On Secure Communication with Constrained Randomization

Matthieu R. Bloch
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332–0250
Email: matthieu.bloch@ece.gatech.edu

Jörg Kliewer
School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, New Mexico 88003-8001
Email: jkliewer@nmsu.edu

***Abstract*—In this paper, we investigate how constraints on the randomization in the encoding process affect the secrecy rates achievable over wiretap channels. In particular, we characterize the secrecy capacity with a rate-limited local source of randomness and a less capable eavesdropper's channel, which shows that limited rate incurs a secrecy rate penalty but does not preclude secrecy. We also show that secure communication is possible when randomizing with a non-uniform source of randomness, which suggests the possibility of designing robust coding schemes.**

I. Introduction

The wiretap channel model [1], [2] has attracted much attention in recent years because of its potential to strengthen the security of communication systems [3], [4]. Although this model provides a convenient abstraction to design codes for secure communication, it relies on two implicit simplifying assumptions. First, the model assumes that the transmitter knows the statistics of the channel. Second, the model assumes that the transmitter has access to an arbitrary local source of randomness, whose statistics can be optimized as part of the code design. In practice, however, these assumptions are unlikely to be perfectly guaranteed. For instance, an eavesdropper has little incentive to help characterize the channel statistics and, realistically, the legitimate parties may only have approximate knowledge of the true statistics. Similarly, the statistics of the local source of randomness may be imperfectly known, or the source may only provide a limited rate of randomness.

Secure communications with imperfect channel knowledge have already been the subject of previous investigations. For instance, several works have studied *compound* wiretap channels (see [5] and references therein), in which the transmitter only knows that its channel belongs to a set of possible channels. Other works have investigated the secrecy capacity of state-dependent channels under different assumptions regarding state information (see [3], [4] and references therein). Finally, some recent works have shown the existence of *universal* wiretap codes that guarantee secrecy as soon as the channel capacity of the eavesdropper's channel is known to be low enough [6]; for multiple-antenna systems, the existence of universal codes is even established without any assumption regarding the eavesdropper's channel statistics [7], [8].

In contrast to the problem of channel knowledge, little attention has been devoted to the problem of imperfect local sources of randomness. In particular, the questions of how much randomness is required to guarantee secrecy and how sensitive are secure communication codes to imperfections in randomness are still largely open.

In this paper, we provide partial answers to these questions. Our main contributions are 1) the characterization of secrecy capacity with a rate-limited source of randomness and a less capable eavesdropper's channel, and 2) the derivation of a sufficient condition for secure communication with a non-uniform randomization.

The remainder of the paper is organized as follows. Section II introduces the wiretap channel model used to analyze the effect of constrained randomization and presents our results on the secrecy-capacity of wiretap channels with a rate-limited local source of randomness. Section III discusses the possibility of secure communication with a non-uniform local source of randomness that cannot be processed. All proofs are relegated to appendices to streamline the presentation; due to space limitations, some proof details are also omitted.

II. Rate-Limited Randomness

We consider a discrete wiretap channel $(\mathcal{X}, W_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y} \times \mathcal{Z})$, characterized by a finite input alphabet $\mathcal{X}$, two finite output alphabets $\mathcal{Y}$ and $\mathcal{Z}$, and transition probabilities $p_{\mathsf{YZ}|\mathsf{X}}$. As illustrated in Figure 1, we assume that the transmitter (Alice) wishes to transmit a secret message to the receiver observing Y^n (Bob), in the presence of an eavesdropper observing Z^n (Eve). The channel $(\mathcal{X}, W_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y})$ is called the main channel while the channel $(\mathcal{X}, W_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z})$ is called the eavesdropper's channel. We assume the eavesdropper's channel is less capable, that is for any input X we have $\mathbb{I}(\mathsf{X};\mathsf{Z}) \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y})$. The encoding process may be stochastic, but the only source of randomness is a discrete memoryless[1] source $(\mathcal{R}, p_{\mathsf{R}})$ with known alphabet $\mathcal{R}$ and known statistics p_{R}. This model captures a situation in

This work was supported in part by the U.S. National Science Foundation under grants CCF-0830666 and CCF-1017632.

[1] The assumption of a memoryless source is a matter of convenience, and the proofs in the appendices generalize to arbitrary sources.

978-1-4673-2579-0/12/$31.00 ©2012 IEEE

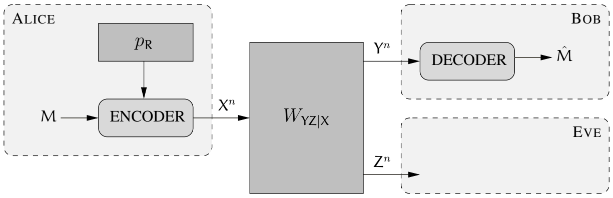

Fig. 1. Communication over a randomness-limited wiretap channel.

which the transmitter does not have access to a infinite pool of random numbers, and those must be generated on-the-fly during encoding from a source of randomness (thermal noise, photon counting). In addition, it forces us to explicitly specify how to use the randomness in the encoding process.

Definition 1: A $(2^{nR}, n)$ wiretap code $\mathcal{C}_n$ for the discrete wiretap channel $\left(\mathcal{X}, p_{\mathsf{YZ|X}}, \mathcal{Y} \times \mathcal{Z}\right)$ with local source of randomness $(\mathcal{R}, p_{\mathsf{R}})$ consists of the following.

- a message alphabet $\mathcal{M} = [\![1, 2^{nR}]\!]$;
- an encoding function $e : \mathcal{M} \times \mathcal{R}^n \to \mathcal{X}^n$;
- a decoding function $f : \mathcal{Y}^n \to \mathcal{M} \cup \{?\}$.

The performance of $\mathcal{C}_n$ is measured in terms of the average probability of error $P_e(\mathcal{C}_n) \triangleq \mathbb{P}\Big(\mathsf{M} \neq \hat{\mathsf{M}} | \mathcal{C}_n\Big)$ and of the secrecy leakage $L(\mathcal{C}_n) \triangleq \mathbb{I}(\mathsf{M}; \mathsf{Z}^n | \mathcal{C}_n)$

Definition 2: A rate R is achievable if there exists a sequence of $(2^{nR}, n)$ wiretap codes $\{\mathcal{C}_n\}_{n \geqslant 1}$ such that

$$\lim_{n\to\infty} P_e(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} L(\mathcal{C}_n) = 0.$$

The (strong) secrecy capacity with rate-limited randomness C_s is defined as the supremum of all achievable rates.

Remark 1: The definition of a wiretap code above implicitly allows the encoder to process the observations obtained from the local source of randomness. In particular, the encoder can remove a possible bias in the randomness. What happens when the encoder does not perfectly process the local source is discussed in Section III.

Remark 2: The model can be viewed as a special case of wiretap channel with channel state known non-causally at the transmitter [9], in which the state is independent of the channel; however, our result does not follow from [9] because we consider a strong secrecy metric.

Proposition 1: The secrecy capacity of a wiretap channel $(\mathcal{X}, W_{\mathsf{YZ|X}}, \mathcal{Y} \times \mathcal{Z})$ with a rate-limited source of local randomness $(\mathcal{R}, p_{\mathsf{R}})$ and a less capable eavesdropper's channel[2] is

$$C_s = \max_{p_{\mathsf{UVXYZ}} \in \mathcal{P}} \left(\mathbb{I}(\mathsf{X}; \mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X}; \mathsf{Z}|\mathsf{U})\right)$$

where the set $\mathcal{P}$ is the set of distributions p_{UXYZ} that factorize as $p_{\mathsf{UXYZ}} = p_{\mathsf{U}} p_{\mathsf{X|U}} W_{\mathsf{YZ|X}}$ and with $\mathbb{I}(\mathsf{X}; \mathsf{Z}|\mathsf{U}) \leqslant \mathbb{H}(\mathsf{R})$.

Proof: See Appendix A and Appendix B. ■

Remark 3: Using standard techniques, one can show that the cardinality of $\mathcal{U}$ is bounded by $|\mathcal{U}| \leqslant 2$.

The expression in Proposition 1 is similar to that obtained in [2, Corollary 2]. The effect of the local source of randomness explicitly appears in the expression through the auxiliary time-sharing random variable U and the constraint $\mathbb{I}(\mathsf{X}; \mathsf{Z}|\mathsf{U}) \leqslant \mathbb{H}(\mathsf{R})$. Proposition 1 confirms the optimal structure of the encoder, which performs two distinct operations:

1) *Uniformization:* the encoder generates nearly-uniform random numbers K at rate $\mathbb{H}(\mathsf{R})$ from the local source of randomness;
2) *Randomization:* the encoder uses a fraction $\mathbb{I}(\mathsf{X}; \mathsf{Z}|\mathsf{U})$ of the randomness rate to randomize the choice of a codeword;

The identification of the optimal encoder structure suggests that non-uniform randomization may affect the performance of a code, which we discuss in Section III. Proposition 1 also highlights that the common folklore in information-theoretic security, according to which secrecy is achievable provided the randomization can exhaust the capacity of the Eve's channel, is somewhat misleading. If the source provides a non-zero rate of randomness ($\mathbb{H}(\mathsf{R}) > 0$), then the secrecy capacity with a rate-limited source of randomness is positive if and only if the secrecy capacity with unlimited randomness is positive. Intuitively, this happens because the channel seen by Eve is an "effective channel", which is partly controlled by Alice through time-sharing and the choice of the codebook.

Also note that if the rate of randomness vanishes, then no secure communication is possible. This confirms that, except for pathological channels (for instance, one for which $\mathbb{I}(\mathsf{X}; \mathsf{Z}) = 0$ for any X), one cannot replace the local source of randomness by a pseudo-random number generator without losing the information-theoretic secrecy guarantees.

III. Non-Uniform Rate-Limited Randomness

The result of Proposition 1 suggests that one should always "uniformize" the local source of randomness to create uniformly distributed random numbers. This operation, however, may be imperfect and one may wonder whether achieving secrecy is then still possible. A situation where the random numbers may not be perfectly uniform is if the local source of randomness is another message source; understanding this setting is crucial to assess whether secrecy constraints incur an overall rate loss or not [4].

For simplicity, we assume that the output of uniformization is a random variable $\mathsf{K} \in [\![1, 2^{nR_r}]\!]$ with perhaps non-uniform distribution p_{K}. In this case, we show that secrecy is still achievable, but at a lower rate limited by the Rényi entropy rate of order two $\frac{1}{n} R_2(\mathsf{K})$ where

$$R_2(\mathsf{K}) \triangleq -\log \left(\sum_{u \in [\![1, 2^{nR_r}]\!]} p_{\mathsf{K}}(u)^2 \right).$$

[2] We used the less capable assumption to avoid dealing with the problem of channel prefixing. Days before submitting the current paper, [10] was posted on ArXiv and independently solved the general case. Proposition 1 appears as [10, Corollary 12].

Proposition 2: A secrecy rate R is achievable when randomization is performed with randomness K if it satisfies

$$R < \max_{p_{\mathsf{UXYZ}}\in\mathcal{P}} \left(\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U})\right),$$

where $\mathcal{P}$ is the set of distributions $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U})$ that factorize as $p_{\mathsf{U}}p_{\mathsf{X}|\mathsf{U}}W_{\mathsf{YZ}|\mathsf{X}}$ and such that $\mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) < \frac{1}{n}R_2(\mathsf{K})$.

Proof: See Appendix C. ■

It is not straightforward to establish a converse for Proposition 2 because typical converse arguments make no assumption regarding the internal structure of the encoder. In particular, it seems difficult to include a constraint that would prevent any processing of K.

In general, $\frac{1}{n}R_2(\mathsf{K}) \leqslant \frac{1}{n}\mathbb{H}(\mathsf{K})$, and the constraint in Proposition 2 is therefore more stringent than in Proposition 1. The effect can be quite dramatic, and the following example shows that the gap between the rates in Proposition 1 and Proposition 2 can be large.

Example 1: Assume the encoder performs randomization with a biased local source of randomness, which produces random numbers $\mathsf{K} \in [\![1, 2^{nR_r}]\!]$ such that

$$\mathbb{P}(\mathsf{K}=1) = 2^{-n\alpha R_r} \text{ and } \mathbb{P}(\mathsf{K}=i) = \frac{1-2^{-n\alpha R_r}}{2^{nR_r}-1} \text{ if } i \neq 1,$$

where $\alpha \in]0;\frac{1}{2}[$ is a parameter that controls the uniformity of the distribution. Note that

$$\lim_{n\to\infty} \tfrac{1}{n}R_2(\mathsf{K}) = \alpha R_r \quad \text{whereas} \quad \lim_{n\to\infty} \tfrac{1}{n}\mathbb{H}(\mathsf{K}) = R_r.$$

Consequently, without proper uniformization, the achievable rates predicted in Proposition 2 could be arbitrarily small.

IV. Conclusion

We have shown for the classical wiretap channel that strong secrecy can be guaranteed even in the presence of non-uniform or rate-limited randomness, albeit at the expense of a lower secrecy capacity. The result of this work enables several interesting applications. For example, if the public message in the wiretap channel model is identified as the output of a source encoder, which is in general not uniformly distributed, extra information can be conveyed publicly while still providing secure communication. Another application is secure transmission in a network, in which multiple links are wiretapped by the same eavesdropper via channels with different capacities and in which only a given amount of randomness exists.

Appendix A
Converse Proof for Proposition 1

Let $\epsilon > 0$ and let R be an achievable rate. Then, there exists a $(2^{nR}, n)$ code $\mathcal{C}_n$ such that $P_e(\mathcal{C}_n) \leqslant \epsilon$ and $L(\mathcal{C}_n) \leqslant \epsilon$. Following the converse technique in [2], we obtain

$$R \leqslant \tfrac{1}{n}\sum_{i=1}^{n}\left(\mathbb{I}\Big(\mathsf{M};\mathsf{Y}_i|\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\Big) - \mathbb{I}\Big(\mathsf{M};\mathsf{Z}_i|\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\Big)\right) + \delta(\epsilon),$$

where $\tilde{\mathsf{Y}}^{i-1} \triangleq \{\mathsf{Y}_j\}_{j=1}^{i-1}$, $\tilde{\mathsf{Z}}^{i+1} \triangleq \{\mathsf{Z}_j\}_{j=i+1}^{n}$ and $\delta(\epsilon)$ is a function of ϵ that goes to zero with ϵ. Next, by definition of the encoder e and by independence of R^n and M,

$$\frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{M}) = \frac{1}{n}\mathbb{H}(e(\mathsf{M},\mathsf{R}^n)|\mathsf{M}) \leqslant \frac{1}{n}\mathbb{H}(\mathsf{R}^n) = \mathbb{H}(\mathsf{R}). \quad (1)$$

Now, we also have

$$\begin{aligned}
&\frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{M})\\
&= \frac{1}{n}\mathbb{H}(\mathsf{X}^n) - \frac{1}{n}\mathbb{H}(\mathsf{M}) + \frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{X}^n)\\
&\geqslant \frac{1}{n}\mathbb{H}(\mathsf{X}^n) - \frac{1}{n}\mathbb{H}(\mathsf{M}) + \frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{X}^n) + \frac{1}{n}\mathbb{I}(\mathsf{M};\mathsf{Z}^n) - \delta(\epsilon)\\
&= \frac{1}{n}\mathbb{H}(\mathsf{X}^n) - \frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{Z}^n) + \frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{X}^n) - \delta(\epsilon)\\
&= \frac{1}{n}\mathbb{H}(\mathsf{X}^n) - \frac{1}{n}\mathbb{H}(\mathsf{M}\mathsf{X}^n|\mathsf{Z}^n) + \frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{M}\mathsf{Z}^n)\\
&\qquad\qquad + \frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{X}^n) - \delta(\epsilon)\\
&= \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n) + \frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{M}\mathsf{Z}^n) - \delta(\epsilon)\\
&\geqslant \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n) - \delta(\epsilon), \qquad (2)
\end{aligned}$$

where the last inequality follows because $\mathsf{M} \to \mathsf{X}^n \to \mathsf{Z}^n$ forms a Markov chain and $\mathbb{H}(\mathsf{M}|\mathsf{X}^n\mathsf{Z}^n) = \mathbb{H}(\mathsf{M}|\mathsf{X}^n)$. Then,

$$\begin{aligned}
&\frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n)\\
&\qquad \geqslant \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{H}\Big(\mathsf{Z}_i|\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\Big) - \mathbb{H}\Big(\mathsf{Z}_i|\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\mathsf{X}_i\Big)\right)\\
&\qquad = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\Big(\mathsf{X}_i;\mathsf{Z}_i|\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\Big), \qquad (3)
\end{aligned}$$

where the inequality follows because conditioning does not increase entropy and $\tilde{\mathsf{Z}}^{i+1}\mathsf{Y}^{i-1} \to \mathsf{X}_i \to \mathsf{Z}_i$ forms a Markov chain. Let us now define a random variable Q independent of all others and uniformly distributed on $[\![1,n]\!]$. For $i \in [\![1,n]\!]$, we also define $\mathsf{U}_i \triangleq \mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}$ and $\mathsf{V}_i \triangleq \mathsf{U}_i\mathsf{M}$. Combining inequalities (1), (2), and (3), and substituting the definition of Q, U_i, V_i above, we obtain

$$\begin{aligned}
R &\leqslant \mathbb{I}(\mathsf{V}_\mathsf{Q};\mathsf{Y}_\mathsf{Q}|\mathsf{Q}\mathsf{U}_\mathsf{Q}) - \mathbb{I}(\mathsf{V}_\mathsf{Q};\mathsf{Z}_\mathsf{Q}|\mathsf{Q}\mathsf{U}_\mathsf{Q}) + \delta(\epsilon) \qquad (4)\\
\mathbb{H}(\mathsf{R}) &\geqslant \mathbb{I}(\mathsf{X}_\mathsf{Q};\mathsf{Z}_\mathsf{Q}|\mathsf{Q}\mathsf{U}_\mathsf{Q}) - \delta(\epsilon). \qquad (5)
\end{aligned}$$

Finally, define $\mathsf{U} \triangleq \mathsf{U}_\mathsf{Q}\mathsf{Q}$, $\mathsf{V} \triangleq \mathsf{V}_\mathsf{Q}\mathsf{Q}$, $\mathsf{X} \triangleq \mathsf{X}_\mathsf{Q}$, $\mathsf{Y} \triangleq \mathsf{Y}_\mathsf{Q}$ and $\mathsf{Z} \triangleq \mathsf{Z}_\mathsf{Q}$. Note that $\mathsf{U} \to \mathsf{V} \to \mathsf{X} \to \mathsf{YZ}$ forms a Markov chain and that the statistics $p_{\mathsf{YZ}|\mathsf{X}}$ are those of the original channel $W_{\mathsf{YZ}|\mathsf{X}}$. Substituting these definitions in (4) and (5), we obtain

$$\begin{aligned}
R &\leqslant \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) + \delta(\epsilon)\\
\mathbb{H}(\mathsf{R}) &\geqslant \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) - \delta(\epsilon).
\end{aligned}$$

Because the eavesdropper's channel is less capable, then $\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U})$. Since ϵ can be chosen arbitrarily small, we obtained the desired converse.

Appendix B
Achievability Proof for Proposition 1

The proof relies on binning, superposition coding, and stochastic encoding as in [2, Lemma 2]; however, since the local source of randomness is explicit and since we impose a strong secrecy criterion, some details must be laid out carefully. We denote the set of ϵ-strongly typical sequences with respect to p_{X} by $T_\epsilon^n(\mathsf{X})$ and the set of conditional ϵ-strongly typical sequence with respect to p_{YX} and $x^n \in T_\epsilon^n(\mathsf{X})$ by $T_\epsilon^n(\mathsf{Y}|x^n)$.

We first show the existence of a code $\mathcal{C}_n$ assuming an unlimited amount of uniform randomness is available. We fix a joint distribution p_{UX} on $\mathcal{U} \times \mathcal{X}$ such that[3] $\mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) \leqslant \mathbb{H}(\mathsf{R})$ and $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) > 0$, and we construct a code $\mathcal{C}_n$ for the broadcast channel with confidential messages $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y} \times \mathcal{Z})$. Let $\epsilon > 0$, $R > 0$, $R_r > 0$, $R_0 > 0$ and $n \in \mathbb{N}$. We randomly construct a code as follows. We generate 2^{nR_0} sequences independently at random according to p_{U}, which we label $u^n(i)$ for $i \in [\![1, 2^{nR_0}]\!]$. For each sequence $u^n(i)$, we generate $2^{n(R+R_r)}$ sequences independently at random according to $p_{\mathsf{X}|\mathsf{U}}$, which we label $x^n(i,j,k)$ with $j \in [\![1, 2^{nR}]\!]$ and $k \in [\![1, 2^{nR_r}]\!]$. To transmit a message $i \in [\![1, 2^{nR_0}]\!]$ and $j \in [\![1, 2^{nR}]\!]$, the transmitter obtains a realization k of a uniform random number $\mathsf{K} \in [\![1, 2^{nR_r}]\!]$, and transmits $x^n(i,j,k)$ over the channel. Upon receiving y^n, Bob decodes i as the received index if it is the unique one such that $(u^n(i), y^n) \in T_\epsilon^n(UY)$; otherwise, he declares an error. Bob then decodes (j,k) as the other pair of indices if it is the unique one such that $(x^n(i,j,k), y^n) \in T_\epsilon^n(UXY)$. Similarly, upon receiving z^n, Eve decodes i as the received index if it is the unique one such that $(u^n(i), z^n) \in T_\epsilon^n(UZ)$; otherwise, she declares an error.

Lemma 1: If $R_0 < \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z}))$ and $R + R_r < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U})$, then $\mathbb{E}(P_e(\mathsf{C}_n)) \leqslant 2^{-\alpha n}$ for some $\alpha > 0$.

Proof: The proof follows from a standard random coding argument and is omitted. ■

Lemma 2: If $R_r > \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U})$, then we have $\mathbb{E}_{\mathsf{C}_n}(\mathbb{V}(p_{\mathsf{MZ}^n}, p_{\mathsf{M}}p_{\mathsf{Z}^n})) \leqslant 2^{-\beta n}$ for some $\beta > 0$, where $\mathbb{V}$ denotes the variational distance.

Proof: See Appendix C. ■

Using Markov's inequality, we conclude that there exists at least one code $\mathcal{C}_n$ satisfying the rate inequalities in Lemma 1 and Lemma 2, such that $P_e(\mathcal{C}_n) \leqslant 3 \cdot 2^{-\alpha n}$ and $\mathbb{V}(p_{\mathsf{MM}_0\mathsf{Z}^n}, p_{\mathsf{M}}p_{\mathsf{M}_0\mathsf{Z}^n}) \leqslant 3 \cdot 2^{-\beta n}$. Finally, the uniform numbers K can be approximately obtained from $(\mathcal{R}, p_{\mathsf{R}})$ with an appropriate function ϕ.

Lemma 3 (adapted from [11]): If $R_r < \mathbb{H}(\mathsf{R})$, then there exists ϕ such that $\mathbb{V}\big(p_{\phi(\mathsf{R}^n)}, p_{\mathsf{K}}\big) \leqslant 2^{-n\eta}$ for some $\eta > 0$.

Consequently, one can show that, even if the code $\mathcal{C}_n$ is used with $\phi(\mathsf{R}^n)$ in place of K, then

$$P_e(\mathcal{C}_n) \leqslant 2^{-\kappa n} \quad \text{and} \quad \mathbb{V}(p_{\mathsf{MM}_0\mathsf{Z}^n}, p_{\mathsf{M}}p_{\mathsf{M}_0\mathsf{Z}^n}) \leqslant 2^{-\kappa n}.$$

for some $\kappa > 0$. The fact that $L(\mathcal{C}_n) \leqslant 2^{-\kappa' n}$ for some $\kappa' > 0$ follows from [12, Lemma 1]. Combining all rate constraints in the previous lemmas, and since ϵ can be chosen arbitrarily small, we see that any rate $R < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U})$ such that $\mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) \leqslant \mathbb{H}(\mathsf{R})$ is achievable. Note that the constraint on R_0 plays no role since it represents a negligible rate of time sharing information to synchronize transmitter and receiver.

Appendix C
Proof of Proposition 2

The proof is similar to that Appendix B, with Lemma 4 in place of Lemma 2. Lemma 2 is obtained in the special case of K uniform.

Lemma 4: If $\frac{1}{n}R_2(\mathsf{K}) > \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U})$, then we have $\mathbb{E}_{\mathsf{C}_n}(\mathbb{V}(p_{\mathsf{MZ}^n}, p_{\mathsf{M}}p_{\mathsf{Z}^n})) \leqslant 2^{-\beta n}$ for some $\beta > 0$.

The proof relies on a careful analysis and a slight generalization of the "cloud-mixing" lemma [13]; the notation is that of Appendix B. We define the distribution $q_{\mathsf{U}^n\mathsf{X}^n\mathsf{Z}^n}$ on $\mathcal{U}^n \times \mathcal{X}^n \times \mathcal{Z}^n$ as

$$q_{\mathsf{U}^n\mathsf{X}^n\mathsf{Z}^n}(u^n, x^n, z^n) = W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)p_{\mathsf{X}^n\mathsf{U}^n}(x^n, u^n).$$

First note that the variational distance $\mathbb{V}(p_{\mathsf{MZ}^n}, p_{\mathsf{M}}p_{\mathsf{Z}^n})$ can be bounded as follows.

$$\begin{aligned}
&\mathbb{V}(p_{\mathsf{MZ}^n}, p_{\mathsf{M}}p_{\mathsf{Z}^n}) \\
&\quad \leqslant \mathbb{V}(p_{\mathsf{MU}^n\mathsf{Z}^n}, p_{\mathsf{M}}p_{\mathsf{U}^n\mathsf{Z}^n}) \\
&\quad = \mathbb{E}_{\mathsf{U}^n\mathsf{M}}\big(\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{MU}^n}, p_{\mathsf{Z}^n|\mathsf{U}^n}\big)\big) \\
&\quad \leqslant \mathbb{E}_{\mathsf{U}^n\mathsf{M}}\big(\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{MU}^n}, q_{\mathsf{Z}^n|\mathsf{U}^n}\big) + \mathbb{V}\big(q_{\mathsf{Z}^n|\mathsf{U}^n}, p_{\mathsf{Z}^n|\mathsf{U}^n}\big)\big) \\
&\quad \leqslant 2\mathbb{E}_{\mathsf{U}^n\mathsf{M}}\big(\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{MU}^n}, q_{\mathsf{Z}^n|\mathsf{U}^n}\big)\big)
\end{aligned}$$

Then, let U_1^n be the sequence in $\mathcal{U}^n$ corresponding to $\mathsf{M}_0 = 1$. By symmetry of the random code construction, the average of the variational distance $\mathbb{V}(p_{\mathsf{MZ}^n}, p_{\mathsf{M}}p_{\mathsf{Z}^n})$ over randomly generated codes C_n satisfies

$$\begin{aligned}
&\mathbb{E}_{\mathsf{C}_n}(\mathbb{V}(p_{\mathsf{MZ}^n}, p_{\mathsf{M}}p_{\mathsf{Z}^n})) \\
&\qquad\qquad \leqslant 2\mathbb{E}_{\mathsf{C}_n}\big(\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{U}^n=\mathsf{U}_1^n\mathsf{M}=1}, q_{\mathsf{Z}^n|\mathsf{U}^n=\mathsf{U}_1^n}\big)\big),
\end{aligned}$$

where

$$p_{\mathsf{Z}^n|\mathsf{U}^n=\mathsf{U}_1^n\mathsf{M}=1}(z^n) = \sum_{k=1}^{2^{nR_r}} W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n(1,1,k))p_{\mathsf{K}}(k).$$

The average over the random codes can be further bounded by splitting the average between U_1^n and the random code $\mathsf{C}_n(u_1^n)$ for a fixed value of u_1^n.

$$\begin{aligned}
&\mathbb{E}_{\mathsf{C}_n}\big(\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{U}^n=\mathsf{U}_1^n\mathsf{M}=1}, q_{\mathsf{Z}^n|\mathsf{U}^n=\mathsf{U}_1^n}\big)\big) \\
&\ = \sum_{u_1^n \in \mathcal{U}^n} p_{\mathsf{U}^n}(u_1^n)\mathbb{E}_{\mathsf{C}_n(u_1^n)}\big(\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n\mathsf{M}=1}, q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}\big)\big) \\
&\ \leqslant 2\mathbb{P}(\mathsf{U}^n \notin T_\epsilon^n(\mathsf{U})) \\
&\ + \sum_{u_1^n \in T_\epsilon^n(\mathsf{U})} p_{\mathsf{U}^n}(u_1^n)\mathbb{E}_{\mathsf{C}_n(u_1^n)}\big(\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n\mathsf{M}=1}, q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}\big)\big).
\end{aligned}$$

where the last inequality follows from the fact that the variational distance is always less than 2. By construction, the first term on the right-hand side vanishes as n gets large; we now

[3] If such a probability distribution does not exist, then the result of Proposition 1 is trivial and there is nothing to prove.

proceed to bound the expectation in the second term. First note that, for any $z^n \in \mathcal{Z}^n$,

$$
\begin{aligned}
&\mathbb{E}_{\mathsf{C}_n(u_1^n)}\big(p_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n\mathsf{M}=1}(z^n)\big)\\
&= \mathbb{E}_{\mathsf{C}_n(u_1^n)}\left(\sum_{k=1}^{2^{nR_r}} W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n(1,1,k))p_{\mathsf{K}}(k)\right)\\
&= \sum_{k=1}^{2^{nR_r}} \mathbb{E}_{\mathsf{C}_n(u_1^n)}\big(W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n(1,1,k))\big)p_{\mathsf{K}}(k)\\
&= q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}(z^n).
\end{aligned}
$$

We now let $\mathbf{1}$ denote the indicator function and we define

$$
\begin{aligned}
p^{(1)}(z^n) &\triangleq \sum_{k=1}^{2^{nR_r}} W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n(1,1,k))p_{\mathsf{K}}(k)\\
&\qquad\qquad \mathbf{1}\{(x^n(1,1,k),z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)\}\\
p^{(2)}(z^n) &\triangleq \sum_{k=1}^{2^{nR_r}} W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n(1,1,k))p_{\mathsf{K}}(k)\\
&\qquad\qquad \mathbf{1}\{(x^n(1,1,k),z^n)\notin T_\epsilon^n(\mathsf{XZ}|u_1^n)\}
\end{aligned}
$$

so that we can upper bound $\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n\mathsf{M}=1}, q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}\big)$ as

$$
\begin{aligned}
&\mathbb{V}\big(p_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n\mathsf{M}=1}, q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}\big)\\
&\leqslant \sum_{z^n\notin T_\epsilon^n(\mathsf{Z}|u_1^n)} \left|p_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n\mathsf{M}=1}(z^n) - q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}(z^n)\right|\\
&\quad + \sum_{z^n\in T_\epsilon^n(\mathsf{Z}|u_1^n)} \left|p^{(1)}(z^n) - \mathbb{E}\Big(p^{(1)}(z^n)\Big)\right|\\
&\quad + \sum_{z^n\in T_\epsilon^n(\mathsf{Z}|u_1^n)} \left|p^{(2)}(z^n) - \mathbb{E}\Big(p^{(2)}(z^n)\Big)\right|.
\end{aligned}
$$

One can show that the average of the first sum and of the last sum vanish as n goes to infinity. We now focus on the average of the second sum. For $z^n \in T_\epsilon^n(\mathsf{Z}|u_1^n)$, Jensen's inequality and the concavity of $x \mapsto \sqrt{x}$ guarantee that

$$
\mathbb{E}\left(\left|p^{(1)}(z^n) - \mathbb{E}\Big(p^{(1)}(z^n)\Big)\right|\right) \leqslant \sqrt{\mathrm{Var}\big(p^{(1)}(z^n)\big)}.
$$

In addition,

$$
\begin{aligned}
\mathrm{Var}\Big(p^{(1)}(z^n)\Big) &= \sum_{k=1}^{2^{nR_r}} p_{\mathsf{K}}(k)^2\mathrm{Var}\left(W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|\mathsf{X}^n(1,1,k))\right.\\
&\qquad\qquad \left.\mathbf{1}\{(\mathsf{X}^n(1,1,k),z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)\}\right)
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\mathrm{Var}\left(W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|\mathsf{X}^n(1,1,k))\mathbf{1}\{(\mathsf{X}^n(1,1,k),z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)\}\right)\\
&= \sum_{x^n\in\mathcal{X}^n} p_{\mathsf{X}^n|\mathsf{U}^n=u_1^n}(x^n)\\
&\qquad \left(W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)\mathbf{1}\{(x^n,z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)\}\right)^2\\
&= \sum_{x^n:(x^n,z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)} p_{\mathsf{X}^n|\mathsf{U}^n=u_1^n}(x^n)W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)^2\\
&\overset{(a)}{\leqslant} 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})-\delta(\epsilon))}\\
&\qquad \sum_{x^n:(x^n,z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)} p_{\mathsf{X}^n|\mathsf{U}^n=u_1^n}(x^n)W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)\\
&\leqslant 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})-\delta(\epsilon))}q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}(z^n)\\
&\overset{(b)}{\leqslant} 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon))},
\end{aligned}
$$

where (a) and (b) follow from the AEP; therefore,

$$
\begin{aligned}
\mathrm{Var}\Big(p^{(1)}(z^n)\Big) &\leqslant 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon))}\sum_{k=1}^{2^{nR_r}} p_{\mathsf{K}}(k)^2\\
&\leqslant 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon))+\frac{R_2(\mathsf{K})}{n}}.
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{z^n\in T_\epsilon^n(\mathsf{Z}|u_1^n)} \mathbb{E}\left(\left|p^{(1)}(z^n) - \mathbb{E}\Big(p^{(1)}(z^n)\Big)\right|\right)\\
&\qquad \leqslant 2^{n\mathbb{H}(\mathsf{Z}|\mathsf{U})}2^{-\frac{n}{2}(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon)+\frac{R_2(\mathsf{K})}{n})}\\
&\qquad = 2^{-\frac{n}{2}(\frac{R_2(\mathsf{K})}{n}-\mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U})-\delta(\epsilon))}
\end{aligned}
$$

Hence, if $\frac{R_2(\mathsf{K})}{n} > \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) + \delta(\epsilon)$, the sum vanishes as n goes to infinity, which concludes the proof. Note that if K is uniform, then $R_2(\mathsf{K}) = nR_r$, and we obtain Lemma 2.

References

[1] A. D. Wyner, “The Wire-Tap Channel,” *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1367, October 1975.

[2] I. Csiszár and J. Körner, “Broadcast Channels with Confidential Messages,” *IEEE. Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[3] Y. Liang, H. V. Poor, and S. S. (Shitz), *Information-Theoretic Security*, ser. Foundations and Trends in Communications and Information Theory. Delft, 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 University Press, October 2011.

[5] Y. Liang, G. Kramer, H. V. Poor, and S. S. (Shitz), “Compound Wiretap Channels,” *EURASIP Journal on Wireless Comm. and Networking*, vol. 142374, pp. 1–12, 2009.

[6] J. Muramatsu and S. Miyake, “Construction of Wiretap Channel Codes by Using Sparse Matrices,” in *Proc. IEEE Information Theory Workshop*, Taormina, Sicily, October 2009, pp. 105–109.

[7] X. He and A. Yener, “Providing secrecy when the eavesdropper channel is arbitrarily varying: A case for multiple antennas,” in *Proc. 48th Annual Allerton Conf. Communication, Control, and Computing*, 2010, pp. 1228–1235.

[8] X. He, A. Khisti, and A. Yener, “MIMO Broadcast Channel with Arbitrarily Varying Eavesdropper Channel: Secrecy Degrees of Freedom,” in *Proc. IEEE Global Telecommunications Conf.*, 2011, pp. 1–5.

[9] Y. Chen and A. J. Han Vinck, “Wiretap Channel With Side Information,” *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 395–402, 2008.

[10] S. Watanabe and Y. Oohama, “Broadcast Channels with Confidential Messages by Randomness Constrained Stochastic Encoder,” preprint, January 2012. [Online]. Available: arXiV:1201.6468

[11] R. Ahlswede and I. Csiszár, “Common Randomness in Information Theory and Cryptography. II. CR Capacity,” *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 225–240, January 1998.

[12] I. Csiszár, “Almost Independence and Secrecy Capacity,” *Problems of Info. Transmission*, vol. 32, no. 1, pp. 40–47, January-March 1996.

[13] P. W. Cuff, “Communication in Networks for Coordinating Behavior,” Ph.D. dissertation, Princeton University, July 2009.