\lim_{n\to\infty} \mathbb{S}_6(\mathcal{E}_n) = 0$. For every $n \in \mathbb{N}^*$, and with probability one, we have

$$\begin{aligned}
\tfrac{1}{n} H(W) &= \tfrac{1}{n} H(W|Z) + \tfrac{1}{n} I(W; Z) \\
&= \tfrac{1}{n} I(W; \mathbf{K}|Z) + \tfrac{1}{n} H(W|Z\mathbf{K}) + \tfrac{1}{n} I(W; Z) \\
&= \tfrac{1}{n} H(\mathbf{K}) - \tfrac{1}{n} H(\mathbf{K}|WZ) - \tfrac{1}{n} I(\mathbf{K}; Z) \\
&\qquad + \tfrac{1}{n} H(W|Z\mathbf{K}) + \tfrac{1}{n} I(W; Z).
\end{aligned}$$

Since $R = \text{p-liminf}_{n\to\infty} \frac{1}{n} H(W)$, $\text{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{K}|WZ) \geqslant 0$, and $\text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{K}; Z) \geqslant 0$, we obtain

$$\begin{aligned}
R \leqslant \text{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{K}) + \text{p-limsup}_{n\to\infty} \frac{1}{n} H(W|Z\mathbf{K}) \\
+ \text{p-limsup}_{n\to\infty} \frac{1}{n} I(W; Z).
\end{aligned}$$

Note that $\text{p-limsup}_{n\to\infty} \frac{1}{n} I(W; Z) = 0$ by assumption. The Verdú–Han Lemma [10], [26] also guarantees that $\text{p-limsup}_{n\to\infty} \frac{1}{n} H(W|Z\mathbf{K}) = 0$; hence, we have

$$C_{SC}^{(6)} \leqslant \text{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{K}). \tag{3}$$

Combining (2) and (3) with Proposition 1, we conclude that, for every $i \in [\![2, 6]\!]$, $C_{SC}^{(i)} = \text{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{K})$. If the source is memoryless, then for every $i \in [\![1, 6]\!]$, $C_{SC}^{(i)} = \mathbb{H}(K)$. ■

The coding scheme used in Theorem 1 extracts the source intrinsic randomness of $(\mathcal{K}, \{p_{\mathbf{K}}\}_{n\geqslant 1})$ to protect the message with a one-time pad. Nevertheless, the message is kept secret from the eavesdropper because the encoder exploits the randomness of the source to control the distribution of the eavesdropper's observation; hence, the coding mechanism for secure communication can be interpreted as channel resolvability, which we confirm in the next section. From a cryptographic perspective, Theorem 1 shows that the secure communication rate is maximized if the legitimate terminals make sure that their keys are almost perfectly uniform. This has operational significance in a practical situation if the mechanism providing secret keys is biased and does not yield perfectly uniform keys. Finally, the fact that $C_{SC}^{(i)}$ remains identical for all metrics $\mathbb{S}_i$ with $i \in [\![2, 6]\!]$ suggests that asymptotic statistical independence is indeed a fundamental measure of secrecy.

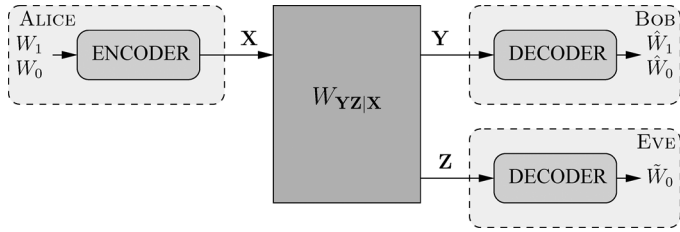

Fig. 3. Broadcast channel with confidential messages.

V. Secrecy From Channel Resolvability Over Noisy Channels

We now turn our attention to the problem of secure communication over noisy channels. We consider a broadcast channel with confidential messages $\big(\mathcal{X}, \mathcal{Y}, \{W_{\mathbf{YZ}|\mathbf{X}}\}_{n\geqslant 1}, \mathcal{Z}\big)$ characterized by an input alphabet $\mathcal{X}$, two output alphabets $\mathcal{Y}$ and $\mathcal{Z}$, and a sequence of transition probabilities $\{W_{\mathbf{YZ}|\mathbf{X}}\}_{n\geqslant 1}$. The channels $\big(\mathcal{X}, \{W_{\mathbf{Y}|\mathbf{X}}\}_{n\geqslant 1}, \mathcal{Y}\big)$ and $\big(\mathcal{X}, \{W_{\mathbf{Z}|\mathbf{X}}\}_{n\geqslant 1}, \mathcal{Z}\big)$ obtained from the marginals are called the *main channel* and the *eavesdropper's channel*, respectively. The inputs to the channels are also subject to cost constraint $P \in \mathbb{R}^+$; specifically, there exists a sequence of cost functions $\{c_n\}_{n\geqslant 1}$ with $c_n : \mathcal{X}^n \to \mathbb{R}_+$, such that any sequence $\mathbf{x} \in \mathcal{X}^n$ transmitted through the channel should satisfy $\frac{1}{n} c_n(\mathbf{x}) \leqslant P$. Following standard practice, the transmitter is named Alice, the receiver observing output $\mathbf{Y}$ is named Bob, and the receiver observing output $\mathbf{Z}$ is named Eve. As illustrated in Fig. 3, Alice wishes to transmit a common message W_0 to both Bob and Eve and an individual message W_1 for Bob alone, viewing Eve as an eavesdropper for message W_1. Bob's estimates of the messages are denoted by $\hat{W}_0$ and $\hat{W}_1$ while Eve's estimate is denoted by $\tilde{W}_0$.

Definition 3: A $\big(2^{nR_0}, 2^{nR_1}, n\big)$ wiretap code $\mathcal{C}_n$ consists of

1) a common message set $\mathcal{W}_0 = [\![1, 2^{nR_0}]\!]$;
2) an individual message set $\mathcal{W}_1 = [\![1, 2^{nR_1}]\!]$;
3) an auxiliary message set $\mathcal{W}' = [\![1, 2^{nR'_n}]\!]$, with $R'_n > 0$,[8] used to randomize the encoding of messages;
4) a source of local randomness $(\mathcal{R}, p_R)$, which is only known to Alice and can be used to further randomize the encoding;
5) an encoding function $f_n : \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}' \times \mathcal{R} \to \mathcal{X}^n$, such that $\frac{1}{n} c_n(f_n(m_0, m_1, m', r)) \leqslant P$;
6) a decoding function $g_n : \mathcal{Y}^n \to \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}'$;
7) a decoding function $h_n : \mathcal{Z}^n \to \mathcal{W}_0$.

The auxiliary message is denoted by W'. All messages W_0, W_1, W' are assumed to be uniformly distributed in their respective sets. The size of the auxiliary message set and the source of local randomness $(\mathcal{R}, p_R)$ can be optimized as part of the code design, and the eavesdropper is assumed to know the code $\mathcal{C}_n$, which includes the statistics p_R of the source of local randomness. In the remainder of the paper, we clearly identify the channel inputs and outputs obtained when using a code $\mathcal{C}_n$ by introducing a bar in the notation of the corresponding random variables. For instance, the random variable representing a codeword chosen in $\mathcal{C}_n$ is denoted $\bar{\mathbf{X}}$, those

[8] Although R_0 and R_1 are fixed parameters, we allow R'_n to vary with n.

representing the corresponding channel outputs are denoted $\bar{\mathbf{Y}}$ and $\bar{\mathbf{Z}}$. The joint distribution between $W_0, W_1, \bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\mathbf{Z}}$ is

$$p_{W_0W_1\bar{\mathbf{X}}\bar{\mathbf{Y}}\bar{\mathbf{Z}}}(m_0,m_1,\mathbf{x},\mathbf{y},\mathbf{z}) = W_{\mathbf{YZ}|\mathbf{X}}(\mathbf{y},\mathbf{z}|\mathbf{x}) \, p_{\bar{\mathbf{X}}|W_0W_1}(\mathbf{x}|m_0,m_1) p_{W_0}(m_0) p_{W_1}(m_1). \quad (4)$$

The reliability of a code $\mathcal{C}_n$ is measured in terms of the average probability of error

$$\mathbb{P}_e(\mathcal{C}_n) \triangleq \mathbb{P}\left[(\hat{W}_0,\hat{W}_1,\hat{W}') \neq (W_0,W_1,W') \text{ or } \tilde{W}_0 \neq W_0 \middle| \mathcal{C}_n \right]$$

while its secrecy is measured in terms of the secrecy metric $\mathbb{S}_i(\mathcal{C}_n) \triangleq \mathbb{S}_i(p_{W_1\bar{\mathbf{Z}}|\mathcal{C}_n}, p_{W_1}p_{\bar{\mathbf{Z}}|\mathcal{C}_n})$ for $i \in [\![1,6]\!]$.

Definition 4: A rate pair (R_0,R_1) is achievable for secrecy metric $\mathbb{S}_i$ over a broadcast channel if there exists a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \mathbb{P}_e(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{S}_i(\mathcal{C}_n) = 0.$$

The secrecy-capacity region $\mathcal{R}_{BCC}^{(i)}$ is the closure of the set of rate pairs achievable for secrecy metric $\mathbb{S}_i$, and the secrecy capacity for secrecy metric $\mathbb{S}_i$ is

$$C_{WT}^{(i)} \triangleq \sup\{R_1 : (0,R_1) \text{ is achievable for secrecy metric } \mathbb{S}_i\}.$$

In the absence of a common message ($R_0 = 0$), a broadcast channel with confidential messages is concisely called a wiretap channel, and a $(1, 2^{nR_1}, n)$ code is simply denoted as a $(2^{nR_1}, n)$ code. Note that our definition of a wiretap code explicitly introduces the randomness used in the encoding process. The randomness is split between a source of local randomness and an auxiliary message with uniform distribution that we require the legitimate receiver to decode. This allows us to distinguish the part of the randomness that merely acts as artificial random noise from the part that helps secrecy without reducing the reliable communication rate. Since the source of local randomness can be arbitrarily chosen, our definition incurs no loss of generality and allows us to explicitly define the class of *capacity-based wiretap codes* in Section V-A.

Remark 1: Csiszár and Körner [2] analyze the fundamental limits of secure communication more precisely by studying the rate-equivocation region (R_0,R_1,R_e), where $R_e \leqslant R_1$ represents the equivocation-rate $\frac{1}{n}\mathbb{H}(W_1|\bar{\mathbf{Z}})$ of the eavesdropper about the individual message. Unlike the rates R_0 and R_1, the notion of equivocation depends on the secrecy metric considered; therefore, we restrict ourselves to the special case of full secrecy rates $R_1 = R_e$, for which we can leverage the result of Proposition 1.

A. Capacity-Based Wiretap Codes and Strong Secrecy

We now define the subclass of *capacity-based* wiretap codes.

Definition 5: A $\left(2^{nR_0}, 2^{nR_1}, n\right)$ capacity-based wiretap code $\mathcal{C}_n$ is a $\left(2^{nR_0}, 2^{nR_1}, n\right)$ wiretap code such that:

1) the auxiliary message rate is $R'_n = C_e - \epsilon_n$, where C_e is the eavesdropper's channel capacity and $\{\epsilon_n\}_{n\geqslant 1}$ satisfies $\lim_{n\to\infty} \epsilon_n = 0$;
2) there exists a decoding function $h'_n : \mathcal{Z}^n \times \mathcal{W}_1 \to \mathcal{W}'$, which allows the eavesdropper to estimate the auxiliary message W' from the observation of $\bar{\mathbf{Z}}$ and W_1.

We let $\tilde{W}'$ denote Eve's estimate of W'. The reliability of a capacity-based wiretap code $\mathcal{C}_n$ is then measured in terms of the modified average probability of error

$$\mathbb{P}_e^*(\mathcal{C}_n) \triangleq \mathbb{P}\left[(\hat{W}_0,\hat{W}_1,\hat{W}') \neq (W_0,W_1,W') \text{ or } (\tilde{W}_0,\tilde{W}') \neq (W_0,W') \middle| \mathcal{C}_n \right].$$

Definition 6: A rate pair (R_0,R_1) is achievable for secrecy metric $\mathbb{S}_i$ with capacity-based wiretap codes if there exists a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ capacity-based wiretap codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \mathbb{P}_e^*(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{S}_i(\mathcal{C}_n) = 0.$$

The constraint $\lim_{n\to\infty} \mathbb{P}_e^*(\mathcal{C}_n) = 0$ ensures that, given knowledge of $\bar{\mathbf{Z}}$ and W_1, the eavesdropper could reliably decode the auxiliary message W'. Nevertheless, since the eavesdropper does not have access to the message W_1, this property is solely used to impose structure on the code. However, note that this also imposes $\lim_{n\to\infty} \epsilon_n\sqrt{n} = \infty$ [30, Th. 49]. The denomination "capacity-based code" is used because the set of codewords associated to a known pair of messages (W_0,W_1) forms a subcode of rate $R'_n = C_e - \epsilon_n$, which stems from a sequence of capacity-achieving codes for Eve's channel.

As formalized in [31, Th. 1], capacity-based wiretap codes are implicitly used in most works that show the existence of wiretap codes achieving secrecy rates for metric $\mathbb{S}_4$. In this section, we show that this may be an intrinsic limitation, by proving that sequences of random capacity-based wiretap codes that achieve the weak secrecy capacity *cannot* achieve the strong secrecy capacity.

Specifically, we consider a discrete memoryless wiretap channel $\left(\mathcal{X},\mathcal{Y},W_{YZ|X},\mathcal{Z}\right)$ without cost constraint ($\forall \mathbf{x} \in \mathcal{X}^n \;\; c_n(\mathbf{x}) = n$ and $P = 1$) in which the eavesdropper's channel and the main channel are both symmetric.[9] We further assume that the main channel is more capable than the eavesdropper's channel and has capacity $C_m < \frac{1}{2}\log|\mathcal{X}|$ bits. The former assumption ensures that, without loss of optimality, we can assume no source of local randomness $(\mathcal{R}, p_R)$ is available [2] and that the secrecy capacity is $C_s = C_m - C_e$; the latter one is a technical assumption required to simplify the analysis.

Proposition 2: Let $\{C_n\}_{n\geqslant 1}$ be a sequence of $(2^{nR}, n)$ random capacity-based wiretap codes, obtained by generating codeword symbols independently and uniformly at random. Let the rate R'_n of the auxiliary message be such that $R'_n = C_e - \epsilon_n$ and $R + R'_n = C_m - \epsilon_n$. Then, there exists $\eta, \alpha > 0$, such that, for n sufficiently large,

$$\mathbb{P}[\mathbb{S}_2(C_n) > \eta, \;\; \mathbb{P}_e^*(C_n) \leqslant \epsilon'_n \;\text{ and }\; \mathbb{S}_4(C_n) \leqslant 3\epsilon'_n] \geqslant 1 - 2^{-\alpha n\epsilon_n^2},$$

with $\epsilon'_n \triangleq \max(\epsilon_n, \log|\mathcal{X}|2^{-\alpha n\epsilon_n^2}, n^{-1})$, i.e., with high probability over the random code ensemble, a sequence of capacity-

[9] More specifically, we use Gallager's notion of symmetry [32, p. 94].

based random codes achieves the weak secrecy capacity but does not achieve the strong secrecy capacity.

Proof: See Appendix B ∎

We conjecture that the inability to achieve strong secrecy holds for any capacity-based wiretap codes, and not just random codes, as well as for any discrete memoryless channel, and not just symmetric channels. Despite its lack of generality, Proposition 2 shows that a random coding argument with capacity-based wiretap codes is not powerful enough to prove strong secrecy results, which suggests exploiting a more powerful mechanism to ensure secrecy. In the remainder of the paper, we derive secrecy from channel resolvability and show that the resulting codes do not suffer from the limitations of capacity-based wiretap codes.

Remark 2: If the main channel is noiseless and the eavesdropper's channel is symmetric, a slight modification of the proof of Proposition 2 shows that no capacity-based wiretap code (including nonrandom codes) achieves secrecy capacity for metrics $\mathbb{S}_2$ and $\mathbb{S}_1$. This fact was independently noted in [33] for metric $\mathbb{S}_1$ using results for finite blocklength channel coding [30]. Our approach builds on a similar result established for secret-key agreement in [34].

B. General Broadcast Channels With Confidential Messages and Cost Constraint

In this section, we establish the secrecy-capacity region of a general broadcast channel with confidential messages for secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$; the alphabets and transition probabilities of the channel $\{W_{\mathbf{YZ}|\mathbf{X}}\}_{n\geqslant 1}$ are arbitrary, so that the model includes continuous channels and channels with memory. Following the conclusions drawn from Proposition 2, we analyze codes that are more powerful than capacity-based wiretap codes and whose secrecy is tied to the notion of channel resolvability.

Theorem 2: The secrecy-capacity region of a broadcast channel $\left(\mathcal{X}, \mathcal{Y}, \{W_{\mathbf{YZ}|\mathbf{X}}\}_{n\geqslant 1}, \mathcal{Z}\right)$ with confidential messages and cost constraint P is the same for all secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$ and is given by

$$\mathcal{R}_{BCC} = \bigcup_{\{\mathbf{UVX}\}_{n\geqslant 1}\in\mathcal{P}} \left\{ \begin{array}{l} (R_0, R_1) \in \mathbb{R}_+^2 : \\ R_0 \leqslant \min\left(\underset{n\to\infty}{\text{p-liminf}}\, \frac{1}{n} I(\mathbf{U};\mathbf{Y}), \right. \\ \qquad\qquad \left. \underset{n\to\infty}{\text{p-liminf}}\, \frac{1}{n} I(\mathbf{U};\mathbf{Z}) \right), \\ R_1 \leqslant \underset{n\to\infty}{\text{p-liminf}}\, \frac{1}{n} I(\mathbf{V};\mathbf{Y}|\mathbf{U}) \\ \qquad\qquad - \underset{n\to\infty}{\text{p-limsup}}\, \frac{1}{n} I(\mathbf{V};\mathbf{Z}|\mathbf{U}) \end{array} \right\} \quad (5)$$

where

$$\mathcal{P} \triangleq \left\{ \{\mathbf{UVX}\}_{n\geqslant 1} : \forall n \in \mathbb{N}^*\ \mathbf{U} \to \mathbf{V} \to \mathbf{X} \to \mathbf{YZ} \text{ forms a Markov chain and } \mathbb{P}\left[\tfrac{1}{n} c_n(\mathbf{X}) \leqslant P\right] = 1 \right\}.$$

Notice that the form of the secrecy-capacity region is the natural generalization of that obtained for memoryless channels in [2, Corollary 1]; however, the main channel statistics affect the secure rate R_1 through their "worst realization" $\underset{n\to\infty}{\text{p-liminf}}\, \frac{1}{n} I(\mathbf{V};\mathbf{Y}|\mathbf{U})$ while the eavesdropper's channel statistics affect it through their "best realization" $\underset{n\to\infty}{\text{p-limsup}}\, \frac{1}{n} I(\mathbf{V};\mathbf{Z}|\mathbf{U})$. Intuitively, as illustrated in Fig. 4, this occurs because the worst case for secure communication is when the main channel conveys the smallest information rate to the legitimate receiver while the eavesdropper's channel leaks the largest information rate to the eavesdropper. It will be apparent in the proof that this asymmetry, which disappears in the case of memoryless channels, arises because the coding mechanisms used to ensure reliability and secrecy are different.

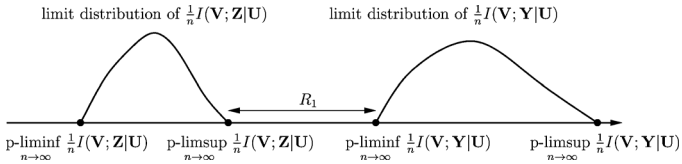

Fig. 4. Illustration of secure rates in Theorem 2.

Proof of Theorem 2: We start with the achievability part of the proof, for which we create a codebook by combining superposition coding and binning schemes. Let $n \in \mathbb{N}^*$ and $\epsilon, \gamma, R_0, R_1, R' > 0$. Define $M_0 \triangleq \lceil 2^{nR_0} \rceil$, $M_1 \triangleq \lceil 2^{nR_1} \rceil$ and $M' \triangleq \lceil 2^{nR'} \rceil$. Let $\mathcal{U}$ be an arbitrary alphabet and fix a distribution $p_{\mathbf{U}}$ on $\mathcal{U}^n$. Fix a conditional distribution $p_{\mathbf{X}|\mathbf{U}}$ on $\mathcal{X}^n \times \mathcal{U}^n$ such that $\mathbb{P}\left[\frac{1}{n} c_n(\mathbf{X}) \leqslant P\right] = 1$. Let $\mathbf{U}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ be the random variables with joint distribution

$$p_{\mathbf{UXYZ}}(\mathbf{u},\mathbf{x},\mathbf{y},\mathbf{z}) \triangleq W_{\mathbf{YZ}|\mathbf{X}}(\mathbf{y},\mathbf{z}|\mathbf{x}) p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}|\mathbf{u}) p_{\mathbf{U}}(\mathbf{u}). \quad (6)$$

1) *Code generation:* Randomly generate M_0 sequences $\mathbf{u}_k \in \mathcal{U}^n$ with $k \in [\![1, M_0]\!]$ according to $p_{\mathbf{U}}$. For each $k \in [\![1, M_0]\!]$, generate $M_1 M'$ sequences $\mathbf{x}_{klm} \in \mathcal{X}^n$ with $(l, m) \in [\![1, M_1]\!] \times [\![1, M']\!]$ according to $p_{\mathbf{X}|\mathbf{U}=\mathbf{u}_k}$. We denote by C_n the random variable representing the generated code and by $\mathcal{C}_n$ one of its realizations.
2) *Encoding:* To transmit a message pair $(k, l) \in [\![1, M_0]\!] \times [\![1, M_1]\!]$, Alice generates an auxiliary message m uniformly at random in $[\![1, M']\!]$ and sends the codeword $\mathbf{x}_{klm}$ through the channel.
3) *Bob's decoding:* Define the sets

$$\mathcal{T}_1^n \triangleq \left\{ (\mathbf{u},\mathbf{y}) \in \mathcal{U}^n \times \mathcal{Y}^n : \frac{1}{n} \log \frac{p_{\mathbf{Y}|\mathbf{U}}(\mathbf{y}|\mathbf{u})}{p_{\mathbf{Y}}(\mathbf{y})} \geqslant \frac{1}{n} \log M_0 + \gamma \right\},$$

$$\mathcal{T}_2^n \triangleq \left\{ (\mathbf{u},\mathbf{x},\mathbf{y}) \in \mathcal{U}^n \times \mathcal{X}^n \times \mathcal{Y}^n : \frac{1}{n} \log \frac{p_{\mathbf{Y}|\mathbf{XU}}(\mathbf{y}|\mathbf{x},\mathbf{u})}{p_{\mathbf{Y}|\mathbf{U}}(\mathbf{y}|\mathbf{u})} \geqslant \frac{1}{n} \log M_1 M' + \gamma \right\}.$$

Upon observing $\mathbf{y}$, Bob decodes k as the received common message if $\mathbf{u}_k$ is the unique sequence in $\mathcal{C}_n$ such that $(\mathbf{u}_k, \mathbf{y}) \in \mathcal{T}_1^n$; otherwise, a random message is chosen. Similarly, he decodes l as the received individual message and m as the received auxiliary message if there exists a unique codeword $\mathbf{x}_{klm}$ such that $(\mathbf{u}_k, \mathbf{x}_{klm}, \mathbf{y}) \in \mathcal{T}_2^n$; otherwise, random messages are chosen.

4) *Eve's decoding:* Define the set

$$\mathcal{T}_3^n \triangleq \{(\mathbf{u},\mathbf{z}) \in \mathcal{U}^n \times \mathcal{Z}^n : \frac{1}{n}\log\frac{p_{\mathbf{Z}|\mathbf{U}}(\mathbf{z}|\mathbf{u})}{p_{\mathbf{Z}}(\mathbf{z})} \geqslant \frac{1}{n}\log M_0 + \gamma\Big\}.$$

Upon observing $\mathbf{z}$, Eve decodes k as the received common message if $\mathbf{u}_k$ is the unique sequence such that $(\mathbf{u}_k,\mathbf{z}) \in \mathcal{T}_3^n$; otherwise, a random message is chosen.

The following lemmas, whose proofs are relegated to Appendix C, provide sufficient conditions to guarantee reliability and secrecy.

Lemma 1 (Reliability Conditions):

$$\text{If } R_0 \leqslant \min\Big(\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{U};\mathbf{Y}) - 2\gamma, \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{U};\mathbf{Z}) - 2\gamma\Big)$$
$$\text{and } R_1 + R' \leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{X};\mathbf{Y}|\mathbf{U}) - 2\gamma,$$

then $\lim_{n\to\infty}\mathbb{E}[\mathbb{P}_e(C_n)] \leqslant \epsilon$.

Lemma 2 (Secrecy From Channel Resolvability Condition):

$$\text{If } R' \geqslant \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U}) + 2\gamma \quad \text{then } \lim_{n\to\infty}\mathbb{E}[\mathbb{S}_2(C_n)] \leqslant \epsilon.$$

Combining Lemmas 1 and 2, we obtain that if

$$R_0 \leqslant \min\Big(\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{U};\mathbf{Y}) - 2\gamma, \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{U};\mathbf{Z}) - 2\gamma\Big)$$
$$\text{and } R_1 \leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{X};\mathbf{Y}|\mathbf{U}) - \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U}) - 4\gamma,$$

then, $\lim_{n\to\infty}\mathbb{E}[\mathbb{P}_e(C_n)] \leqslant \epsilon$ and $\lim_{n\to\infty}\mathbb{E}[\mathbb{S}_2(C_n)] \leqslant \epsilon$. By Markov's inequality and the union bound, there exists at least one sequence of $(2^{nR_0}, 2^{nR_1}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that $\lim_{n\to\infty}\mathbb{P}_e(\mathcal{C}_n) \leqslant 3\epsilon$ and $\lim_{n\to\infty}\mathbb{S}_2(\mathcal{C}_n) \leqslant 3\epsilon$. Since ϵ and γ can be chosen arbitrarily small, we conclude that

$$\bigcup_{\{\mathbf{UX}\}_{n\geqslant 1}\in\mathcal{P}} \left\{ \begin{array}{l} (R_0,R_1) \in \mathbb{R}_+^2 : \\ R_0 \leqslant \min\Big(\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{U};\mathbf{Y}), \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{U};\mathbf{Z})\Big), \\ R_1 \leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{X};\mathbf{Y}|\mathbf{U}) - \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U}) \end{array} \right\} \subseteq \mathcal{R}_{BCC}^{(2)} \tag{7}$$

where

$$\mathcal{P} \triangleq \{\{\mathbf{UX}\}_{n\geqslant 1} : \forall n \in \mathbb{N}^*\ \mathbf{U} \to \mathbf{X} \to \mathbf{YZ} \text{ forms a Markov chain and } \mathbb{P}\big[\tfrac{1}{n}c_n(\mathbf{X}) \leqslant P\big] = 1\}.$$

Finally, note that the source of local randomness $(\mathcal{R}, p_R)$ can be used to prefix an arbitrary channel $\big(\mathcal{V}, \{p_{\mathbf{X}|\mathbf{V}}\}_{n\geqslant 1}, \mathcal{X}\big)$ to the broadcast channel $\big(\mathcal{X}, \mathcal{Y}, \{W_{\mathbf{YZ}|\mathbf{X}}\}_{n\geqslant 1}, \mathcal{Z}\big)$. That this prefix is useful for secrecy applications is well established [2]. By applying the proof above to the concatenated channel $\big(\mathcal{V}, \mathcal{Y}, \{p_{\mathbf{YZ}|\mathbf{V}}\}_{n\geqslant 1}, \mathcal{Z}\big)$, we conclude that the region given in Theorem 2 is included in the capacity region $\mathcal{R}_{BCC}^{(2)}$.

We now turn to the converse part of the proof. Consider a sequence of codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ achieving the rate pair (R_0, R_1) for secrecy metric $\mathbb{S}_6$. For $n \in \mathbb{N}^*$, let $\bar{\mathbf{U}}$ denote the choice of a common message uniformly at random in $[\![1, 2^{nR_0}]\!]$ and let $\bar{\mathbf{W}}$ denote the choice of an individual message uniformly at random in $[\![1, 2^{nR_1}]\!]$. Let $\bar{\mathbf{Y}}$ and $\bar{\mathbf{Z}}$ denote the channel outputs corresponding to the transmission of the message pair $(\bar{\mathbf{U}}, \bar{\mathbf{W}})$. As shown in Appendix D, the following lemma holds.

Lemma 3: If $\lim_{n\to\infty}\mathbb{P}_e(\mathcal{C}_n) = 0$ and $\lim_{n\to\infty}\mathbb{S}_6(\mathcal{C}_n) = 0$, then

$$R_0 \leqslant \min\Big(\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{U}};\bar{\mathbf{Y}}\big), \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{U}};\bar{\mathbf{Z}}\big)\Big)$$
$$R_1 \leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{W}};\bar{\mathbf{Y}}|\bar{\mathbf{U}}\big) - \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}\big).$$

Note that, by assumption, $\bar{\mathbf{U}}\bar{\mathbf{W}} \to \bar{\mathbf{X}} \to \bar{\mathbf{Y}}\bar{\mathbf{Z}}$ forms a Markov chain. Define $\bar{\mathbf{V}} \triangleq (\bar{\mathbf{U}}, \bar{\mathbf{W}})$, which is such that $\bar{\mathbf{U}} \to \bar{\mathbf{V}} \to \bar{\mathbf{X}} \to \bar{\mathbf{Y}}\bar{\mathbf{Z}}$ forms a Markov chain. With probability one, we have

$$I\big(\bar{\mathbf{W}};\bar{\mathbf{Y}}|\bar{\mathbf{U}}\big) = I\big(\bar{\mathbf{V}};\bar{\mathbf{Y}}|\bar{\mathbf{U}}\big) \quad \text{and} \quad I\big(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}\big) = I\big(\bar{\mathbf{V}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}\big);$$

therefore, an achievable pair (R_0, R_1) must satisfy

$$R_0 \leqslant \min\Big(\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{U}};\bar{\mathbf{Y}}\big), \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{U}};\bar{\mathbf{Z}}\big)\Big),$$
$$\text{and} \quad R_1 \leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{V}};\bar{\mathbf{Y}}|\bar{\mathbf{U}}\big) - \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I\big(\bar{\mathbf{V}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}\big),$$

where $\bar{\mathbf{U}} \to \bar{\mathbf{V}} \to \bar{\mathbf{X}} \to \bar{\mathbf{Y}}\bar{\mathbf{Z}}$ forms a Markov chain, $p_{\bar{\mathbf{Y}}\bar{\mathbf{Z}}|\bar{\mathbf{X}}} = W_{\mathbf{YZ}|\mathbf{X}}$, and $\mathbb{P}\big[\frac{1}{n}c_n(\bar{\mathbf{X}}) \leqslant P\big] = 1$. Taking the union over all possible processes $\{\bar{\mathbf{U}}\bar{\mathbf{V}}\bar{\mathbf{X}}\}_{n\geqslant 1}$ gives the desired outer bound for the secrecy-capacity region $\mathcal{R}_{BCC}^{(6)}$.

Since the outer bound for $\mathcal{R}_{BCC}^{(6)}$ and the inner bound for $\mathcal{R}_{BCC}^{(2)}$ match, we conclude using Proposition 1 that the secrecy-capacity region is the same for all metrics $i \in [\![2, 6]\!]$. ■

A few comments regarding Theorem 2 are now in order. First, the achievability part of the proof is based on an explicit operational interpretation of secrecy in terms of channel resolvability; in Lemma 2, codes are constructed so that, for a given message W_0 and taking the average over the random codebook selection, the probability distribution induced at the eavesdropper's channel output by all messages W_1 is asymptotically the same in the sense of variational distance. Second, the existence of a sequence of codes simultaneously satisfying the reliability and secrecy conditions is obtained by handling the constraints separately, as illustrated by the separate results of Lemma 1 and Lemma 2. This contrasts with the approach of [1], [2], in which the two constraints are handled somewhat simultaneously by using capacity-based wiretap codes. As should be clear from the condition $R' > \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U})$ obtained in Lemma 2,

the codes constructed are not capacity-based wiretap codes, for which the condition would read $R' < \text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Z}|\mathbf{U})$; essentially, channel resolvability enables the analysis of codes operating at rates beyond the capacity of the eavesdropper's channel. Finally, we note that, as in Section IV, the secrecy-capacity region is invariant with respect to the metrics $\mathbb{S}_i$ for $i \in [\![2,6]\!]$; nevertheless, practical coding schemes should be designed to provide secrecy with respect to the strongest metric.

Remark 3: If the eavesdropper's channel is exponentially information stable, so that

$$\mathbb{P}_{\mathbf{UXZ}}\left[\frac{1}{n} I(\mathbf{X};\mathbf{Z}|\mathbf{U}) > \frac{1}{n}\log M' + \epsilon\right]$$

decays exponentially fast with n for any $\epsilon > 0$, then a closer look at the proof of Theorem 2 shows that $\mathbb{S}_2(\mathcal{C}_n)$, and consequently $\mathbb{S}_1(\mathcal{C}_n)$, would also decay exponentially fast with n. We do not explore this issue further for arbitrary channels but we analyze it more precisely in the next section for memoryless channels.

Without a common message ($R_0 = 0$), we obtain in a similar way the secrecy capacity of a general wiretap channel established by Hayashi [6, Th. 5].

Corollary 1: The secrecy capacity of a wiretap channel $\left(\mathcal{X}, \mathcal{Y}, \{W_{\mathbf{YZ}|\mathbf{X}}\}_{n\geqslant 1}, \mathcal{Z}\right)$ with cost constraint P is identical for secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$ and is given by

$$C_s = \sup_{\{\mathbf{VX}\}_{n\geqslant 1}\in\mathcal{P}} \left(\text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{V};\mathbf{Y}) - \text{p-limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{V};\mathbf{Z}) \right), \quad (8)$$

where

$$\mathcal{P} \triangleq \left\{ \{\mathbf{VX}\}_{n\geqslant 1} : \forall n \in \mathbb{N}^*, \mathbf{V} \to \mathbf{X} \to \mathbf{YZ} \text{ forms a Markov chain and } \mathbb{P}\left[\tfrac{1}{n} c_n(\mathbf{X}) \leqslant P\right] = 1 \right\}.$$

C. Memoryless Broadcast Channels With Additive Cost Constraint

We now consider memoryless channels (not necessarily discrete) with an additive cost constraint. This is a special case of the general model, in which the transition probabilities factor as

$$W_{\mathbf{YZ}|\mathbf{X}}(\mathbf{y},\mathbf{z}|\mathbf{x}) = \prod_{i=1}^{n} W_{YZ|X}(y_i, z_i|x_i)$$

and the cost constraint satisfies $c_n(\mathbf{x}) = \sum_{i=1}^{n} c(x_i)$ for some cost function $c : \mathcal{X} \to \mathbb{R}^+$. For this special class of channels and constraints, and under mild conditions, the result of Section V-C extends to metric $\mathbb{S}_1$. For discrete memoryless channels without cost constraint, this result was obtained independently in [35] and [36] using secure multiplex coding and privacy amplification.

Theorem 3: The secrecy-capacity region of a memoryless broadcast channel $\left(\mathcal{X}, \mathcal{Y}, W_{YZ|X}, \mathcal{Z}\right)$ with confidential messages and additive cost constraint P is the same for all secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,4]\!]$ and is given by

$$\mathcal{R}_{BCC} = \bigcup_{(UVX)\in\mathcal{P}} \left\{ \begin{array}{l} (R_0, R_1) \in \mathbb{R}_+^2 : \\ R_0 \leqslant \min\left(\mathbb{I}(U;Y), \mathbb{I}(U;Z)\right) \\ R_1 \leqslant \mathbb{I}(V;Y|U) - \mathbb{I}(V;Z|U) \end{array} \right\}, \quad (9)$$

where

$$\mathcal{P} \triangleq \{(UVX) : U \to V \to X \to YZ \text{ forms a Markov chain and } \mathbb{E}[c(X)] \leqslant P\}.$$

If the rates on the boundary of $\mathcal{R}_{BCC}$ are obtained for some random variables $UVXYZ$ such that the integrals defining the moment generating functions of $I(V;Z|U)$ and $c(X)$ converge uniformly in a neighborhood of 0 and are differentiable at 0, then $\mathcal{R}_{BCC}$ is also the secrecy-capacity region for metric $\mathbb{S}_1$.

Proof: See Appendix E. ∎

The conditions that yield $\mathcal{R}_{BCC}$ for metric $\mathbb{S}_1$ are sufficient conditions required to obtain exponential upper bounds when applying Chernov bounds. These conditions are not too restrictive and are automatically satisfied for discrete memoryless channels and for Gaussian channels with additive power constraint. Improved exponents can be obtained in such cases using techniques as in [37].

In the absence of a common message ($R_0 = 0$), we obtain in a similar way the following result, which was already obtained for discrete memoryless channels by Csiszár [5] and Maurer and Wolf [16] with different techniques.

Corollary 2: The secrecy capacity of a memoryless wiretap channel $\left(\mathcal{X}, \mathcal{Y}, W_{YZ|X}, \mathcal{Z}\right)$ with additive cost constraint P is the same for all secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,4]\!]$ and is given by

$$C_s = \sup_{(VX)\in\mathcal{P}} \left(\mathbb{I}(V;Y) - \mathbb{I}(V;Z)\right),$$

where

$$\mathcal{P} \triangleq \{(VX) : V \to X \to YZ \text{ forms a Markov chain and } \mathbb{E}[c(X)] \leqslant P\}.$$

If the random variables $VXYZ$ maximizing C_s are such that the integrals defining the moment generating functions of $I(V;Z)$ and $c(X)$ converge uniformly in a neighborhood of 0 and are differentiable at 0, then C_s is also the secrecy capacity for metric $\mathbb{S}_1$.

For general memoryless channels, the converse part of Theorem 3 and Corollary 2 follows from standard arguments with metric $\mathbb{S}_4$ [2]; however, for discrete memoryless channels, the converse is obtained by specializing Theorem 2 and holds for metric $\mathbb{S}_6$.

Remark 4: In the proof of Theorem 3, we can actually establish a stronger result than the one stated. If the conditions for the moment generating functions of $I(V;Z|U)$ and $c(X)$ are satisfied, we can show that $\mathbb{S}_1(\mathcal{C}_n)$ vanishes exponentially fast with n.

VI. Applications

In this section, we illustrate the usefulness of deriving secrecy from channel resolvability by considering several problems in which the derivation of achievable secrecy rates is tremendously simplified. In particular, results for wireless channels, mixed wiretap channels, and compound wiretap channels come almost "for free." For simplicity, we only consider cases in which the common message rate is zero ($R_0 = 0$).

A. Ergodic Wireless Channels With Full CSI

We consider the situation in which Alice and Bob communicate over an ergodic-fading wiretap channel and have access to the instantaneous fading gains for both the main channel and the eavesdropper's channel. Specifically, at each time $k \geqslant 1$, the relationships between input and outputs are given by

$$Y_k = H_{m,k} X_k + N_{m,k},$$
$$Z_k = H_{e,k} X_k + N_{e,k},$$

where $\{H_{m,k}\}_{k\geqslant 1}$, $\{H_{e,k}\}_{k\geqslant 1}$ are fading gains known to all parties and $\{N_{m,k}\}_{k\geqslant 1}$, $\{N_{e,k}\}_{k\geqslant 1}$ are i.i.d. complex Gaussian zero-mean noise processes with respective variance σ_m^2 and σ_e^2. In addition, the channel inputs are subject to the long-term power constraint $\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[X_k^2\right] \leqslant P$.

Proposition 3: The secrecy capacity of the ergodic wireless channel with full CSI for secrecy metric $\mathbb{S}_1$ is

$$C_s = \max_{\gamma} \mathbb{E}\left[\log\left(1 + \frac{|H_m|^2 \gamma(H_m, H_e)}{\sigma_m^2}\right) - \log\left(1 + \frac{|H_e|^2 \gamma(H_m, H_e)}{\sigma_e^2}\right)\right], \quad (10)$$

where the maximization is over all power allocation functions $\gamma : \mathbb{C}^2 \to \mathbb{R}^+$ such that $\mathbb{E}[\gamma(H_m, H_e)] \leqslant P$.

Sketch of Proof: We only sketch the achievability part of the proof; the converse for secrecy metric $\mathbb{S}_4$ is established in [12]. Because the channel gains are instantaneously known to all parties, the ergodic wireless channel can be demultiplexed into a set of independent Gaussian wiretap channels, each characterized by a specific realization (h_m, h_e) of the channel gains and subject to a power constraint $\gamma(h_m, h_e)$. Upon substituting $V = 0$ and $X \sim \mathcal{N}(0, \gamma(h_m, h_e))$ in Corollary 2, we obtain the following achievable rate for metric $\mathbb{S}_1$ and for each channel:

$$\log\left(1 + \frac{|h_m|^2 \gamma(h_m, h_e)}{\sigma_m^2}\right) - \log\left(1 + \frac{|h_e|^2 \gamma(h_m, h_e)}{\sigma_e^2}\right).$$

Hence, using the ergodicity of the channel, we conclude that all the rates $R \geqslant 0$ such that

$$R < \max_{\gamma} \mathbb{E}\left[\log\left(1 + \frac{|H_m|^2 \gamma(H_m, H_e)}{\sigma_m^2}\right) - \log\left(1 + \frac{|H_e|^2 \gamma(H_m, H_e)}{\sigma_e^2}\right)\right]$$

are achievable for metric $\mathbb{S}_1$, where $\gamma : \mathbb{C}^2 \to \mathbb{R}^+$ satisfies $\mathbb{E}[\gamma(H_m, H_e) \leqslant P]$. ■

The result of Proposition 3 has already been established in [28] with a completely different approach; deriving secrecy from channel resolvability and leveraging Corollary 2 provides a much simpler and direct proof, which can be generalized to include the effect of imperfect CSI [38], [39].

B. Mixed and Compound Channels With Receiver CSI

As another application, we study mixed and compound wiretap channels with receiver CSI. These models have practical relevance since they allow one to analyze situations in which the channel is imperfectly known to the transmitter, either because the channel estimation mechanism is imperfect or because the channel is partially controlled by the eavesdropper.

Let $K \in \mathbb{N}^*$ and let $\{\alpha_k\}_{k\in[\![1,K]\!]}$ be such that $\forall k \in [\![1,K]\!]$ $\alpha_k > 0$ and $\sum_{k=1}^{K} \alpha_k = 1$. Consider the wiretap channels $(\mathcal{X}, \mathcal{Y}, \{W_{\mathbf{Y}_k\mathbf{Z}_k|\mathbf{X}}\}_{n\geqslant 1}, \mathcal{Z})$ for $k \in [\![1,K]\!]$. The mixed wiretap channel is the channel $(\mathcal{X}, \mathcal{Y}, W_{\mathbf{YZ}|\mathbf{X}}, \mathcal{Z})$ whose transition probabilities satisfy

$$W_{\mathbf{YZ}|\mathbf{X}}(\mathbf{y}, \mathbf{z}|\mathbf{x}) = \sum_{k=1}^{K} \alpha_k W_{\mathbf{Y}_k\mathbf{Z}_k|\mathbf{X}}(\mathbf{y}, \mathbf{z}|\mathbf{x}).$$

Proposition 4: The secrecy capacity of the mixed wiretap channel with power constraint P is the same for all secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$ and is given by

$$\sup_{\{\mathbf{V},\mathbf{X}\}_{n\geqslant 1}\in\mathcal{P}} \left(\min_{k\in[\![1,K]\!]} \text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{V}; \mathbf{Y}_k) - \max_{k\in[\![1,K]\!]} \text{p-limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{V}; \mathbf{Z}_k) \right), \quad (11)$$

where

$$\mathcal{P} \triangleq \{\{\mathbf{VX}\}_{n\geqslant 1} : \forall n \in \mathbb{N}^*,\ \forall k \in [\![1,K]\!],\ \mathbf{V} \to \mathbf{X} \to \mathbf{Y}_k\mathbf{Z}_k \text{ forms a Markov chain and } \mathbb{P}\left[\tfrac{1}{n} c_n(\mathbf{X}) \leqslant P\right] = 1\}.$$

Proof: Using [10, Lemma 1.4.2], we obtain

$$\text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{V}; \mathbf{Y}) = \min_{k\in[\![1,K]\!]} \left(\text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{V}; \mathbf{Y}_k) \right)$$
$$\text{p-limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{V}; \mathbf{Z}) = \max_{k\in[\![1,K]\!]} \left(\text{p-limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{V}; \mathbf{Z}_k) \right).$$

The result follows by substituting these equalities in Corollary 1. ■

Note that for $i \in [\![1,2]\!]$, we have $\mathbb{S}_i(p_{W\bar{\mathbf{Z}}}, p_W p_{\bar{\mathbf{Z}}}) \leqslant \sum_{k=1}^{K} \alpha_k \mathbb{S}_i(p_{W\bar{\mathbf{Z}}_k}, p_W p_{\bar{\mathbf{Z}}_k})$. Therefore, a code ensuring secrecy for the mixed wiretap channel may not guarantee secrecy over each individual wiretap channel. If one wants to ensure secrecy over all possible K channels, one must consider a compound wiretap channel, in which the transmitter has no knowledge (even statistical knowledge) of which channel in the set is used for transmission; however, to avoid unnecessary mathematical complications, we assume that receivers can estimate channel statistics perfectly and always know from which channel they obtain observations; hence, we refer to this model as a compound channel with receiver CSI. For every channel $k \in [\![1,K]\!]$, the performance of a code $\mathcal{C}_n$ is measured in terms of the average probability of error $\mathbb{P}_e^{(k)}(\mathcal{C}_n)$ and in

terms of the secrecy metric $\mathbb{S}_i^{(k)}(\mathcal{C}_n) \triangleq \mathbb{S}_i(p_{W\bar{\mathbf{Z}}_k}, p_W p_{\bar{\mathbf{Z}}_k})$; the notion of achievable rate is accordingly modified as follows.

Definition 7: A rate R is achievable over a compound wiretap channel with receiver CSI for secrecy metric $\mathbb{S}_i$ if there exists a sequence of $(2^{nR}, n)$ wiretap codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\forall k \in [\![1,K]\!] \quad \lim_{n\to\infty} \mathbb{P}_e^{(k)}(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{S}_i^{(k)}(\mathcal{C}_n) = 0.$$

Unlike the mixed wiretap channel, there is no distribution associated to the choice of the channel in the set, and secrecy and reliability must be guaranteed for any realized channel.

Proposition 5: The secrecy capacity of a compound wiretap channel with receiver CSI and with cost constraint P is the same for all secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$ and is given by

$$\sup_{\{\mathbf{V},\mathbf{X}\}_{n\geqslant 1}\in\mathcal{P}} \left(\min_{k\in[\![1,K]\!]} \operatorname*{p\text{-}liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{V};\mathbf{Y}_k) - \max_{k\in[\![1,K]\!]} \operatorname*{p\text{-}limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{V};\mathbf{Z}_k) \right), \quad (12)$$

where

$$\mathcal{P} \triangleq \{\{\mathbf{VX}\}_{n\geqslant 1} : \forall n \in \mathbb{N}^*,\ \forall k \in [\![1,K]\!],\ \mathbf{V} \to \mathbf{X} \to \mathbf{Y}_k\mathbf{Z}_k \text{ forms a Markov chain and } \mathbb{P}\left[\tfrac{1}{n} c_n(\mathbf{X}) \leqslant P\right] = 1\}.$$

Proof: We start with the achievability part of the proof, which is similar to that of Theorem 2. Let $n \in \mathbb{N}^*$ and $\epsilon, \gamma, R_1, R' > 0$. Define $M_1 \triangleq \lceil 2^{nR_1} \rceil$ and $M' \triangleq \lceil 2^{nR'} \rceil$. Fix a distribution $p_{\mathbf{X}}$ on $\mathcal{X}^n$ such that $\mathbb{P}\left[\frac{1}{n} c_n(\mathbf{X}) \leqslant P\right] = 1$. Let $\mathbf{X}$, $\{\mathbf{Y}_k\}_{k\in[\![1,K]\!]}$, $\{\mathbf{Z}_k\}_{k\in[\![1,K]\!]}$ be the random variables with joint distribution

$$\forall k \in [\![1,K]\!] \ \ p_{\mathbf{X}\mathbf{Y}_k\mathbf{Z}_k}(\mathbf{x},\mathbf{y},\mathbf{z}) \triangleq W_{\mathbf{Y}_k\mathbf{Z}_k|\mathbf{X}}(\mathbf{y},\mathbf{z}|\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}).$$

1) *Code generation:* Randomly generate $M_1 M'$ sequences $\mathbf{x}_{lm} \in \mathcal{X}^n$ with $(l,m) \in [\![1,M_1]\!] \times [\![1,M']\!]$ according to $p_{\mathbf{X}}$. We denote by C_n the random variable representing the generated code and by $\mathcal{C}_n$ one of its realizations.
2) *Encoding:* To transmit a message $l \in [\![1,M_1]\!]$, Alice generates an auxiliary message m uniformly at random in $[\![1,M']\!]$ and transmits the codeword $\mathbf{x}_{lm}$ through the channel.
3) *Bob's decoding for channel $k \in [\![1,K]\!]$:* Define the set

$$\mathcal{T}_k^n \triangleq \{(\mathbf{x},\mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}_k^n : \frac{1}{n}\log \frac{W_{\mathbf{Y}_k|\mathbf{X}}(\mathbf{y}|\mathbf{x})}{p_{\mathbf{Y}_k}(\mathbf{y})} \geqslant \frac{1}{n}\log M_1 M' + \gamma\}.$$

Note that the decoding rule depends on the channel index k since we have assumed that Bob knows which channel is being observed. Upon observing $\mathbf{y}_k$, Bob decodes l as the received individual message and m as the received auxiliary message if there exists a unique codeword $\mathbf{x}_{lm}$ such that $(\mathbf{x}_{lm}, \mathbf{y}_k) \in \mathcal{T}_k^n$; otherwise, a random message is chosen.

The following lemmas provide sufficient conditions to guarantee reliability and secrecy. Their proofs are similar to those provided in Appendix C and are omitted.

Lemma 4 (Reliability Conditions): For each $k \in [\![1,K]\!]$,

$$\text{If } R_1 + R' \leqslant \operatorname*{p\text{-}liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Y}_k) - 2\gamma \quad \text{then } \lim_{n\to\infty} \mathbb{E}[\mathbb{P}_e^{(k)}(C_n)] \leqslant \epsilon.$$

Lemma 5 (Secrecy From Channel Resolvability Condition): For each $k \in [\![1,K]\!]$,

$$\text{If } R' \geqslant \operatorname*{p\text{-}limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Z}_k) + 2\gamma \text{ , then } \lim_{n\to\infty} \mathbb{E}\left[\mathbb{S}_2^{(k)}(C_n)\right] \leqslant \epsilon.$$

Using Lemmas 4 and 5, we obtain that if

$$R_1 \leqslant \min_{k\in[\![1,K]\!]} \operatorname*{p\text{-}liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Y}_k) - \max_{k\in[\![1,K]\!]} \operatorname*{p\text{-}liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Z}_k) - 4\gamma$$
$$\text{then } \forall k \in [\![1,K]\!] \ \begin{cases} \lim_{n\to\infty} \mathbb{E}[\mathbb{P}_e^{(k)}(C_n)] \leqslant \epsilon \\ \lim_{n\to\infty} \mathbb{E}\left[\mathbb{S}_2^{(k)}(C_n)\right] \leqslant \epsilon. \end{cases}$$

Using Markov's inequality and the union bound, we can show there exists at least one sequence of $(2^{nR_1}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\forall k \in [\![1,K]\!] \quad \lim_{n\to\infty} \mathbb{P}_e^{(k)}(\mathcal{C}_n) \leqslant (K+1)\epsilon \quad \text{and } \lim_{n\to\infty} \mathbb{S}_2^{(k)}(\mathcal{C}_n) \leqslant (K+1)\epsilon.$$

Since K is fixed and ϵ, γ can be chosen arbitrarily small, we conclude that all rates R_1 such that

$$0 \leqslant R_1 < \sup_{\{\mathbf{X}\}_{n\geqslant 1}\in\mathcal{P}} \left(\min_{k\in[\![1,K]\!]} \operatorname*{p\text{-}liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Y}_k) - \max_{k\in[\![1,K]\!]} \operatorname*{p\text{-}limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Z}_k) \right) \quad (13)$$

are achievable, where

$$\mathcal{P} \triangleq \left\{\{X_n\}_{n\geqslant 1} : \mathbb{P}\left[\tfrac{1}{n} c_n(\mathbf{X}) \leqslant P\right] = 1\right\}.$$

The achievability of the rates below the secrecy capacity $C_s^{(2)}$ in (12) is then obtained by introducing a prefix channel as in the proof of Theorem 2.

We now turn to the converse part of the proof. Consider a sequence of wiretap codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ achieving rate R_1 for secrecy metric $\mathbb{S}_6$. For $n \in \mathbb{N}^*$, let $\bar{\mathbf{V}}$ denote the choice of a message uniformly at random in $[\![1, 2^{nR_1}]\!]$. By definition, for every $n \in \mathbb{N}^*$ and $k \in [\![1,K]\!]$, $\bar{\mathbf{V}} \to \bar{\mathbf{X}} \to \bar{\mathbf{Y}}_k\bar{\mathbf{Z}}_k$ forms a Markov chain and $\mathbb{P}\left[\frac{1}{n} c(\bar{\mathbf{X}}) \leqslant P\right] = 1$. By the Verdú–Han Lemma [26, Th. 4], we obtain

$$R_1 \leqslant \min_{k\in[\![1,K]\!]} \operatorname*{p\text{-}liminf}_{n\to\infty} \frac{1}{n} I\big(\bar{\mathbf{V}};\bar{\mathbf{Y}}_k\big). \quad (14)$$

By definition of the metric $\mathbb{S}_6$, we also have

$$\max_{k\in[\![1,K]\!]} \operatorname*{p\text{-}limsup}_{n\to\infty} \frac{1}{n} I\big(\bar{\mathbf{V}};\bar{\mathbf{Z}}_k\big) = 0. \quad (15)$$

Subtracting (15) from (14), and maximizing over all processes $\{\bar{\mathbf{V}}\bar{\mathbf{X}}\}$, we obtain the desired result. ∎

Although the secrecy capacity of a compound wiretap channel with receiver CSI is identical to that of a mixed wiretap channel, the coding schemes achieving it may be fundamentally different.

Proposition 6: Given a memoryless compound wiretap channel with receiver CSI and additive cost constraint P, all rates R_1 such that

$$0 \leqslant R_1 < \sup_{(VX)\in\mathcal{P}} \left(\min_{k\in[\![1,K]\!]} \mathbb{I}(V;Y_k) - \max_{k\in[\![1,K]\!]} \mathbb{I}(V;Z_k) \right) \tag{16}$$

are achievable for secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$, where

$$\mathcal{P} \triangleq \{VX : \forall k \in [\![1,K]\!],\, V \to X \to Y_k Z_k \text{ forms a Markov chain and } \mathbb{E}[c(X)] \leqslant P\}.$$

If the random variables maximizing (16) are such that, for all $k \in [\![1,K]\!]$, the integrals defining the moment generating functions of $I(V;Y_k)$ and $c(X)$ converge uniformly in a neighborhood of 0 and are differentiable at 0, then the rates are also achievable for metric $\mathbb{S}_1$.

Proof: The proof of Proposition 6 follows from steps similar to those used in the proof of Proposition 5 and Theorem 3 and is omitted. ∎

If the receivers do not know which channel they observe, the counterpart of Proposition 6 was independently derived in [40]. Note that deriving secrecy from channel resolvability circumvents the enhancement argument used in [41, Th. 1], which is required to show achievability using capacity-based wiretap codes. Similarly, when applied to Gaussian compound wiretap channels with power constraint, Proposition 6 strengthens [42, Th. 1] with receiver CSI.

Remark 5: The general result of Proposition 5 holds provided the number of channels K is fixed and independent of the number n of channel uses; nevertheless, in the special case of Proposition 6, for which we establish secrecy for metric $\mathbb{S}_1$, we can show that, for each $k \in [\![1,K]\!]$, $\mathbb{S}_1^{(k)} \leqslant (K+1)2^{-\epsilon_k n}$ for some $\epsilon_k > 0$. Therefore, Proposition 6 also holds if the number of compound channels grows exponentially with n as $K = 2^{\beta n}$ with $\beta < \min_{k\in[\![1,K]\!]} \epsilon_k$.

C. Secret-Key Agreement From General Sources

As a last application, we exploit a connection between secret-key agreement and wiretap coding to analyze the fundamental limits of secret-key agreement for a general source model. Specifically, we consider a *discrete* source $(\mathcal{X},\mathcal{Y},\mathcal{Z},\{p_{\mathbf{XYZ}}\}_{n\geqslant 1})$ with three components taking values in discrete alphabets. As illustrated in Fig. 5, Alice and Bob attempt to distill a secret key from their correlated observations $\mathbf{X}$ and $\mathbf{Y}$, respectively, and a message transmitted by Alice over public authenticated channel with unlimited capacity. The key should remain secret from an eavesdropper who observes $\mathbf{Z}$ and the public message.

Definition 8: A $(2^{nR}, n)$ key-distillation strategy $\mathcal{S}_n$ consists of

1) a key alphabet $\mathcal{K} = [\![1, 2^{nR}]\!]$;
2) an alphabet $\mathcal{A}$ used by Alice to communicate over the public channel;

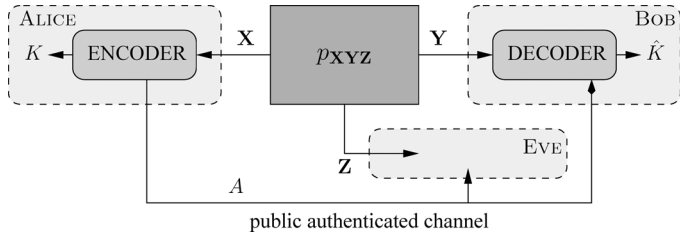

Fig. 5. Secret-key agreement from general source.

3) a source of local randomness for Alice $(\mathcal{R}_\mathcal{X}, p_{R_X})$;
4) a source of local randomness for Bob $(\mathcal{R}_\mathcal{Y}, p_{R_Y})$;
5) an encoding function $f : \mathcal{X}^n \times \mathcal{R}_\mathcal{X} \to \mathcal{A}$;
6) a key-distillation function $\kappa_a : \mathcal{X}^n \times \mathcal{R}_\mathcal{X} \to \mathcal{K}$;
7) a key-distillation function $\kappa_b : \mathcal{Y}^n \times \mathcal{A} \times \mathcal{R}_\mathcal{Y} \to \mathcal{K}$;

The random variables corresponding to the public message, Alice's key, and Bob's key are denoted by A, K, and $\hat{K}$, respectively. The performance of a secret-key distillation strategy $\mathcal{S}_n$ is measured in terms of the average probability of error $\mathbb{P}_e(\mathcal{S}_n) \triangleq \mathbb{P}\left[K \neq \hat{K} \middle| \mathcal{S}_n\right]$, the secrecy of the key $\mathbb{S}_i(\mathcal{S}_n) \triangleq \mathbb{S}_i(p_{K\mathbf{Z}A|\mathcal{S}_n}, p_{K|\mathcal{S}_n} p_{\mathbf{Z}A|\mathcal{S}_n})$ for $i \in [\![1,6]\!]$, and the uniformity of the key $\mathbb{U}(\mathcal{S}_n) \triangleq \log\lceil 2^{nR}\rceil - \mathbb{H}(K)$.

Definition 9: A key rate R is achievable for secrecy metric $\mathbb{S}_i$ for a source if there exists a sequence $\{\mathcal{S}_n\}_{n\geqslant 1}$ of $\left(2^{nR}, n\right)$ key-distillation strategies such that

$$\lim_{n\to\infty} \mathbb{P}_e(\mathcal{S}_n) = 0, \quad \lim_{n\to\infty} \mathbb{S}_i(\mathcal{S}_n) = 0, \quad \lim_{n\to\infty} \mathbb{U}(\mathcal{S}_n) = 0.$$

The forward secret-key capacity $C_{SK}^{(i)}$ is the supremum of achievable key rates for metric $\mathbb{S}_i$.

Proposition 7: The forward secret-key capacity of a discrete source $(\mathcal{X},\mathcal{Y},\mathcal{Z},\{p_{\mathbf{XYZ}}\}_{n\geqslant 1})$ for secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$ satisfies

$$\begin{aligned} &\operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Z}) - \operatorname*{p\text{-limsup}}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Y}) \leqslant C_{SK}^{(i)} \\ &\leqslant \min\left(\operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Y}), \operatorname*{p\text{-liminf}}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Y}|\mathbf{Z}) \right). \end{aligned} \tag{17}$$

If the discrete source is i.i.d., Proposition 7 holds for secrecy metric $\mathbb{S}_1$, as already known from [5] and [16].

Corollary 3: The secret-key capacity of an i.i.d discrete source $(\mathcal{X},\mathcal{Y},\mathcal{Z},p_{XYZ})$ for secrecy metric $\mathbb{S}_1$ satisfies

$$\begin{aligned} &\max\left(\mathbb{I}(X;Y) - \mathbb{I}(X;Z), \mathbb{I}(X;Y) - \mathbb{I}(Y;Z)\right) \\ &\qquad \leqslant C_{SK}^{(1)} \leqslant \min\left(\mathbb{I}(X;Y), \mathbb{I}(X,Y|Z)\right). \end{aligned}$$

Proof of Theorem 7 and Corollary 3: The achievability part of Theorem 7 is based on the construction of a conceptual wiretap channel as in [13]. Assume that Alice, Bob, and Eve observe n realizations $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ of the source, respectively. Consider an arbitrary process $\{\mathbf{U}\}_{n\geqslant 1}$ such that $\mathbf{U} \in \mathcal{X}^n$. Assume that Alice forms the signal $\mathbf{U} \oplus \mathbf{X}$ on the public channel, in which $\oplus$ denotes the symbol-wise modulo-$\mathcal{X}$ addition. This operation creates a conceptual wiretap channel with input $\mathbf{U}$, in which Bob observes the outputs $\mathbf{Y}$ and $\mathbf{U} \oplus \mathbf{X}$ while Eve observes the outputs $\mathbf{Z}$ and $\mathbf{U} \oplus \mathbf{X}$. From Corollary 1, the secrecy

capacity of this conceptual channel for secrecy metrics $\mathbb{S}_i$ with $i \in [\![2,6]\!]$ is at least

$$\sup_{\mathbf{U}} \left(\operatorname*{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{U};\mathbf{Y},\mathbf{U}\oplus\mathbf{X}) - \operatorname*{p-limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{U};\mathbf{Z},\mathbf{U}\oplus\mathbf{X}) \right).$$

In particular, we can choose for $\mathbf{U}$ an i.i.d. process such that, for all $j \in \mathbb{N}^*$, U_j is independent of $\mathbf{XYZ}$ and uniformly distributed in $\mathcal{X}$. Then, with probability 1,

$$\begin{aligned} I(\mathbf{U};\mathbf{Y},\mathbf{U}\oplus\mathbf{X}) &= \log \frac{p_{\mathbf{U}\oplus\mathbf{X},\mathbf{Y}|\mathbf{U}}(\mathbf{U}\oplus\mathbf{X},\mathbf{Y}|\mathbf{U})}{p_{\mathbf{U}\oplus\mathbf{X},\mathbf{Y}}(\mathbf{U}\oplus\mathbf{X},\mathbf{Y})} \\ &= \log \frac{p_{\mathbf{X}|\mathbf{YU}}(\mathbf{X}|\mathbf{YU}) p_{\mathbf{Y}|\mathbf{U}}(\mathbf{Y}|\mathbf{U})}{p_{\mathbf{U}\oplus\mathbf{X}|\mathbf{Y}}(\mathbf{U}\oplus\mathbf{X}|\mathbf{Y}) p_{\mathbf{Y}}(\mathbf{Y})} \\ &= \log p_{\mathbf{X}|\mathbf{Y}}(\mathbf{X}|\mathbf{Y}) - \log \frac{1}{|\mathcal{X}|^n}, \end{aligned}$$

where the last inequality follows from $p_{\mathbf{Y}|\mathbf{U}}(\mathbf{Y}|\mathbf{U}) = p_{\mathbf{Y}}(\mathbf{Y})$, $p_{\mathbf{X}|\mathbf{YU}}(\mathbf{X}|\mathbf{YU}) = p_{\mathbf{X}|\mathbf{Y}}(\mathbf{X}|\mathbf{Y})$ since $\mathbf{U}$ is independent of $\mathbf{XY}$ and $p_{\mathbf{U}\oplus\mathbf{X}|\mathbf{Y}}(\mathbf{U}\oplus\mathbf{X}|\mathbf{Y}) = \frac{1}{|\mathcal{X}|^n}$ by the crypto lemma [9]. Therefore,

$$\operatorname*{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{U};\mathbf{Y},\mathbf{U}\oplus\mathbf{X}) = \log|\mathcal{X}| - \operatorname*{p-limsup}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Y}). \quad (18)$$

Similarly, one obtains

$$\operatorname*{p-limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{U};\mathbf{Z},\mathbf{U}\oplus\mathbf{X}) = \log|\mathcal{X}| - \operatorname*{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Z}). \quad (19)$$

Combining (18) and (19), we conclude that any rate R such that

$$R < \operatorname*{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Z}) - \operatorname*{p-limsup}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Y})$$

is an achievable rate for the conceptual wiretap channel. Since this channel allows one to transmit uniformly distributed messages, R is also an achievable secret-key rate for the source model. For i.i.d. discrete sources, a similar proof based on Corollary 2 in place of Corollary 1 shows that the result holds for metric $\mathbb{S}_1$ as well. The proof of the converse is an information-spectrum version of the converse in [14] and is omitted for brevity. ∎

In Proposition 7, achievable key rates are expressed in terms of conditional entropy; except in some special cases, such as i.i.d. sources, this is rather different from the achievable secrecy rates for wiretap channels in Corollary 1, which are expressed in terms of mutual information. In particular, if $\operatorname*{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{X}) = \operatorname*{p-limsup}_{n\to\infty} \frac{1}{n} H(\mathbf{X})$, then

$$\begin{aligned} &\operatorname*{p-liminf}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Z}) - \operatorname*{p-limsup}_{n\to\infty} \frac{1}{n} H(\mathbf{X}|\mathbf{Y}) \\ &\qquad \geqslant \operatorname*{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Y}) - \operatorname*{p-limsup}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Z}). \end{aligned}$$

This distinction suggests that the coding mechanism for secret-key distillation, which one would have to exploit to design secret-key distillation strategies without relying on the existence of wiretap codes, is not linked to channel resolvability; indeed, the first author has argued in a previous work that secret-key distillation is more easily understood in terms of channel intrinsic randomness [5], [24] and privacy amplification [17], [37]. In that respect, the proof of Proposition 7 provides limited insight into the design of practical secret-key distillation strategies.

VII. Conclusion

We have analyzed several models of secure communication by building upon the work of Csiszár [5] and Hayashi [6] and by exploiting the idea that the coding mechanism to ensure secrecy can be tied to channel resolvability. This approach has allowed us to establish several results for generic channels and for stronger secrecy metrics than the usual average mutual information rate between messages and eavesdropper's observations.

From a technical point of view, deriving secrecy from channel resolvability provides a conceptually simple approach to analyze the secure achievable rates of many models. Although we have limited applications to mixed, compound, and wireless channels, the connection between secrecy and channel resolvability is useful in many other settings. Examples of secure communication models for which deriving secrecy from channel resolvability simplifies the analysis include queuing channels [43], wireless channels with imperfect state information [38], [39], runlength-limited channels [44], and two-way wiretap channels [45].

From a practical perspective, we believe that the connection between strong secrecy and channel resolvability opens intriguing perspectives for code design. In particular, we have provided evidence that this connection circumvents a weakness of capacity-based wiretap codes, which cannot always achieve the strong secrecy capacity. This observation is consistent with practical code constructions achieving strong secrecy rates [33], [46] and other approaches based on privacy amplification [16], [29], [35].

Our results can be extended in several directions. For instance, the coding mechanisms for secrecy presented in Section IV for Shannon's cipher system and in Section V for wiretap channels can be combined without much difficulty using a coding scheme similar to that proposed in [47]. One could also further investigate the nature of the coding mechanisms for secrecy in secret-key agreement models. Some results along these lines are already available, for instance in [5], [24] and [34].

Appendix A
Supporting Lemmas

Lemma 6 (Chernov Bound): Let X be a real-valued random variable with moment generating function $\phi_X : \mathbb{R} \to \mathbb{R} : s \mapsto \mathbb{E}\left[e^{sX}\right]$. Let $\{X_i\}_{i=1}^n$ be i.i.d. with distribution p_X. If the integral defining ϕ_X converges uniformly in a neighborhood of 0 and is differentiable at 0, then $\forall \epsilon > 0\ \exists \alpha_\epsilon > 0$ such that

$$\mathbb{P}\left[\frac{1}{n}\sum_{i=1}^n X_i > \mathbb{E}[X] + \epsilon\right] \leqslant 2^{-\alpha_\epsilon n}.$$

Lemma 7 (Basic Properties of Variational Distance): Let X_1, X_2, and X_3 be random variables defined on the same alphabet $\mathcal{X}$. Then,

$$\begin{aligned} &\mathbb{V}(p_{X_1};p_{X_3}) \leqslant \mathbb{V}(p_{X_1};p_{X_2}) + \mathbb{V}(p_{X_2};p_{X_3}),\\ \text{and}\quad &\mathbb{V}(p_{X_1};p_{X_2}) \leqslant \mathbb{V}(p_{X_1}p_{X_3};p_{X_2X_3})\\ &\qquad = \mathbb{E}_{X_3}\left[\mathbb{V}\big(p_{X_1},p_{X_2|X_3}\big)\right]. \end{aligned}$$

Lemma 8 (Data-Processing Inequality for Variational Distance): Let X_1 and X_2 be random variables defined on the same alphabet $\mathcal{X}$. Let $W_{Z|X}$ be the transition probability from $\mathcal{X}$ to $\mathcal{Z}$ and define the random variables Z_1 and Z_2 such that

$$\begin{aligned} \forall (z,x) \in \mathcal{Z}\times\mathcal{X} \quad & p_{Z_1X_1}(z,x) = W_{Z|X}(z|x)p_{X_1}(x)\\ &\text{and } p_{Z_2X_2}(z,x) = W_{Z|X}(z|x)p_{X_2}(x). \end{aligned}$$

Then, $\mathbb{V}(p_{Z_1},p_{Z_2}) \leqslant \mathbb{V}(p_{X_1},p_{X_2})$.

Appendix B
Proof of Proposition 2

Let C_n be the random variable that denotes a randomly generated capacity-based wiretap code, whose codeword symbols are generated i.i.d. according to the uniform distribution q_X. Let p_Z be the output distribution of the eavesdropper's channel corresponding to the input on $\mathcal{X}$, i.e.,

$$\forall z \in \mathcal{Z},\ p_Z(z) = \sum_{x\in\mathcal{X}} W_{Z|X}(z|x)\frac{1}{|\mathcal{X}|}.$$

Let $p_{\mathbf{Z}}$ be the distribution of n i.i.d. random variables distributed according to p_Z. The proof of the proposition relies on the following lemmas.

Lemma 9: Consider $M_n \triangleq 2^{nR}$ codewords of length n, obtained by generating codeword symbols independently and uniformly at random in $\mathcal{X}$. If $R < \frac{1}{2}\log|\mathcal{X}|$, there exists $\alpha_0 > 0$ such that the probability that all M_n codewords are distinct satisfies

$$\mathbb{P}[\text{all } M_n \text{ codewords are distinct}] \geqslant 1 - 2^{-\alpha_0 n}.$$

Proof: The proof follows from the same technique as in [48, Lemma 6], which we recall for convenience. Note that,

$$\begin{aligned} &\mathbb{P}\left(\text{all } M_n \text{ codewords are distinct}\right)\\ &\qquad = \prod_{i=0}^{M_n-1}\frac{|\mathcal{X}|^n - i}{|\mathcal{X}|^n} = \prod_{i=0}^{M_n-1}\left(1-\frac{i}{|\mathcal{X}|^n}\right). \end{aligned}$$

Since $\ln(1-x) \geqslant \frac{-x}{1-x}$ for $x\in[0,1)$, we have

$$\begin{aligned} &\mathbb{P}(\text{all } M_n \text{ codewords are distinct})\\ &\qquad \geqslant \exp\left(-\sum_{i=0}^{M_n-1}\frac{i}{|\mathcal{X}|^n - i}\right)\\ &\qquad \geqslant \exp\left(-\frac{(M_n-1)(M_n-1)}{|\mathcal{X}|^n-(M_n-1)}\right). \end{aligned}$$

Since $e^{-x} \geqslant 1-x$, we obtain

$$\begin{aligned} &\mathbb{P}\left(\text{all } M_n \text{ codewords are distinct}\right)\\ &\qquad \geqslant 1 - \frac{(M_n-1)^2}{|\mathcal{X}|^n-(M_n-1)} \geqslant 1 - \frac{M_n^2}{|\mathcal{X}|^n - M_n}. \end{aligned}$$

Substituting $M_n = 2^{nR}$, we obtain

$$\mathbb{P}\left(\text{all } M_n \text{ codewords are distinct}\right) \geqslant 1 - \frac{2^{2nR}}{|\mathcal{X}|^n - 2^{nR}},$$

which goes to 1 as n goes to infinity provided $R < \frac{1}{2}\log|\mathcal{X}|$. ■

Lemma 10: There exists $\alpha_1 > 0$, such that, for n sufficiently large,

$$\mathbb{P}[\mathbb{P}_e^*(C_n) \leqslant \epsilon_n' \text{ and } \mathbb{S}_4(C_n) \leqslant 3\epsilon_n'] \geqslant 1 - 2^{-\alpha_1 n\epsilon_n^2},$$

with $\epsilon_n' \triangleq \max(\epsilon_n, \log|\mathcal{X}|2^{-\alpha_1 n\epsilon_n^2}, n^{-1})$.

Proof: The existence of $\alpha_1 > 0$ such that $\mathbb{P}\left[\mathbb{P}_e^*(C_n) \leqslant 2^{-\alpha_1 n\epsilon_n^2}\right] \geqslant 1 - 2^{-\alpha_1 n\epsilon_n^2}$ follows from a standard random coding argument. Consider a code $\mathcal{C}_n$ such that $\mathbb{P}_e^*(\mathcal{C}_n) \leqslant 2^{-\alpha_1 n\epsilon_n^2}$. Then, for n large enough,

$$\begin{aligned} \mathbb{S}_4(\mathcal{C}_n) &= \tfrac{1}{n}\mathbb{I}\big(W_1;\bar{\mathbf{Z}}\big)\\ &= \tfrac{1}{n}\mathbb{I}\big(W_1W';\bar{\mathbf{Z}}\big) - \tfrac{1}{n}\mathbb{I}\big(W';\bar{\mathbf{Z}}|W_1\big)\\ &\overset{(a)}{\leqslant} \tfrac{1}{n}\mathbb{I}\big(\bar{\mathbf{X}};\bar{\mathbf{Z}}\big) - \tfrac{1}{n}\mathbb{H}(W'|W_1) + \tfrac{1}{n}\mathbb{H}\big(W'|W_1\bar{\mathbf{Z}}\big)\\ &\overset{(b)}{\leqslant} C_e - (C_e-\epsilon_n) + R'\mathbb{P}_e^*(\mathcal{C}_n) + \tfrac{1}{n}\\ &\leqslant 3\epsilon_n' \end{aligned}$$

where (a) follows because $W_1W' \to \bar{\mathbf{X}} \to \bar{\mathbf{Z}}$ forms a Markov chain, and (b) follows from Fano's inequality. ■

Lemma 11: There exists $\beta, \alpha_2 > 0$, such that, for n large enough

$$\mathbb{P}\left[\mathbb{V}(p_{\bar{\mathbf{Z}}}, p_{\mathbf{Z}}) \leqslant 2^{-\beta n}\right] \geqslant 1 - 2^{-\alpha_2 n}.$$

Proof: This result follows from [10, Th. 6.3.1] by remarking that memoryless channels are exponentially information stable or, alternatively, from [49, Lemma 19]. ■

For $n \in \mathbb{N}^*$, let $\mathcal{C}_n$ denote a randomly generated code such that all codewords are distinct and

$$\mathbb{P}_e^*(\mathcal{C}_n) \leqslant \epsilon_n',\ \mathbb{S}_4(\mathcal{C}_n) \leqslant 3\epsilon_n',\ \text{and } \mathbb{V}(p_{\bar{\mathbf{Z}}}, p_{\mathbf{Z}}) \leqslant 2^{-\beta n}. \quad (20)$$

For n large enough, Lemmas 9–11 guarantee that this occurs with probability at least $1-2^{-\alpha n\epsilon_n^2}$ for some $\alpha < \alpha_1$ and n large enough. With a slight abuse of notation, we also let $\mathcal{C}_n \subset \mathcal{X}^n$ denote the codebook and let $f_n^{-1} : \mathcal{C}_n \to \mathcal{M}_1$ be the restriction to $\mathcal{M}_1$ of the inverse mapping of f_n, which is well defined since codewords are distinct. Let us introduce the functions ϕ_n and ψ_n as

$$\begin{aligned} &\phi_n : \mathcal{C}_n \to \mathcal{M}_1 : \mathbf{x} \mapsto f_n^{-1}(\mathbf{x})\\ \text{and}\quad &\psi_n : \mathcal{Z}^n \times \mathcal{M}_1 \to \mathcal{C}_n : (\mathbf{z}, m) \mapsto f_n(m, h_n(\mathbf{z}, m)). \end{aligned}$$

The functions ϕ_n and ψ_n define the encoder and decoder of a source code for the compression of the source $\bar{\mathbf{X}} \in \mathcal{C}_n$ (the

choice of codewords uniformly at random in the code) with $\bar{\mathbf{Z}}$ as correlated side information at the receiver, whose probability of decoding error is $\mathbb{P}_e^*(\mathcal{C}_n)$. We now leverage the results obtained by Hayashi [50] and generalized by Watanabe *et al.* [34] that establish a tradeoff between probability and error and resolvability for source coding of arbitrary sources. Combining [34, Th. 10] and the proof of [34, Th. 11], we obtain, $\forall b > 0$, $\forall n \in \mathbb{N}^*$

$$\mathbb{P}_e^*(\mathcal{C}_n) + \mathbb{S}_2(\mathcal{C}_n) \geqslant 1 - \left(2^{-b\sqrt{n}+1} + \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{A}_n]\right), \quad (21)$$

with

$$\mathcal{A}_n \triangleq \left\{(\mathbf{x},\mathbf{z}) \in \mathcal{C}_n \times \mathcal{Z}^n : \frac{2^{-b\sqrt{n}}}{|\mathcal{M}_1|} < p_{\bar{\mathbf{X}}|\bar{\mathbf{Z}}}(\mathbf{x}|\mathbf{z}) \leqslant \frac{2^{b\sqrt{n}}}{|\mathcal{M}_1|}\right\}.$$

Note that $|\mathcal{M}'| = 2^{n(C_e-\epsilon_n)}$ and $p_{\bar{\mathbf{X}}}(\bar{\mathbf{X}}) = \frac{1}{|\mathcal{M}_1||\mathcal{M}'|}$. Therefore, by Bayes's rule

$$\begin{aligned}\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{A}_n] &= \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\frac{2^{-b\sqrt{n}}}{|\mathcal{M}_1|} < p_{\bar{\mathbf{X}}|\bar{\mathbf{Z}}}(\bar{\mathbf{X}}|\bar{\mathbf{Z}}) \leqslant \frac{2^{b\sqrt{n}}}{|\mathcal{M}_1|}\right] \\ &= \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\frac{2^{-b\sqrt{n}}}{|\mathcal{M}_1|} < W_{\mathbf{Z}|\mathbf{X}}(\bar{\mathbf{Z}}|\bar{\mathbf{X}})\frac{p_{\bar{\mathbf{X}}}(\bar{\mathbf{X}})}{p_{\bar{\mathbf{Z}}}(\bar{\mathbf{Z}})} \leqslant \frac{2^{b\sqrt{n}}}{|\mathcal{M}_1|}\right] \\ &= \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^+\right] - \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^-\right],\end{aligned}$$

where we have defined

$$\mathcal{Q}_n^{\pm} \triangleq \left\{(\mathbf{x},\mathbf{z}) \in \mathcal{C}_n \times \mathcal{Z}^n : \log\frac{W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x})}{p_{\bar{\mathbf{Z}}}(\mathbf{z})} \leqslant \pm b\sqrt{n} + n(C_e - \epsilon_n)\right\}.$$

We analyze $\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{Q}_n^+]$ and $\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{Q}_n^-]$ by introducing the sets

$$\begin{aligned}\mathcal{A}_n^{\pm} &\triangleq \left\{(\mathbf{x},\mathbf{z}) \in \mathcal{C}_n \times \mathcal{Z}^n : \log\frac{W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x})}{p_{\mathbf{Z}}(\mathbf{z})} \leqslant \pm 2b\sqrt{n} + n(C_e - \epsilon_n)\right\} \\ \mathcal{B}_n &\triangleq \left\{(\mathbf{x},\mathbf{z}) \in \mathcal{C}_n \times \mathcal{Z}^n : \log\frac{p_{\bar{\mathbf{Z}}}(\mathbf{z})}{p_{\mathbf{Z}}(\mathbf{z})} < b\sqrt{n}\right\} \\ \text{and} \quad \mathcal{D}_n &\triangleq \left\{(\mathbf{x},\mathbf{z}) \in \mathcal{C}_n \times \mathcal{Z}^n : \log\frac{p_{\bar{\mathbf{Z}}}(\mathbf{z})}{p_{\mathbf{Z}}(\mathbf{z})} > -b\sqrt{n}\right\}.\end{aligned}$$

Using the law of total probability and the fact that $\mathcal{Q}_n^+ \cap \mathcal{B}_n \subset \mathcal{A}_n^+$, we now upper bound $\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{Q}_n^+]$ as follows:

$$\begin{aligned}\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^+\right] &= \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^+ \cap \mathcal{B}_n\right] + \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^+ \cap \mathcal{B}_n^c\right] \\ &\leqslant \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^+\right] + \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{B}_n^c]. \quad (22)\end{aligned}$$

We first establish a bound on $\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{B}_n^c]$.

$$\begin{aligned}\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{B}_n^c] &= \frac{1}{b\sqrt{n}}\sum_{\mathbf{z}\in\mathcal{Z}^n} b\sqrt{n}\, p_{\bar{\mathbf{Z}}}(\mathbf{z})\mathbb{1}\left\{\log\frac{p_{\bar{\mathbf{Z}}}(\mathbf{z})}{p_{\mathbf{Z}}(\mathbf{z})} \geqslant b\sqrt{n}\right\} \\ &\leqslant \frac{1}{b\sqrt{n}}\sum_{\mathbf{z}\in\mathcal{Z}^n} p_{\bar{\mathbf{Z}}}(\mathbf{z})\log\frac{p_{\bar{\mathbf{Z}}}(\mathbf{z})}{p_{\mathbf{Z}}(\mathbf{z})} \\ &= \frac{1}{b\sqrt{n}}\mathbb{D}(p_{\bar{\mathbf{Z}}}\|p_{\mathbf{Z}}). \quad (23)\end{aligned}$$

We define $\mu_Z \triangleq \min_{z\in\mathcal{Z}:p_Z(z)>0} p_Z(z)$ and we upper bound the divergence as follows:

$$\begin{aligned}\mathbb{D}(p_{\bar{\mathbf{Z}}}\|p_{\mathbf{Z}}) &= -\mathbb{H}(\bar{\mathbf{Z}}) + \sum_{\mathbf{z}\in\mathcal{Z}^n} p_{\bar{\mathbf{Z}}}(\mathbf{z})\log\frac{1}{p_{\mathbf{Z}}(\mathbf{z})} \\ &= \mathbb{H}(\mathbf{Z}) - \mathbb{H}(\bar{\mathbf{Z}}) + \sum_{\mathbf{z}\in\mathcal{Z}^n}(p_{\bar{\mathbf{Z}}}(\mathbf{z}) - p_{\mathbf{Z}}(\mathbf{z}))\log\frac{1}{p_{\mathbf{Z}}(\mathbf{z})} \\ &\leqslant \left|\mathbb{H}(\mathbf{Z}) - \mathbb{H}(\bar{\mathbf{Z}})\right| + n\mathbb{V}(p_{\bar{\mathbf{Z}}}, p_{\mathbf{Z}})\log\frac{1}{\mu_Z} \\ &\overset{(a)}{\leqslant} \mathbb{V}(p_{\bar{\mathbf{Z}}}, p_{\mathbf{Z}})\log\frac{|\mathcal{Z}|^n}{\mathbb{V}(p_{\bar{\mathbf{Z}}}, p_{\mathbf{Z}})} + n\mathbb{V}(p_{\bar{\mathbf{Z}}}, p_{\mathbf{Z}})\log\frac{1}{\mu_Z} \\ &\overset{(b)}{\leqslant} \left(\log|\mathcal{Z}| + \beta + \log\frac{1}{\mu_Z}\right) n2^{-\beta n} \quad (24)\end{aligned}$$

where (a) follows from [36, Lemma 2.7] and (b) follows from the fact that $x \mapsto x\log\frac{|\mathcal{Z}|^n}{x}$ is monotonously increasing for x small enough.

To upper bound $\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{A}_n^+]$, recall that the eavesdropper's channel is symmetric; hence, there exists a partition $\{\mathcal{Z}_i\}_{i\in[\![1,k]\!]}$ of $\mathcal{Z}$ such that:

1) $\forall x, \tilde{x} \in \mathcal{X}$, there exists a permutation $\pi_{x\tilde{x}} : \mathcal{Z} \to \mathcal{Z}$ that satisfies

$$\forall i \in [\![1,k]\!]\ \pi_{x\tilde{x}}(\mathcal{Z}_i) = \mathcal{Z}_i \qquad \forall z \in \mathcal{Z}\ W_{Z|X}(z|x) = W_{Z|X}(\pi_{x\tilde{x}}(z)|\tilde{x}).$$

2) The output distribution p_Z correponding to a uniform input distribution is locally uniform, i.e.,

$$\forall i \in [\![1,k]\!], \forall z, z' \in \mathcal{Z}_i\ p_Z(z) = p_Z(z').$$

Consequently, upon defining $b_n \triangleq 2b\sqrt{n} + n(C_e - \epsilon_n)$ and for any $\tilde{x} \in \mathcal{X}$, we can rewrite $\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{A}_n^+]$ as shown at the bottom of the next page, where (a) follows because the eavesdropper's channel is symmetric, (b) follows because the functions $\pi_{x_i\tilde{x}}$ are permutations, and (c) follows by defining the i.i.d. random variables $\tilde{Z}_i$ as the eavesdropper's channel output when the channel input is the symbol $\tilde{x}$. Note that the random variables $\log\frac{W_{Z|X}(\tilde{Z}_i|\tilde{x})}{p_Z(\tilde{Z}_i)}$ are also i.i.d., with mean C_e since the channel is symmetric, variance $\sigma > 0$, and third moment $\rho < \infty$; therefore,

$$\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^+\right] = \mathbb{P}\left[\frac{1}{\sqrt{n}\sigma}\sum_{i=1}^n\left(\log\frac{W_{Z|X}(\tilde{Z}_i|\tilde{x})}{p_Z(\tilde{Z}_i)} - C_e\right) \leqslant \frac{2b}{\sigma} - \frac{\sqrt{n}\epsilon_n}{\sigma}\right].$$

From the Berry–Esseen Theorem [51], there exists a universal constant $c > 0$ such that

$$\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^+\right] \leqslant \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{2b}{\sigma}-\frac{\sqrt{n}\epsilon_n}{\sigma}} e^{-\frac{x^2}{2}}\mathrm{d}x + \frac{c}{\sqrt{n}}\frac{\rho}{\sigma^3}. \quad (25)$$

Similarly, using the law of total probability, the fact that $\mathcal{A}_n^- \cap \mathcal{D}_n \subset \mathcal{Q}_n^- \cap \mathcal{D}_n$, and the inclusion–exclusion principle, we lower bound $\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{Q}_n^-]$ as follows:

$$\begin{aligned}\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^-\right] &= \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^- \cap \mathcal{D}_n\right] + \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^- \cap \mathcal{D}_n^c\right] \\ &\geqslant \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^- \cap \mathcal{D}_n\right] \\ &= \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^-\right] + \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{D}_n] - \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^- \cup \mathcal{D}_n\right] \\ &\geqslant \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^-\right] + \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{D}_n] - 1 \\ &\geqslant \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^-\right] - \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{D}_n^c]. \end{aligned} \tag{26}$$

Note that,

$$\begin{aligned}\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{D}_n^c] &= \sum_{\mathbf{z}\in\mathcal{Z}^n} p_{\bar{\mathbf{Z}}}(\mathbf{z}) \mathbb{1}\left\{\log\frac{p_{\bar{\mathbf{Z}}}(\mathbf{z})}{p_{\mathbf{Z}}(\mathbf{z})} \leqslant -b\sqrt{n}\right\} \\ &\leqslant 2^{-b\sqrt{n}} \sum_{\mathbf{z}\in\mathcal{Z}^n} p_{\mathbf{Z}}(\mathbf{z}) \\ &\leqslant 2^{-b\sqrt{n}} \end{aligned} \tag{27}$$

and, following the reasoning leading to (25),

$$\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^-\right] \geqslant \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{2b}{\sigma}-\frac{\sqrt{n}\epsilon_n}{\sigma}} e^{-\frac{x^2}{2}}\mathrm{d}x - \frac{c}{\sqrt{n}}\frac{\rho}{\sigma^3}. \tag{28}$$

Combining (22)–(28), we obtain

$$\begin{aligned}\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}[\mathcal{A}_n] &= \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^+\right] - \mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{Q}_n^-\right] \\ &\leqslant \frac{1}{\sqrt{2\pi}} \int_{-\frac{2b}{\sigma}-\frac{\sqrt{n}\epsilon_n}{\sigma}}^{\frac{2b}{\sigma}-\frac{\sqrt{n}\epsilon_n}{\sigma}} e^{-\frac{x^2}{2}}\mathrm{d}x + \frac{2c}{\sqrt{n}}\frac{\rho}{\sigma^3} \\ &\quad + \frac{\sqrt{n}}{b}\left(\log|\mathcal{Z}| + \beta + \log\frac{1}{\mu_Z}\right)2^{-\beta n} + 2^{-b\sqrt{n}} \\ &\leqslant \frac{4b}{\sigma\sqrt{2\pi}} + \frac{2c}{\sqrt{n}}\frac{\rho}{\sigma^3} \\ &\quad + \frac{\sqrt{n}}{b}\left(\log|\mathcal{Z}| + \beta + \log\frac{1}{\mu_Z}\right)2^{-\beta n} + 2^{-b\sqrt{n}}. \end{aligned} \tag{29}$$

Combining (29) with (21), and using the assumption $\lim_{n\to\infty}\mathbb{P}_e^*(\mathcal{C}_n) = 0$ from (20), we have

$$\forall b > 0 \quad \lim_{n\to\infty} \mathbb{S}_2(\mathcal{C}_n) \geqslant 1 - \frac{4b}{\sigma\sqrt{2\pi}}.$$

Therefore, there exists $\eta > 0$ such that, for n large enough, $\mathbb{S}_2(\mathcal{C}_n) \geqslant \eta$. Notice that Proposition 1 immediately implies that there exists $\eta^* > 0$ such that $\lim_{n\to\infty}\mathbb{S}_1(\mathcal{C}_n) \geqslant \eta^*$.

Appendix C
Lemmas Used in the Achievability Proof of Theorem 2

The following notation is used throughout this appendix. We recall that $\mathbf{U}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ are the random variables defined by the random code generation with distribution given in (6). For any $(k, l, m) \in [\![1, M_0]\!] \times [\![1, M_1]\!] \times [\![1, M']\!]$, the random variables representing the codewords $\mathbf{u}_k$ and $\mathbf{x}_{klm}$ obtained with the random code generation are denoted by $\mathbf{U}_k$ and $\mathbf{X}_{klm}$.

The random variables that correspond to the use of a specific code $\mathcal{C}_n$ are denoted by $\bar{\mathbf{U}}, \bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\mathbf{Z}}$ with distribution given by (4). The channel outputs that correspond to the transmission of $\mathbf{u}_k$ and $\mathbf{x}_{klm}$ are denoted by $\bar{\mathbf{Y}}_{klm}$ and $\bar{\mathbf{Z}}_{klm}$, respectively.

1) Proof of Lemma 1: By symmetry of the random code construction, we have

$$\begin{aligned}&\mathbb{E}[\mathbb{P}_e(C_n)] \\ &= \sum_{k=1}^{M_0}\sum_{l=1}^{M_1}\sum_{m=1}^{M'} \frac{\mathbb{E}[\mathbb{P}_e(C_n|W_0 = k, W_1 = l, W' = m)]}{M_0 M_1 M'} \\ &= \mathbb{E}[\mathbb{P}_e(C_n|W_0 = 1, W_1 = 1, W' = 1)],\end{aligned}$$

which can be analyzed in terms of the events

$$\begin{aligned}E_1(k) &\triangleq \left\{(\bar{\mathbf{U}}_k, \bar{\mathbf{Y}}_{111}) \in \mathcal{T}_1^n | W_0 = W_1 = W' = 1\right\} \\ E_2(k) &\triangleq \left\{(\bar{\mathbf{U}}_k, \bar{\mathbf{Z}}_{111}) \in \mathcal{T}_3^n | W_0 = W_1 = W' = 1\right\} \\ E_3(k,l,m) &\triangleq \left\{(\bar{\mathbf{U}}_k, \bar{\mathbf{X}}_{klm}, \bar{\mathbf{Y}}_{111}) \in \mathcal{T}_2^n | W_0 = W_1 = W' = 1\right\}.\end{aligned}$$

It follows from standard arguments (see, for instance, [10, Ch. 3]) that $\mathbb{E}[\mathbb{P}_e(C_n)] < \epsilon$ for n large enough provided

$$\begin{aligned}\tfrac{1}{n}\log M_0 &\leqslant \operatorname*{p\text{-}liminf}_{n\to\infty} \tfrac{1}{n} I(\mathbf{U};\mathbf{Y}) - 2\gamma \\ \tfrac{1}{n}\log M_0 &\leqslant \operatorname*{p\text{-}liminf}_{n\to\infty} \tfrac{1}{n} I(\mathbf{U};\mathbf{Z}) - 2\gamma \\ \tfrac{1}{n}\log M_1 M' &\leqslant \operatorname*{p\text{-}liminf}_{n\to\infty} \tfrac{1}{n} I(\mathbf{X};\mathbf{Y}|\mathbf{U}) - 2\gamma.\end{aligned} \tag{30}$$

$$\begin{aligned}\mathbb{P}_{\bar{\mathbf{X}}\bar{\mathbf{Z}}}\left[\mathcal{A}_n^+\right] &= \sum_{\mathbf{x}\in\mathcal{C}_n}\frac{1}{|\mathcal{C}_n|}\sum_{\mathbf{z}\in\mathcal{Z}^n} W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x})\mathbb{1}\left\{\log\frac{W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x})}{p_{\mathbf{Z}}(\mathbf{z})} \leqslant b_n\right\} \\ &= \sum_{\mathbf{x}\in\mathcal{C}_n}\frac{1}{|\mathcal{C}_n|}\sum_{\mathbf{z}\in\mathcal{Z}^n}\left(\prod_{i=1}^n W_{Z|X}(z_i|x_i)\right)\mathbb{1}\left\{\sum_{i=1}^n \log\frac{W_{Z|X}(z_i|x_i)}{p_Z(z_i)} \leqslant b_n\right\} \\ &\overset{(a)}{=} \sum_{\mathbf{x}\in\mathcal{C}_n}\frac{1}{|\mathcal{C}_n|}\sum_{\mathbf{z}\in\mathcal{Z}^n}\left(\prod_{i=1}^n W_{Z|X}(\pi_{x_i\tilde{x}}(z_i)|\tilde{x})\right)\mathbb{1}\left\{\sum_{i=1}^n \log\frac{W_{Z|X}(\pi_{x_i\tilde{x}}(z_i)|\tilde{x})}{p_Z(\pi_{x_i\tilde{x}}(z_i))} \leqslant b_n\right\} \\ &\overset{(b)}{=} \sum_{\mathbf{x}\in\mathcal{C}_n}\frac{1}{|\mathcal{C}_n|}\sum_{\mathbf{z}\in\mathcal{Z}^n}\left(\prod_{i=1}^n W_{Z|X}(z_i|\tilde{x})\right)\mathbb{1}\left\{\sum_{i=1}^n \log\frac{W_{Z|X}(z_i|\tilde{x})}{p_Z(z_i)} \leqslant b_n\right\} \\ &\overset{(c)}{=} \mathbb{P}\left[\sum_{i=1}^n \log\frac{W_{Z|X}(\tilde{Z}_i|\tilde{x})}{p_Z(\tilde{Z}_i)} \leqslant b_n\right]\end{aligned}$$

2) Proof of Lemma 2: We start by developing an upper bound for $\mathbb{S}_2(\mathcal{C}_n)$ that will be simpler to analyze. First, we have

$$\begin{aligned}\mathbb{S}_2(\mathcal{C}_n) &\triangleq \mathbb{V}\big(p_{W_1\bar{\mathbf{Z}}}, p_{W_1}p_{\bar{\mathbf{Z}}}\big) \leqslant \mathbb{V}\big(p_{\bar{\mathbf{U}}W_1\bar{\mathbf{Z}}}, p_{W_1}p_{\bar{\mathbf{U}}\bar{\mathbf{Z}}}\big)\\ &= \mathbb{E}_{\bar{\mathbf{U}}W_1}\big[\mathbb{V}\big(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}W_1}, p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}}\big)\big].\end{aligned}$$

Next, we use Lemma 7 to further bound $\mathbb{S}_2(\mathcal{C}_n)$ as follows:

$$\begin{aligned}\mathbb{S}_2(\mathcal{C}_n) &\leqslant \mathbb{E}_{\bar{\mathbf{U}}W_1}\big[\mathbb{V}\big(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}W_1}, p_{\mathbf{Z}|\mathbf{U}}\big) + \mathbb{V}\big(p_{\mathbf{Z}|\mathbf{U}}, p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}}\big)\big]\\ &= \mathbb{E}_{\bar{\mathbf{U}}W_1}\big[\mathbb{V}\big(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}W_1}, p_{\mathbf{Z}|\mathbf{U}}\big)\big] + \mathbb{E}_{\bar{\mathbf{U}}}\big[\mathbb{V}\big(p_{\mathbf{Z}|\mathbf{U}}, p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}}\big)\big]\\ &\leqslant \mathbb{E}_{\bar{\mathbf{U}}W_1}\big[\mathbb{V}\big(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}W_1}, p_{\mathbf{Z}|\mathbf{U}}\big)\big]\\ &\qquad + \mathbb{E}_{\bar{\mathbf{U}}}\big[\mathbb{V}\big(p_{W_1}p_{\mathbf{Z}|\mathbf{U}}, p_{\bar{\mathbf{Z}}W_1|\bar{\mathbf{U}}}\big)\big]\\ &= 2\mathbb{E}_{\bar{\mathbf{U}}W_1}\big[\mathbb{V}\big(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}W_1}, p_{\mathbf{Z}|\mathbf{U}}\big)\big]. \qquad (31)\end{aligned}$$

Notice that the term in brackets on the right hand side is a variational distance between the following two distributions:

1) $p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{u}_k,W_1=l}(\mathbf{z}) = \sum_{m=1}^{M'}\frac{1}{M'}W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x}_{klm})$, which represents the distribution induced at the eavesdropper's channel output by the M' codewords $\{\mathbf{x}_{kli}\}_{i\in[\![1,M']\!]}$ selected with a uniform distribution;
2) $p_{\mathbf{Z}|\mathbf{U}=\mathbf{u}_k}(\mathbf{z}) = \sum_{\mathbf{x}} W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x})p_{\mathbf{X}|\mathbf{U}=\mathbf{u}_k}(\mathbf{x})$, which represents the distribution induced at the eavesdropper's channel output by an input process with distribution $p_{\mathbf{X}|\mathbf{U}=\mathbf{u}_k}(\mathbf{x})$.

Therefore, a *sufficient condition* for $\mathbb{S}_2(\mathcal{C}_n)$ to vanish is that, for every pair $(k,l)\in[\![1,M_0]\!]\times[\![1,M_1]\!]$, the variational distance between the two distributions vanishes as well. This is possible if each set of codewords $\{\mathbf{x}_{kli}\}_{i\in[\![1,M']\!]}$ approximates the same process with distribution $p_{\mathbf{Z}|\mathbf{U}=\mathbf{u}_k}(\mathbf{z})$ at the eavesdropper's output, which is exactly what the concept of channel resolvability reviewed in Section II is about. In other words, a sufficient condition to guarantee secrecy is for each subcodebook $\{\mathbf{x}_{kli}\}_{i\in[\![1,M']\!]}$ to be a "channel resolvability code."

We establish the existence of such codebooks with a random coding argument following that used in [7] and [10]. The presence of a common message makes the proof slightly more involved but the steps remain essentially the same. On taking the average over C_n for both sides of (31), we obtain

$$\mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] \leqslant 2\mathbb{E}_{\bar{\mathbf{U}}W_1}\big[\mathbb{E}_{C_n}\big[\mathbb{V}\big(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}W_1}, p_{\mathbf{Z}|\mathbf{U}}\big)\big]\big]. \qquad (32)$$

By symmetry of the random code construction, the inner expectation in (32) is the same for all values of $\bar{\mathbf{U}}=\mathbf{u}_k$ and $W_1=l$; hence, we have

$$\mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] \leqslant 2\mathbb{E}_{C_n}\big[\mathbb{V}\big(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1W_1=1}, p_{\mathbf{Z}|\mathbf{U}=\mathbf{U}_1}\big)\big]. \qquad (33)$$

Let $\tau>0$. On using [10, Lemma 6.3.1], we finally upper bound (33) as

$$\mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] \leqslant 4\frac{\tau}{\log e} + 4A_n \qquad (34)$$

with

$$A_n \triangleq \mathbb{E}_{C_n}\left[\mathbb{P}_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1W_1=1}\left[\log\frac{p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1W_1=1}(\bar{\mathbf{Z}})}{p_{\mathbf{Z}|\mathbf{U}=\mathbf{U}_1}(\bar{\mathbf{Z}})} > \tau\right]\right].$$

Note that the expectation over C_n reduces to the expectation over $\mathbf{U}_1$ and $\{\mathbf{X}_{11j}\}_{j\in[\![1,M']\!]}$. Writing A_n explicitly, we obtain (35) shown at the bottom of the page, where equality (a) follows from the definition of $p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{u}_1W_1=1,\mathcal{C}_n}(\mathbf{z})$, equality

$$\begin{aligned}A_n &= \sum_{\mathbf{u}_1\in\mathcal{U}^n} p_{\mathbf{U}}(\mathbf{u}_1)\sum_{\mathbf{x}_{111}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{111}|\mathbf{u}_1)\cdots\sum_{x_{11M'}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{11M'}|\mathbf{u}_1)\\ &\qquad\sum_{\mathbf{z}\in\mathcal{Z}^n} p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{u}_1W_1=1}(\mathbf{z})\mathbb{1}\left\{\log\frac{p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{u}_1W_1=1}(\mathbf{z})}{p_{\mathbf{Z}|\mathbf{U}}(\mathbf{z}|\mathbf{u}_1)} > \tau\right\}\\ &\overset{(a)}{=} \frac{1}{M'}\sum_{m=1}^{M'}\sum_{\mathbf{u}_1\in\mathcal{U}^n} p_{\mathbf{U}}(\mathbf{u}_1)\sum_{\mathbf{x}_{111}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{111}|\mathbf{u}_1)\cdots\sum_{x_{11M'}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{11M'}|\mathbf{u}_1)\\ &\qquad\sum_{\mathbf{z}\in\mathcal{Z}^n} W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x}_{11m})\mathbb{1}\left\{\log\frac{p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{u}_1W_1=1}(\mathbf{z})}{p_{\mathbf{Z}|\mathbf{U}}(\mathbf{z}|\mathbf{u}_1)} > \tau\right\}\\ &\overset{(b)}{=} \sum_{\mathbf{u}_1\in\mathcal{U}^n} p_{\mathbf{U}}(\mathbf{u}_1)\sum_{\mathbf{x}_{112}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{112}|\mathbf{u}_1)\cdots\sum_{x_{11M'}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{11M'}|\mathbf{u}_1)\\ &\qquad\sum_{\mathbf{x}_{111}\in\mathcal{X}^n}\sum_{\mathbf{z}\in\mathcal{Z}^n} W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x}_{111})p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{111}|\mathbf{u}_1)\mathbb{1}\left\{\log\frac{p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{u}_1W_1=1}(\mathbf{z})}{p_{\mathbf{Z}|\mathbf{U}}(\mathbf{z}|\mathbf{u}_1)} > \tau\right\}\\ &\overset{(c)}{=} \sum_{\mathbf{u}_1\in\mathcal{U}^n} p_{\mathbf{U}}(\mathbf{u}_1)\sum_{\mathbf{x}_{112}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{112}|\mathbf{u}_1)\cdots\sum_{x_{11M'}\in\mathcal{X}^n} p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{11M'}|\mathbf{u}_1)\\ &\qquad\sum_{\mathbf{x}_{111}\in\mathcal{X}^n}\sum_{\mathbf{z}\in\mathcal{Z}^n} p_{\mathbf{ZX}|\mathbf{U}}(\mathbf{z},\mathbf{x}_{111}|\mathbf{u}_1)\mathbb{1}\left\{\log\left(\frac{1}{M'}\sum_{m=1}^{M'}\frac{p_{\mathbf{Z}|\mathbf{XU}}(\mathbf{z}|\mathbf{x}_{11m}\mathbf{u}_1)}{p_{\mathbf{Z}|\mathbf{U}}(\mathbf{z}|\mathbf{u}_1)}\right) > \tau\right\} \qquad (35)\end{aligned}$$

(b) follows by remarking that all codewords are generated according to the same density $p_{\mathbf{X}|\mathbf{U}}$ and equality (c) follows by noting that

1) $W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x}_{111})p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{111}|\mathbf{u}_1) = p_{\mathbf{Z}\mathbf{X}|\mathbf{U}}(\mathbf{z},\mathbf{x}_{111}|\mathbf{u}_1)$ according to (6);
2) for any $\mathbf{u}_1$ such that $p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}_{11m}|\mathbf{u}_1) > 0$,

$$p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{u}_1 W_1=1}(\mathbf{z}) = \frac{1}{M'}\sum_{m=1}^{M'} W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x}_{11m}) = \frac{1}{M'}\sum_{m=1}^{M'} p_{\mathbf{Z}|\mathbf{X}\mathbf{U}}(\mathbf{z}|\mathbf{x}_{11m}\mathbf{u}_1).$$

By adapting the proof technique developed in [10, Ch. 6] and after some calculations, one can further bound A_n to obtain

$$\begin{aligned}\mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] \leqslant\ & 4\frac{\tau}{\log e} + 4\mathbb{P}_{\mathbf{U}\mathbf{X}\mathbf{Z}}\left[\frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U}) \geqslant \frac{\log M'}{n} + \frac{\log \rho}{n}\right] \\ & + 4\mathbb{P}_{\mathbf{U}\mathbf{X}\mathbf{Z}}\left[\frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U}) \geqslant \frac{\log M'}{n}\right] + \frac{4\cdot 2^{-n\gamma}}{\rho^2} \\ & + \frac{4}{\rho^2}\mathbb{P}_{\mathbf{U}\mathbf{X}\mathbf{Z}}\left[\frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U}) \geqslant \frac{\log M'}{n} - \gamma\right] \quad (36)\end{aligned}$$

where $\rho \triangleq \frac{2^{\tau}-1}{2}$. Therefore, $\mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] < \epsilon$ for n large enough provided

$$\frac{1}{n}\log M' \geqslant \operatorname*{p-limsup}_{n\to\infty} \frac{1}{n}I(\mathbf{X};\mathbf{Z}|\mathbf{U}) + 2\gamma. \quad (37)$$

Appendix D
Lemma Used in the Converse Proof of Theorem 2

To prove Lemma 3, note that, with probability 1,

$$\begin{aligned}\tfrac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}) &= \tfrac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}\bar{\mathbf{U}}) - \tfrac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{U}}|\bar{\mathbf{Z}}) \\ &= \tfrac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) - \tfrac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{U}}|\bar{\mathbf{Z}}) \\ &= \tfrac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) - \tfrac{1}{n}H(\bar{\mathbf{U}}|\bar{\mathbf{Z}}) + \tfrac{1}{n}H(\bar{\mathbf{U}}|\bar{\mathbf{W}}\bar{\mathbf{Z}}),\end{aligned}$$

where the second equality follows from the independence of $\bar{\mathbf{W}}$ and $\bar{\mathbf{U}}$. Consequently,

$$\begin{aligned}\lim_{n\to\infty}\mathbb{S}_6(\mathcal{C}_n) &= \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}) \\ &\geqslant \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) - \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}H(\bar{\mathbf{U}}|\bar{\mathbf{Z}}) \\ &\quad + \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}H(\bar{\mathbf{U}}|\bar{\mathbf{W}}\bar{\mathbf{Z}}).\end{aligned}$$

Since $\lim_{n\to\infty}\mathbb{P}_e(\mathcal{C}_n)=0$, note that $\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}H(\bar{\mathbf{U}}|\bar{\mathbf{W}}\bar{\mathbf{Z}})=0$ and $\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}H(\bar{\mathbf{U}}|\bar{\mathbf{Z}}) = 0$ by the Verdú–Han lemma. As $\lim_{n\to\infty}\mathbb{S}_6(\mathcal{C}_n) = 0$, we finally obtain

$$\operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) = 0.$$

Note that, with probability 1,

$$\begin{aligned}H(\bar{\mathbf{W}}) &= H(\bar{\mathbf{W}}) - H(\bar{\mathbf{W}}|\bar{\mathbf{Y}}\bar{\mathbf{U}}) + H(\bar{\mathbf{W}}|\bar{\mathbf{Y}}\bar{\mathbf{U}}) \\ &= I(\bar{\mathbf{W}};\bar{\mathbf{Y}}|\bar{\mathbf{U}}) - I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) + I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) \\ &\quad + H(\bar{\mathbf{W}}|\bar{\mathbf{Y}}\bar{\mathbf{U}}).\end{aligned}$$

Hence,

$$\begin{aligned}R_1 &\leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}H(\bar{\mathbf{W}}) \\ &\leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Y}}|\bar{\mathbf{U}}) - \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) \\ &\quad + \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) + \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}H(\bar{\mathbf{W}}|\mathbf{Y}\bar{\mathbf{U}}).\end{aligned}$$

As seen above, $\operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\bar{\mathbf{W}};\bar{\mathbf{Z}}|\bar{\mathbf{U}}) = 0$ and, since $\lim_{n\to\infty}\mathbb{P}_e(\mathcal{C}_n) = 0$, we have $\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}H(\bar{\mathbf{W}}|\bar{\mathbf{Y}}\bar{\mathbf{U}}) = 0$. Therefore,

$$R_1 \leqslant \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}I(\mathbf{W};\bar{\mathbf{Y}}|\bar{\mathbf{U}}) - \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}I(\mathbf{W};\bar{\mathbf{Z}}|\bar{\mathbf{U}}).$$

Finally, with probability 1,

$$H(\bar{\mathbf{U}}) = I(\bar{\mathbf{U}};\bar{\mathbf{Y}}) + H(\bar{\mathbf{Y}}|\bar{\mathbf{U}}) = I(\bar{\mathbf{U}};\bar{\mathbf{Z}}) + H(\bar{\mathbf{Z}}|\bar{\mathbf{U}})$$

from which conclude after a similar reasoning that

$$R_0 \leqslant \min\left(\operatorname*{p-liminf}_{n\to\infty} I(\bar{\mathbf{U}};\bar{\mathbf{Y}}), \operatorname*{p-liminf}_{n\to\infty} I(\bar{\mathbf{U}};\bar{\mathbf{Z}})\right).$$

Appendix E
Proof of Theorem 3

We prove Theorem 3 with small modifications of the proof of Theorem 2. Specifically, we establish secrecy for $\mathbb{S}_1$ by showing that there exist sequences of codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ for which $\mathbb{S}_2(\mathcal{C}_n)$ decreases exponentially fast with n and by using [5, Lemma 1] to obtain an upper bound for $\mathbb{S}_1(\mathcal{C}_n)$. We handle the power constraint by using an appropriate distribution during the random code generation process as in [10, Sec. 3.2]. We note that a similar technique has been used by He and Yener in [52].

Let $\gamma, \delta, \epsilon > 0$. Let $\mathcal{U}$ be an arbitrary discrete alphabet and fix a distribution $p_{\tilde{U}}$ on $\mathcal{U}$. Fix a conditional distribution $p_{\tilde{X}|\tilde{U}}$ on $\mathcal{X}\times\mathcal{U}$ such that $\mathbb{E}\left[c(\tilde{X})\right] \leqslant P - \delta$. Let $\tilde{\mathbf{U}}, \tilde{\mathbf{X}}, \tilde{\mathbf{Z}}$ be the random variables with joint distribution

$$p_{\tilde{\mathbf{Z}}\tilde{\mathbf{X}}\tilde{\mathbf{U}}}(\mathbf{z},\mathbf{x},\mathbf{u}) = \prod_{i=1}^{n} W_{Z|X}(z_i|x_i)p_{\tilde{X}|\tilde{U}}(x_i|u_i)p_{\tilde{U}}(u_i).$$

We assume that $\tilde{\mathbf{U}}, \tilde{\mathbf{X}}, \tilde{\mathbf{Z}}$ are such that the integrals defining the moment generating functions of $c(\tilde{X})$ and $I(\tilde{X};\tilde{Z}|\tilde{U})$ converge uniformly in a neighborhood of 0 and are differentiable at 0.

Define the set $\mathcal{P}_n$ as

$$\mathcal{P}_n \triangleq \left\{\mathbf{x}\in\mathcal{X}^n : \frac{1}{n}\sum_{i=1}^{n} c(x_i) \leqslant P\right\}.$$

Lemma 6 shows that there exists $\alpha_\delta > 0$ such that $\mathbb{P}\left[\tilde{\mathbf{X}} \in \mathcal{P}_n\right] \geqslant 1 - 2^{-\alpha_\delta n}$. In the sequel, we define $\gamma_n \triangleq 1 - 2^{-n\frac{\alpha_\delta}{2}}$. Define the set $\mathcal{G}_n \subset \mathcal{U}^n$ as follows:

$$\mathcal{G}_n \triangleq \left\{\mathbf{u} : \mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}}\left[\tilde{\mathbf{X}} \notin \mathcal{P}_n|\tilde{\mathbf{U}} = \mathbf{u}\right] < 2^{-n\frac{\alpha_\delta}{2}}\right\}.$$

Upon using Markov's inequality, we obtain

$$\begin{aligned}\mathbb{P}_{\tilde{\mathbf{U}}}\left[\tilde{\mathbf{U}} \notin \mathcal{G}_n\right] &= \mathbb{P}_{\tilde{\mathbf{U}}}\left[\mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}}\left[\tilde{\mathbf{X}} \notin \mathcal{P}_n|\tilde{\mathbf{U}}\right] \geqslant 2^{-n\frac{\alpha_\delta}{2}}\right] \\ &\leqslant \mathbb{E}_{\tilde{\mathbf{U}}}\left[\mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}}\left[\tilde{\mathbf{X}} \notin \mathcal{P}_n|\tilde{\mathbf{U}}\right]\right]2^{n\frac{\alpha_\delta}{2}} \\ &= \mathbb{P}_{\tilde{\mathbf{X}}}\left[\tilde{\mathbf{X}} \notin \mathcal{P}_n\right]2^{n\frac{\alpha_\delta}{2}} \\ &\leqslant 2^{-n(\alpha_\delta - \frac{\alpha_\delta}{2})} \\ &= 1 - \gamma_n. \end{aligned} \tag{38}$$

Now, we define the random variables $\mathbf{U}, \mathbf{X}, \mathbf{Z}$ as follows. First,

$$\forall \mathbf{u} \in \mathcal{U}^n \quad p_{\mathbf{U}}(\mathbf{u}) = \begin{cases} \frac{1}{\mathbb{P}_{\tilde{\mathbf{U}}}\left[\tilde{\mathbf{U}} \in \mathcal{G}_n\right]} p_{\tilde{\mathbf{U}}}(\mathbf{u}) \text{ if } \mathbf{u} \in \mathcal{G}_n \\ 0 \text{ else.} \end{cases}$$

From (38), we have

$$\forall \mathbf{u} \in \mathcal{U}^n \quad p_{\mathbf{U}}(\mathbf{u}) \leqslant \frac{p_{\tilde{\mathbf{U}}}(\mathbf{u})}{\gamma_n}. \tag{39}$$

Next, $\forall (\mathbf{x}, \mathbf{u}) \in \mathcal{X}^n \times \mathcal{G}_n$

$$p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}|\mathbf{u}) = \begin{cases} \frac{1}{\mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}}\left[\tilde{\mathbf{X}} \in \mathcal{P}_n|\tilde{\mathbf{U}}=\mathbf{u}\right]} p_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}}(\mathbf{x}|\mathbf{u}), \text{ if } \mathbf{x} \in \mathcal{P}_n \\ 0, \text{ else.} \end{cases}$$

By construction, we have

$$\forall (\mathbf{x}, \mathbf{u}) \in \mathcal{X}^n \times \mathcal{G}_n \quad p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}|\mathbf{u}) \leqslant \frac{p_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}}(\mathbf{x}|\mathbf{u})}{\gamma_n}. \tag{40}$$

Finally, $\forall (\mathbf{z}, \mathbf{x}, \mathbf{u}) \in \mathcal{Z}^n \times \mathcal{X}^n \times \mathcal{G}_n$

$$p_{\mathbf{ZXU}}(\mathbf{z}, \mathbf{x}, \mathbf{u}) = W_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x}) p_{\mathbf{X}|\mathbf{U}}(\mathbf{x}|\mathbf{u}) p_{\mathbf{U}}(\mathbf{u}). \tag{41}$$

We repeat the random coding argument in the proof of Theorem 2 using the distribution $p_{\mathbf{XU}}$ defined by (41) and with the following lemmas.

Lemma 12 (Reliability Conditions):

$$\begin{aligned}&\text{If } R_0 \leqslant \min\left(\mathbb{I}\left(\tilde{U};\tilde{Y}\right) - 2\gamma, \mathbb{I}\left(\tilde{U};\tilde{Z}\right) - 2\gamma\right) \\ &\text{and } R_1 + R_1' \leqslant \mathbb{I}\left(\tilde{X};\tilde{Y}|\tilde{U}\right) - 2\gamma,\end{aligned}$$

then $\lim_{n\to\infty} \mathbb{E}[\mathbb{P}_e(C_n)] \leqslant \epsilon$.

Proof: Following [10, Proof of Theorem 3.6.2], one can show that

$$\begin{aligned}\text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{U};\mathbf{Y}) &\geqslant \mathbb{I}\left(\tilde{U};\tilde{Y}\right), \\ \text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{U};\mathbf{Z}) &\geqslant \mathbb{I}\left(\tilde{U};\tilde{Z}\right) \\ \text{and} \quad \text{p-liminf}_{n\to\infty} \frac{1}{n} I(\mathbf{X};\mathbf{Y}|\mathbf{U}) &\geqslant \mathbb{I}\left(\tilde{X};\tilde{Y}|\tilde{U}\right).\end{aligned}$$

Hence, the result follows directly from Lemma 1. ■

Lemma 13 (Secrecy From Channel Resolvability Conditions): There exists $\alpha_{\delta,\gamma} > 0$, such that

$$\text{If } R_1' \geqslant \mathbb{I}\left(\tilde{X};\tilde{Z}|\tilde{U}\right) + 2\gamma, \text{ then } \lim_{n\to\infty} \mathbb{E}[\mathbb{S}_2(C_n)] \leqslant 2^{-\alpha_{\delta,\gamma}}.$$

Proof: Note that (31) still holds. Upon using Lemma 7, we obtain

$$\begin{aligned}\mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] &\leqslant 2\mathbb{E}_{C_n}\left[\mathbb{V}\left(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1, W_1=1, C_n}, p_{\mathbf{Z}|\mathbf{U}=\mathbf{U}_1}\right)\right] \\ &\leqslant 2\mathbb{E}_{C_n}\left[\mathbb{V}\left(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1, W_1=1, C_n}, p_{\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}=\mathbf{U}_1}\right)\right] \\ &\quad + 2\mathbb{E}_{C_n}\left[\mathbb{V}\left(p_{\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}=\mathbf{U}_1}, p_{\mathbf{Z}|\mathbf{U}=\mathbf{U}_1}\right)\right].\end{aligned} \tag{42}$$

$$\begin{aligned}\mathbb{V}\left(p_{\bar{\mathbf{Z}}|\tilde{\mathbf{U}}=\mathbf{u}_1}, p_{\mathbf{Z}|\mathbf{U}=\mathbf{u}_1}\right) &\overset{(a)}{\leqslant} \mathbb{V}\left(p_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}_1}, p_{\mathbf{X}|\mathbf{U}=\mathbf{u}_1}\right) \\ &= 2\sup_{\mathcal{A}\subseteq\mathcal{X}^n} \left|\mathbb{P}_{\mathbf{X}|\mathbf{U}=\mathbf{u}_1}[\mathcal{A}] - \mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}_1}[\mathcal{A}]\right| \\ &\leqslant \sup_{\mathcal{A}\subseteq\mathcal{X}^n} \sum_{\mathcal{B}\in\{\mathcal{A},\mathcal{A}^c\}} \left(\left|\mathbb{P}_{\mathbf{X}|\mathbf{U}=\mathbf{u}_1}[\mathcal{B}] - \mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}_1}[\mathcal{B}]\right|\right) \\ &\overset{(b)}{\leqslant} \sup_{\mathcal{A}\subseteq\mathcal{X}^n} \sum_{\mathcal{B}\in\{\mathcal{A},\mathcal{A}^c\}} \left(\left|\mathbb{P}_{\mathbf{X}|\mathbf{U}=\mathbf{u}_1}[\mathcal{B}\cap\mathcal{P}_n] - \mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}_1}[\mathcal{B}\cap\mathcal{P}_n^c] - \mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}_1}[\mathcal{B}\cap\mathcal{P}_n]\right|\right) \\ &\overset{(c)}{\leqslant} \sup_{\mathcal{A}\subseteq\mathcal{X}^n} \sum_{\mathcal{B}\in\{\mathcal{A},\mathcal{A}^c\}} \left(\mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}_1}[\mathcal{B}\cap\mathcal{P}_n]\left(\frac{1}{\gamma_n} - 1\right) + \mathbb{P}_{\tilde{\mathbf{X}}|\tilde{\mathbf{U}}=\mathbf{u}_1}[\mathcal{B}\cap\mathcal{P}_n^c]\right) \\ &\leqslant (\frac{1}{\gamma_n} - 1) + (1 - \gamma_n)\end{aligned} \tag{43}$$

First, we bound the second term on the right-hand side of (42). For all $\mathbf{u}_1 \in \mathcal{G}_n$, we obtain the bound shown in (43) at the bottom of the previous page, where (a) follows from Lemma 8, (b) follows because $\mathbb{P}_{\mathbf{X}|\mathbf{U}=\mathbf{u}_1}[\mathcal{B} \cap \mathcal{P}_n^c] = 0$ by (41), and (c) follows from the bound in (40); therefore, for n large enough, there exists $\beta_\delta > 0$, such that

$$\mathbb{E}_{C_n}\left[\mathbb{V}\left(p_{\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}=\mathbf{U}_1}, p_{\mathbf{Z}|\mathbf{U}=\mathbf{U}_1}\right)\right] \leqslant 2^{-\beta_\delta n}. \tag{44}$$

We now bound the first term on the right-hand side of (42). Applying [10, Lemma 6.3.1], we obtain

$$\begin{aligned} &2\mathbb{E}_{C_n}\left[\mathbb{V}\left(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1 W_1=1}, p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1}\right)\right] \leqslant 4\frac{\tau}{\log e} \\ &+4\mathbb{E}_{C_n}\left[\mathbb{P}_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1 W_1=1}\left[\log\frac{p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1 W_1=1}(\bar{\mathbf{Z}})}{p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1}(\bar{\mathbf{Z}})} > \tau\right]\right]. \end{aligned} \tag{45}$$

Note that (45) is similar to (34), and the only difference is the presence of $p_{\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}}$ instead of $p_{\mathbf{Z}|\mathbf{U}}$ in the denominator; using the definition of $p_{\mathbf{ZXU}}$ in (41), the bounds in (39) and (40), and repeating the steps leading from (34) to (36), one obtains after some calculations

$$\begin{aligned} &2\mathbb{E}_{C_n}\left[\mathbb{V}\left(p_{\bar{\mathbf{Z}}|\bar{\mathbf{U}}=\mathbf{U}_1 W_1=1}, p_{\tilde{Z}_1^n|\tilde{\mathbf{U}}=\mathbf{U}_1}\right)\right] \\ &\leqslant 4\frac{\tau}{\log e} + \frac{4}{\gamma_n^2}\mathbb{P}_{\tilde{\mathbf{U}}\tilde{\mathbf{X}}\tilde{\mathbf{Z}}}\left[\frac{1}{n}I\left(\tilde{\mathbf{X}};\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}\right) \geqslant \frac{\log M'}{n} + \frac{\log\rho}{n}\right] \\ &+\frac{4}{\gamma_n^3}\mathbb{P}_{\tilde{\mathbf{U}}\tilde{\mathbf{X}}\tilde{\mathbf{Z}}}\left[\frac{1}{n}I\left(\tilde{\mathbf{X}};\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}\right) \geqslant \frac{\log M'}{n}\right] + \frac{4\cdot 2^{-n\gamma}}{\gamma_n(\gamma_n\rho+\gamma_n-1)^2} \\ &+\frac{4}{\gamma_n(\gamma_n\rho+\gamma_n-1)^2}\mathbb{P}_{\tilde{\mathbf{U}}\tilde{\mathbf{X}}\tilde{\mathbf{Z}}}\left[\frac{1}{n}I\left(\tilde{\mathbf{X}};\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}\right) \geqslant \frac{\log M'}{n} - \gamma\right]. \end{aligned} \tag{46}$$

If $\frac{1}{n}\log M' \geqslant \mathbb{I}\left(\tilde{X};\tilde{Z}|\tilde{U}\right)+2\gamma$, then Lemma 6 guarantees there exists $\alpha_\gamma > 0$ such that

$$\mathbb{P}_{\tilde{\mathbf{U}}\tilde{\mathbf{X}}\tilde{\mathbf{Z}}}\left[\frac{1}{n}I\left(\tilde{\mathbf{X}};\tilde{\mathbf{Z}}|\tilde{\mathbf{U}}\right) \geqslant \frac{\log M'}{n} - \gamma\right] \leqslant 2^{-\alpha_\gamma n}. \tag{47}$$

Set $\tau = 2^{-\eta n}$ for some η such that $0 < 2\eta < \min(\gamma, \alpha_\gamma)$; note that $\rho = \frac{\ln 2}{2}2^{-\eta n} + o(2^{-\eta n})$. Therefore, for n large enough,

$$\frac{1}{n}\log\rho \geqslant -\gamma, \quad \frac{1}{\gamma_n(\gamma_n\rho+\gamma_n-1)} \leqslant 2\cdot 2^{\eta n}, \quad \frac{1}{\gamma_n^3} \leqslant 2. \tag{48}$$

Consequently, combining (42), (44), (46)–(48), we obtain for n large enough,

$$\begin{aligned} \mathbb{E}_{C_n}[\mathbb{S}_2(C_n)] \leqslant\ & 4\cdot\frac{2^{-\eta n}}{\log e} + 8\cdot 2^{-\alpha_\gamma n} + 8\cdot 2^{-\alpha_\gamma n} \\ &+16\cdot 2^{-(\gamma-2\eta)n} + 16\cdot 2^{-(\alpha_\gamma-2\eta)n} + 2\cdot 2^{-\beta_\delta n}. \end{aligned}$$

Therefore, for n large enough, there exists $\alpha_{\gamma,\delta} > 0$ such that $\mathbb{E}[\mathbb{S}_2(C_n)] \leqslant 2^{-\alpha_{\gamma,\delta} n}$. ■

Using Markov's inequality and for n sufficiently large, we conclude that if

$$\begin{aligned} R_0 &\leqslant \min\left(\mathbb{I}\left(\tilde{U};\tilde{Y}\right) - 2\gamma, \mathbb{I}\left(\tilde{U};\tilde{Z}\right) - 2\gamma\right) \\ R_1 &\leqslant \mathbb{I}\left(\tilde{X};\tilde{Y}|\tilde{U}\right) - \mathbb{I}\left(\tilde{X};\tilde{Z}|\tilde{U}\right) - 4\gamma, \end{aligned}$$

then there exists a specific code $\mathcal{C}_n$ such that $\mathbb{P}_e(\mathcal{C}_n) \leqslant 2\epsilon$ and $\mathbb{S}_2(\mathcal{C}_n) \leqslant 2^{-\frac{\alpha_{\gamma,\delta}}{2}n}$. Using [5, Lemma 1] with n large enough, we obtain $\mathbb{S}_1(\mathcal{C}_n) \leqslant 2^{-\beta_{\gamma,\delta} n}$ for some $\beta_{\gamma,\delta} > 0$.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1367, Oct. 1975.

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[3] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Information-Theoretic Security*, ser. Foundations and Trends in Communications and Information Theory. Delft, The Netherlands: Now Publishers, 2009, vol. 5, no. 1–5.

[4] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, Oct. 2011.

[5] I. Csiszár, "Almost independence and secrecy capacity," *Probl. Inf. Transmiss.*, vol. 32, no. 1, pp. 40–47, Jan.–Mar. 1996.

[6] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, Apr. 2006.

[7] T. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[8] T. S. Han and S. Verdu, "The resolvability and the capacity of AWGN channels are equal," presented at the IEEE Int. Symp. Inf. Theory, Trondheim, Norway, 1994.

[9] G. D. Forney, Jr., "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proc. 41st Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2003, pp. 430–439.

[10] T. S. Han, *Information-Spectrum Methods in Information Theory*. New York, NY, USA: Springer-Verlag, 2002.

[11] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.

[12] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[13] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[14] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. I. Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[15] U. Maurer, "The strong secret key rate of discrete random triples," in *Communications and Cryptography: Two Sides of One Tapestry*. Norwell, MA, USA: Kluwer, 1994, pp. 271–285.

[16] U. M. Maurer and S. Wolf, B. Preneel, Ed., "Information-theoretic key agreement: From weak to strong secrecy for free," in *Proc. 19th Int. Conf. Theory and Appl. Cryptographic Tech.*, 2000, pp. 351–368.

[17] C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.

[18] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 44–55, Jan. 2005.

[19] H. Koga, "Coding theorems on Shannon's cipher system with a general source," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, p. 158.

[20] H. Koga and N. Sato, "On an upper bound of the secrecy capacity for a general wiretap channel," in *Proc. Int. Symp. Inf. Theory*, Adelaide, Australia, Sep. 2005, pp. 1641–1645.

[21] D. Kobayashi, H. Yamamoto, and T. Ogawa, "How to attain the ordinary channel capacity securely in wiretap channels," in *Proc. IEEE Inf. Theory Workshop Theory Pract. Inf.-Theoret. Security*, Awaji Island, Japan, Oct. 2005, pp. 13–18.

[22] M. Bloch and J. N. Laneman, "On the secrecy capacity of arbitrary wiretap channels," in *Proc. 46th Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2008, pp. 818–825.

[23] M. R. Bloch, "Achieving secrecy: Capacity vs. resolvability," in *Proc. IEEE Int. Symp. Inf. Theory*, Saint Petersburg, Russia, Aug. 2011, pp. 632–636.

[24] M. Bloch, "Channel intrinsic randomness," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 2607–2611.

[25] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. San Francisco, CA, USA: Holden Day, 1964.
[26] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.
[27] S. Vembu and S. Verdú, "Generating random bits from an arbitrary source: Fundamental limits," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1322–1332, Sep. 1995.
[28] J. Barros and M. Bloch, "Strong secrecy for wireless channels," in *Information Theoretic Security*, ser. Lecture Notes in Computer Science. Calgary, AB, Canada: Springer, Aug. 2008, pp. 40–53, to be published.
[29] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3989–4001, Jun. 2011.
[30] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
[31] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.
[32] R. G. Gallager, *Information Theory and Reliable Communications*. New York, NY, USA: Wiley, 1968.
[33] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.
[34] S. Watanabe, R. Matsumoto, and T. Uyematsu, "Strongly secure privacy amplification cannot be obtained by encoder of Slepian-Wolf code," *IEICE Trans. Fund. Electron., Commun. Comput. Sci.*, vol. E93-A, no. 9, pp. 1650–1659, Sep. 2010.
[35] R. Matsumoto and M. Hayashi, Strong Security and Separated Code Constructions for the Broadcast Channel with Confidential Messages Oct. 2010 [Online]. Available: arXiv:1010.0743
[36] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[37] M. Hayashi, Tight Exponential Evaluation for Universal Composability with Privacy Amplification and its Applications Feb. 2012 [Online]. Available: arXiv preprint:10.10.1358
[38] M. Bloch and J. N. Laneman, "Information-spectrum methods for information-theoretic security," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, USA, Feb. 2009, pp. 23–28.
[39] M. R. Bloch and J. N. Laneman, "Exploiting partial channel state information for secrecy over wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1840–1849, Sep. 2013.
[40] I. Bjelakovic, H. Boche, and J. Sommerfeld, "Capacity results for compound wiretap channels," in *Proc. IEEE Inf. Theory Workshop*, Paraty, Brazil, Oct. 2011, pp. 60–64.
[41] Y. Liang, G. Kramer, H. V. Poor, and S. S. (Shitz), "Compound wiretap channels," *EURASIP J. Wireless Commun. Netw.*, vol. 142374, pp. 1–12, 2009.
[42] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 116–120.
[43] B. P. Dunn, M. Bloch, and J. N. Laneman, "Secure bits through queues," in *Proc. IEEE Inf. Theory Workshop Netw. Inf. Theory*, Volos, Greece, Jun. 2009, pp. 37–41.
[44] Y. Sankarasubramaniam, A. Thangaraj, and K. Viswanathan, "Finite-state wiretap channels: Secrecy under memory constraints," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2009, pp. 115–119.
[45] A. J. Pierrot and M. R. Bloch, "Strongly secure communications over the two-way wiretap channel," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 595–605, Sep. 2011.
[46] A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong secrecy on the binary erasure wiretap channel using large-girth LDPC codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 585–594, Sep. 2011.
[47] H. Yamamoto, "Rate-distortion theory for the Shannon cipher systems," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 827–835, May 1997.
[48] L. H. Ozarow and A. D. Wyner, "Wire tap channel II," *AT&T Bell Lab. Tech. J.*, vol. 63, no. 10, pp. 2135–2157, Dec. 1984.
[49] P. W. Cuff, "Communication in networks for coordinating behavior," Ph.D. dissertation, Princeton Univ., Princeton, NJ, USA, Jul. 2009.
[50] M. Hayashi, "Second-order asymptotics in fixed-length source coding and intrinsic randomness," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4619–4637, Oct. 2008.
[51] A. N. Shiryaev, *Probability*, 2nd ed. New York, NY, USA: Springer-Verlag, 1995.
[52] X. He and A. Yener, "MIMO wiretap channels with arbitrarily varying eavesdropper channel states," *IEEE Trans. Inf. Theory*, Sep. 2010, submitted for publication.

Matthieu R. Bloch (M'08) received the Engineering degree from Supélec, Gif-sur-Yvette, France, the M.S. degree in Electrical Engineering from the Georgia Institute of Technology, Atlanta, in 2003, the Ph.D. degree in Engineering Science from the Université de Franche-Comté, Besançon, France, in 2006, and the Ph.D. degree in Electrical Engineering from the Georgia Institute of Technology in 2008. In 2008–2009, he was a postdoctoral research associate at the University of Notre Dame, South Bend, IN, USA. Since July 2009, Dr. Bloch has been on the faculty of the School of Electrical and Computer Engineering at the Georgia Institute of Technology, where he is currently an Assistant Professor. His research interests are in the areas of information theory, error-control coding, wireless communications, and cryptography. Dr. Bloch is a member of the IEEE and has served on the organizing committee of several international conferences; he is the current chair of the Online Committee of the IEEE Information Theory Society. He is the co-recipient of the IEEE Communications Society and IEEE Information Theory Society 2011 Joint Paper Award and the co-author of the textbook Physical-Layer Security: From Information Theory to Security Engineering published by Cambridge University Press.

J. Nicholas Laneman (SM'07) is Founding Director of the Wireless Institute in the College of Engineering, an Associate Professor of Electrical Engineering, and a Fellow of the John J. Reilly Center for Science, Technology, and Values at the University of Notre Dame. He joined the faculty in August 2002 shortly after earning a Ph.D. in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology (MIT). His research and teaching interests are in communications architecture – a blend of information theory, error-control coding, signal processing for communications, network protocols, and hardware design – with current emphasis on wireless systems.

Dr. Laneman has received a 2006 Presidential Early-Career Award for Scientists and Engineers (PECASE), a 2006 National Science Foundation (NSF) CAREER Award, a 2003 Oak Ridge Associated Universities (ORAU) Ralph E. Powe Junior Faculty Enhancement Award, and the 2001 MIT EECS Harold L. Hazen Graduate Teaching Award. He is an IEEE Senior Member and has served as an Associate Editor for IEEE Transactions on Communications, as a Guest Editor for Special Issues of IEEE TRANSACTIONS ON INFORMATION THEORY and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, and as the first Online Editor for the IEEE Information Theory Society.