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 discuss a more practical aspect of rate-limited randomization in the context of cooperative jamming. Finally, we show that secure communication is possible with a non-uniform source for randomness; this 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 (see [5] and reference therein), 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 [6] and references therein), in which the transmitter only knows that its channel belongs to a set of possible channels. Secure communication is often possible but the best channel to the eavesdropper usually limits secrecy rates. Other works have investigated the secrecy capacity of state-dependent channels under different assumptions regarding state information (see [3], [4] and references therein). In another approach, [7] has shown the existence of *universal* wiretap codes, which guarantee secrecy and reliability as soon as the channel capacity of the eavesdropper's channel is low enough.

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, 2) practical considerations on the effect of limited randomness for cooperative jamming, and 3) 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 rate-limited randomness in the context of cooperative jamming. Finally, Section IV discusses the possibility of secure communication with a non-uniform local source of randomness that cannot be processed.

II. Rate-Limited Randomness: Theoretical Considerations

Unless otherwise specified, 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 $(\mathcal{R}, p_R)$ with known alphabet $\mathcal{R}$ and known statistics p_R. This model captures a situation in which the transmitter

---

[1]The assumption of a memoryless source is a matter of convenience, and the proofs in the appendices generalize easily to arbitrary sources.
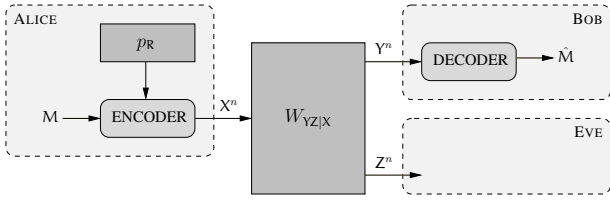
Fig. 1. Communication over a randomness-limited wiretap channel.

Fig. 2. Cooperative jamming with rate-limited randomness.

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 specify explicitly how to use the randomness provided by the source in the encoding process.

Definition 1: A $(2^{nR}, n)$ wiretap code $\mathcal{C}_n$ for the discrete wiretap channel $(\mathcal{X}, p_{YZ|X}, \mathcal{Y} \times \mathcal{Z})$ with local source of randomness $(\mathcal{R}, p_R)$ 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 $P_e(\mathcal{C}_n) \triangleq \mathbb{P}\left(\mathsf{M} \neq \hat{\mathsf{M}}|\mathcal{C}_n\right)$ 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 IV.

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

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

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

Proof: See Appendix A and Appendix B. ∎

Remark 2: 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

[2]We used the less capable assumption to avoid dealing with the problem of channel prefixing. Days before submitting the current paper, [8] was posted on ArXiv and independently solved the general case. Proposition 1 appears as [8, Corollary 12].
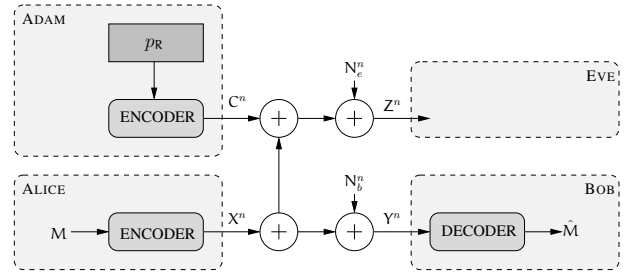
auxiliary time-sharing random variable U and the constraint $\mathbb{I}(\mathsf{X}; \mathsf{Z}|\mathsf{U}) \leqslant \mathbb{H}(\mathsf{R})$. Proposition 1 confirms the optimal structure of the encoder, which performs two distinct operations:

1) *Uniformization:* the encoder generates nearly-uniform random numbers U_r 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 suggest that non-uniform randomization may affect the performance of a code, which we discuss in Section IV. 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. RATE-LIMITED RANDOMNESS: PRACTICAL CONSIDERATIONS

It is legitimate to wonder how the results of previous sections generalize to continuous channels and, in particular, to Gaussian channels. There are no conceptual difficulties in analyzing the randomization part of the encoder since $\mathbb{I}(\mathsf{X}; \mathsf{Z}|\mathsf{U})$ remains finite with a power constraint; however, the simulation of Gaussian noise plays a key role in multi-user wiretap channels [9] as a means to perform cooperative jamming.

We analyze the situation illustrated in Figure 2, in which an eavesdropper observes the output of an AWGN channel with noise variance σ^2 and suffers from the added interference of a cooperative jammer (Adam). The signal obtained by the

eavesdropper is then

$$Z^n = X^n + C^n + N_e^n,$$

where X^n is the codeword transmitted by Alice, C^n is the interference introduced by Adam, and N_e^n is the channel noise. Cooperative jamming would consist in generating C^n i.i.d. according to a Gaussian distribution. With a local source of randomness, the following results holds.

Proposition 2: With a local source of randomness $(\mathcal{R}, p_R)$, a cooperative jammer can induce artificial Gaussian noise with power $\rho \leqslant \sigma^2 2^{2\mathbb{H}(R)-1}$.

Sketch of proof: The result follows by remarking that the objective of cooperative jamming is to increase the variance of Gaussian noise at the eavesdropper's terminal; therefore, the distribution of $C^n + N_e^n$ should be close to Gaussian, but C^n itself need not be Gaussian. In particular, the sequences C^n can be chosen from a codebook with average power constraint ρ; the result of channel resolvability over Gaussian channels [10] guarantees there exists a codebook with rate arbitrarily close to $\frac{1}{2} \log(1 + \frac{\rho}{\sigma^2})$ so that the distribution of $C^n + N_e^n$ is arbitrarily close to $\mathcal{N}(\sigma^2 + \rho)$. Since the rate of the codebook is given by the rate of the source $\mathbb{H}(R)$, the result follows. ∎
Consequently, rate-limited randomness effectively translates into a power constraint on the Gaussian artificial noise that the encoder introduces to jam the eavesdropper. Therefore, rate-limited randomness reduces the effectiveness of cooperating jamming but does not preclude it.

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

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

Proposition 3: A secrecy rate R is achievable when randomization is performed with randomness U_r if it satisfies

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

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

Proof: See Appendix C. ∎
It is not straightforward to establish a converse for Proposition 3 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 U_r.

In general, $\frac{1}{n} R_2(U_r) \leqslant \frac{1}{n}\mathbb{H}(U_r)$, and the constraint in Proposition 3 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 3 can be large.

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

$$\mathbb{P}(U_r = 1) = 2^{-n\alpha R} \text{ and } \mathbb{P}(U_r = i) = \frac{1 - 2^{-n\alpha R}}{2^{nR} - 1} \text{ if } i \neq 1,$$

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

$$\lim_{n \to \infty} \frac{1}{n} R_2(U_r) = \alpha R \quad \text{whereas} \quad \lim_{n \to \infty} \frac{1}{n}\mathbb{H}(U_r) = R.$$

Consequently, without proper uniformization, the achievable rates predicted in Proposition 3 could be arbitrarily small.

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

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

where $\tilde{Y}^{i-1} \triangleq \{Y_j\}_{j=1}^{i-1}$, $\tilde{Z}^{i+1} \triangleq \{Z_j\}_{j=i+1}^{n}$ 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,

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

Now, we also have

$$\frac{1}{n}\mathbb{H}(X^n|M)$$
$$= \frac{1}{n}\mathbb{H}(X^n) - \frac{1}{n}\mathbb{H}(M) + \frac{1}{n}\mathbb{H}(M|X^n)$$
$$\geqslant \frac{1}{n}\mathbb{H}(X^n) - \frac{1}{n}\mathbb{H}(M) + \frac{1}{n}\mathbb{H}(M|X^n) + \frac{1}{n}\mathbb{I}(M; Z^n) - \delta(\epsilon)$$
$$= \frac{1}{n}\mathbb{H}(X^n) - \frac{1}{n}\mathbb{H}(M|Z^n) - \delta(\epsilon) + \frac{1}{n}\mathbb{H}(M|X^n)$$
$$= \frac{1}{n}\mathbb{H}(X^n) - \frac{1}{n}\mathbb{H}(MX^n|Z^n) + \frac{1}{n}\mathbb{H}(X^n|MZ^n)$$
$$\qquad\qquad + \frac{1}{n}\mathbb{H}(M|X^n) - \delta(\epsilon)$$
$$= \frac{1}{n}\mathbb{I}(X^n; Z^n) + \frac{1}{n}\mathbb{H}(X^n|MZ^n) - \delta(\epsilon)$$
$$\geqslant \frac{1}{n}\mathbb{I}(X^n; Z^n) - \delta(\epsilon), \quad (2)$$

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

$$\frac{1}{n}\mathbb{I}(X^n; Z^n)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{H}\left(Z_i|\tilde{Z}^{i+1}\right) - \mathbb{H}\left(Z_i|X^n\tilde{Z}^{i+1}\right)\right)$$
$$\geqslant \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{H}\left(Z_i|Y^{i-1}\tilde{Z}^{i+1}\right) - \mathbb{H}\left(Z_i|Y^{i-1}\tilde{Z}^{i+1}X_i\right)\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\left(X_i; Z_i|Y^{i-1}\tilde{Z}^{i+1}\right), \tag{3}$$

where the inequality follows because conditioning does not increase entropy and $\tilde{Z}^{i+1}Y^{i-1} \to X_i \to Z_i$ 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 $U_i \triangleq Y^{i-1}\tilde{Z}^{i+1}$ and $V_i \triangleq U_iM$. Combining inequalities (1), (2), and (3), and substituting the definition of Q, U_i, V_i above, we obtain

$$R \leqslant \mathbb{I}(V_Q; Y_Q|QU_Q) - \mathbb{I}(V_Q; Z_Q|QU_Q) + \delta(\epsilon) \tag{4}$$
$$\mathbb{H}(R) \geqslant \mathbb{I}(X_Q; Z_Q|QU_Q) - \delta(\epsilon). \tag{5}$$

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

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

Because the eavesdropper's channel is less capable, then $\mathbb{I}(V; Y|U) - \mathbb{I}(V; Z|U) \leqslant \mathbb{I}(X; Y|U) - \mathbb{I}(X; Z|U)$. 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 $T_\epsilon^n(X)$ and the set of conditional ϵ-strongly typical sequence with respect to p_{YX} and $x^n \in T_\epsilon^n(X)$ by $T_\epsilon^n(Y|x^n)$.

We first show the existence of a code 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] $\mathbb{I}(X; Z|U) \leqslant \mathbb{H}(R)$ and $\mathbb{I}(X; Y|U) - \mathbb{I}(X; Z|U) > 0$, and we construct a code C_n for the broadcast channel with confidential messages $(\mathcal{X}, p_{YZ|X}, \mathcal{Y} \times \mathcal{Z})$. 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

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

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 $p_{X|U}$, 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 $U_r \in [\![1, 2^{nR_r}]\!]$, 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 decode (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 $R_0 < \min(\mathbb{I}(U; Y), \mathbb{I}(U; Z))$ and $R + R_r < \mathbb{I}(X; Y|U)$, then $\mathbb{E}(P_e(C_n)) \leqslant 2^{-\alpha n}$ for some $\alpha > 0$.

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

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

Proof: Lemma 2 is a special case of Lemma 4 proved in Appendix C. ∎

Using Markov's inequality, we conclude that there exists at least one code C_n satisfying the rate inequalities in Lemma 1 and Lemma 2, such that $P_e(C_n) \leqslant 3 \cdot 2^{-\alpha n}$ and $\mathbb{V}(p_{MM_0Z^n}, p_Mp_{M_0Z^n}) \leqslant 3 \cdot 2^{-\beta n}$. Finally, the uniform numbers U_r can be approximately obtained from $(\mathcal{R}, p_R)$ with an appropriate function ϕ.

Lemma 3 (adapted from [11]): If $R_r < \mathbb{H}(R)$, then there exists ϕ such that $\mathbb{V}(p_{\phi(R^n)}, p_{U_r}) \leqslant 2^{-n\eta}$ for some $\eta > 0$. Consequently, it is not hard to show that, even if the code C_n is used with $\phi(R^n)$ in place of U_r, then

$$P_e(C_n) \leqslant 2^{-\kappa n} \quad \text{and} \quad \mathbb{V}(p_{MM_0Z^n}, p_Mp_{M_0Z^n}) \leqslant 2^{-\kappa n}.$$

for some $\kappa > 0$. The fact that $L(C_n) \leqslant 2^{-\kappa' n}$ 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 $R < \mathbb{I}(X; Y|U) - \mathbb{I}(X; Z|U)$ such that $\mathbb{I}(X; Z|U) \leqslant \mathbb{H}(R)$ 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 3

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 U_r uniform.

Lemma 4: If $\frac{1}{n}R_2(U_r) > \mathbb{I}(X; Z|U)$, then we have $\mathbb{E}_{C_n}(\mathbb{V}(p_{MZ^n}, p_Mp_{Z^n})) \leqslant 2^{-\beta n}$ for some $\beta > 0$.
The proof relies on a careful analysis and modification of the "cloud-mixing" lemma [13] and the notation is that of Appendix B. We define the distribution $q_{U^nX^nZ^n}$ on $\mathcal{U}^n \times \mathcal{X}^n \times \mathcal{Z}^n$ as

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

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

$$\mathbb{V}(p_{\mathsf{M}Z^n}, p_{\mathsf{M}}p_{Z^n})$$
$$\leqslant \mathbb{V}(p_{\mathsf{M}\mathsf{U}^n Z^n}, p_{\mathsf{M}}p_{\mathsf{U}^n Z^n})$$
$$= \mathbb{E}_{\mathsf{U}^n \mathsf{M}}\big(\mathbb{V}(p_{Z^n|\mathsf{M}\mathsf{U}^n}, p_{Z^n|\mathsf{U}^n})\big)$$
$$\leqslant \mathbb{E}_{\mathsf{U}^n \mathsf{M}}\big(\mathbb{V}(p_{Z^n|\mathsf{M}\mathsf{U}^n}, q_{Z^n|\mathsf{U}^n}) + \mathbb{V}(q_{Z^n|\mathsf{U}^n}, p_{Z^n|\mathsf{U}^n})\big)$$
$$\leqslant 2\mathbb{E}_{\mathsf{U}^n \mathsf{M}}\big(\mathbb{V}(p_{Z^n|\mathsf{M}\mathsf{U}^n}, q_{Z^n|\mathsf{U}^n})\big)$$

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 $\mathbb{V}(p_{\mathsf{M}Z^n}, p_{\mathsf{M}}p_{Z^n})$ over randomly generated codes C_n satisfies

$$\mathbb{E}_{C_n}(\mathbb{V}(p_{\mathsf{M}Z^n}, p_{\mathsf{M}}p_{Z^n}))$$
$$\leqslant 2\mathbb{E}_{C_n}\big(\mathbb{V}(p_{Z^n|\mathsf{U}^n=\mathsf{U}_1^n \mathsf{M}=1}, q_{Z^n|\mathsf{U}^n=\mathsf{U}_1^n})\big),$$

where

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

The average over the random codes can be split between the average of U_1^n and the random code $C_n(u_1^n)$ for a fixed value of u_1^n, so that

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

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

$$\mathbb{E}_{C_n(u_1^n)}\big(p_{Z^n|\mathsf{U}^n=u_1^n \mathsf{M}=1}(z^n)\big)$$
$$= \mathbb{E}_{C_n(u_1^n)}\left(\sum_{k=1}^{2^{nR_r}} W_{Z^n|X^n}(z^n|x^n(1,1,k))p_{\mathsf{U}_r}(k)\right)$$
$$= \sum_{k=1}^{2^{nR_r}} \mathbb{E}_{C_n(u_1^n)}\big(W_{Z^n|X^n}(z^n|x^n(1,1,k))\big)p_{\mathsf{U}_r}(k)$$
$$= q_{Z^n|\mathsf{U}^n=u_1^n}(z^n).$$

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

$$p^{(1)}(z^n) \triangleq \sum_{k=1}^{2^{nR_r}} W_{Z^n|X^n}(z^n|x^n(1,1,k))p_{\mathsf{U}_r}(k)$$
$$\mathbf{1}\{(x^n(1,1,k), z^n) \in T_\epsilon^n(\mathsf{X}\mathsf{Z}|u_1^n)\},$$

$$p^{(2)}(z^n) \triangleq \sum_{k=1}^{2^{nR_r}} W_{Z^n|X^n}(z^n|x^n(1,1,k))p_{\mathsf{U}_r}(k)$$
$$\mathbf{1}\{(x^n(1,1,k), z^n) \notin T_\epsilon^n(\mathsf{X}\mathsf{Z}|u_1^n)\},$$

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

$$\mathbb{V}(p_{Z^n|\mathsf{U}^n=u_1^n \mathsf{M}=1}, q_{Z^n|\mathsf{U}^n=u_1^n})$$
$$\leqslant \sum_{z^n \notin T_\epsilon^n(\mathsf{Z}|u_1^n)} \big|p_{Z^n|\mathsf{U}^n=u_1^n \mathsf{M}=1}(z^n) - q_{Z^n|\mathsf{U}^n=u_1^n}(z^n)\big| \quad (6)$$
$$+ \sum_{z^n \in T_\epsilon^n(\mathsf{Z}|u_1^n)} \big|p^{(1)}(z^n) - \mathbb{E}\big(p^{(1)}(z^n)\big)\big| \quad (7)$$
$$+ \sum_{z^n \in T_\epsilon^n(\mathsf{Z}|u_1^n)} \big|p^{(2)}(z^n) - \mathbb{E}\big(p^{(2)}(z^n)\big)\big|. \quad (8)$$

Taking the expectation of the term in (6) over $C_n(u_1^n)$, we obtain

$$\mathbb{E}\left(\sum_{z^n \notin T_\epsilon^n(\mathsf{Z}|u_1^n)} \big|p_{Z^n|\mathsf{U}^n=u_1^n \mathsf{M}=1}(z^n) - q_{Z^n|\mathsf{U}^n=u_1^n}(z^n)\big|\right)$$
$$\leqslant \sum_{z^n \notin T_\epsilon^n(\mathsf{Z}|u_1^n)} \mathbb{E}\big(\max(p_{Z^n|\mathsf{U}^n=u_1^n \mathsf{M}=1}(z^n), q_{Z^n|\mathsf{U}^n=u_1^n}(z^n))\big)$$
$$= \sum_{z^n \notin T_\epsilon^n(\mathsf{Z}|u_1^n)} q_{Z^n|\mathsf{U}^n=u_1^n}(z^n),$$

which vanishes as n goes to infinity. Similarly, taking the expectation of the term in (8) over $C_n(u_1^n)$, we obtain

$$\mathbb{E}\left(\sum_{z^n \in T_\epsilon^n(\mathsf{Z}|u_1^n)} \big|p^{(2)}(z^n) - \mathbb{E}\big(p^{(2)}(z^n)\big)\big|\right)$$
$$\leqslant \mathbb{E}\left(\sum_{z^n \in \mathcal{Z}^n} \big|p^{(2)}(z^n) - \mathbb{E}\big(p^{(2)}(z^n)\big)\big|\right)$$
$$\leqslant \sum_{z^n \in \mathcal{Z}^n} \mathbb{E}\big(p^{(2)}(z^n)\big)$$
$$= \sum_{z^n \in \mathcal{Z}^n} \mathbb{E}\big(W_{Z^n|X^n}(z^n|X^n(1,1,1))$$
$$\mathbf{1}\{(X^n(1,1,1), z^n) \notin T_\epsilon^n(\mathsf{X}\mathsf{Z}|u_1^n))$$
$$= \sum_{(x^n, z^n) \notin T_\epsilon^n(\mathsf{X}\mathsf{Z}|u_1^n)} q_{Z^n X^n|\mathsf{U}^n=u_1^n}(z^n, x^n),$$

which vanishes as n goes to infinity. Finally, we focus on the expectation of the term in (7) over $C_n(u_1^n)$. 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

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

In addition,

$$\mathrm{Var}\big(p^{(1)}(z^n)\big) = \sum_{k=1}^{2^{nR_r}} p_{\mathsf{U}_r}(k)^2 \mathrm{Var}\big(W_{Z^n|X^n}(z^n|X^n(1,1,k))$$
$$\mathbf{1}\{(X^n(1,1,k), z^n) \in T_\epsilon^n(\mathsf{X}\mathsf{Z}|u_1^n)\}\big)$$

Note that

$$\mathrm{Var}\left(W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|\mathsf{X}^n(1,1,k))\mathbf{1}\{(\mathsf{X}^n(1,1,k),z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)\}\right)$$

$$= \sum_{x^n\in\mathcal{X}^n} p_{\mathsf{X}^n|\mathsf{U}^n=u_1^n}(x^n)$$

$$\left(W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)\mathbf{1}\{(x^n,z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)\}\right)^2$$

$$= \sum_{x^n:(x^n,z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)} p_{\mathsf{X}^n|\mathsf{U}^n=u_1^n}(x^n)W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)^2$$

$$\overset{(a)}{\leqslant} 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})-\delta(\epsilon))}$$

$$\sum_{x^n:(x^n,z^n)\in T_\epsilon^n(\mathsf{XZ}|u_1^n)} p_{\mathsf{X}^n|\mathsf{U}^n=u_1^n}(x^n)W_{\mathsf{Z}^n|\mathsf{X}^n}(z^n|x^n)$$

$$\leqslant 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})-\delta(\epsilon))}q_{\mathsf{Z}^n|\mathsf{U}^n=u_1^n}(z^n)$$

$$\overset{(b)}{\leqslant} 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon))},$$

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

$$\mathrm{Var}\left(p^{(1)}(z^n)\right) \leqslant 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon))}\sum_{k=1}^{2^{nR_r}} p_{\mathsf{U}_r}(k)^2$$

$$\leqslant 2^{-n(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon))+\frac{R_2(\mathsf{U}_r)}{n}}.$$

and

$$\sum_{z^n\in T_\epsilon^n(\mathsf{Z}|u_1^n)} \mathbb{E}\left(\left|p^{(1)}(z^n)-\mathbb{E}\left(p^{(1)}(z^n)\right)\right|\right)$$

$$\leqslant 2^{n\mathbb{H}(\mathsf{Z}|\mathsf{U})}2^{-\frac{n}{2}(\mathbb{H}(\mathsf{Z}|\mathsf{X})+\mathbb{H}(\mathsf{Z}|\mathsf{U})-\delta(\epsilon)+\frac{R_2(\mathsf{U}_r)}{n})}$$

$$= 2^{-\frac{n}{2}(\frac{R_2(\mathsf{U}_r)}{n}-\mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U})-\delta(\epsilon))}$$

Hence, if $\frac{R_2(\mathsf{U}_r)}{n} > \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) + \delta(\epsilon)$, the sum vanishes as n goes to infinity, which concludes the proof. Note that if U_r is uniform, then $R_2(\mathsf{U}_r) = nR_r$, 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] 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, 2011.

[6] 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.

[7] 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.

[8] S. Watanabe and Y. Oohama, "Broadcast Channels with Confidential Messages by Randomness Constrained Stochastic Encoder," preprint, January 2012. [Online]. Available: arXiv:1201.6468

[9] E. Tekin and A. Yener, "The General Gaussian Multiple-Access and Two-Way Wiretap Channels: Achievable Rates and Cooperative Jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, June 2008.

[10] T. S. Han and S. Verdu, "The resolvability and the capacity of AWGN channels are equal," in *Proc. Symp. IEEE Int Information Theory*, 1994.

[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.