

# AN INTRODUCTION TO SECRECY CAPACITY

BRIAN DUNN

## 1. OVERVIEW

This paper introduces the reader to several information theoretic aspects of covert communications. In particular, it discusses fundamental limits on the amount of information that can be reliably communicated in such a way that a malevolent third-party observer cannot decode the messages. This situation is considered first for a system with a single intended receiver, and then extensions of secrecy capacity to more general multi-terminal settings are presented. We begin in Section 2 with a discussion of the wiretap channel introduced by Wyner in [1], along with the classical generalization due to Csiszár<sup>1</sup> and Körner [2]. Next, Section 3 ties the older work in with topics receiving recent attention through a discussion of network secrecy capacity and common randomness. The presentation of newer results centers around Csiszár and Narayan's contemporary work relating secrecy generation to multiterminal source coding [3], which in turn relies on some interesting results from Ahlswede [4, 5].

## 2. POINT-TO-POINT SECRECY CAPACITY

Wyner was the first to study reliably sending information over a DMC such that a secondary 'wiretap' observer could not decode the message [1]. As depicted in Figure 1, in Wyner's model the wire-tapper has access to information that passes through both the main DMC and an additional wiretap channel. Thus, the composite (main + wiretap) channel is an explicit degraded version of the main channel in the traditional sense that the channel law factors as  $p(z^n|x^n) = p(z^n|y^n)p(y^n|x^n)$ . Although Wyner refers to the secondary observer as

---

*Date:* December 13, 2006.

<sup>1</sup>Pronounced *CHEE-sar*.

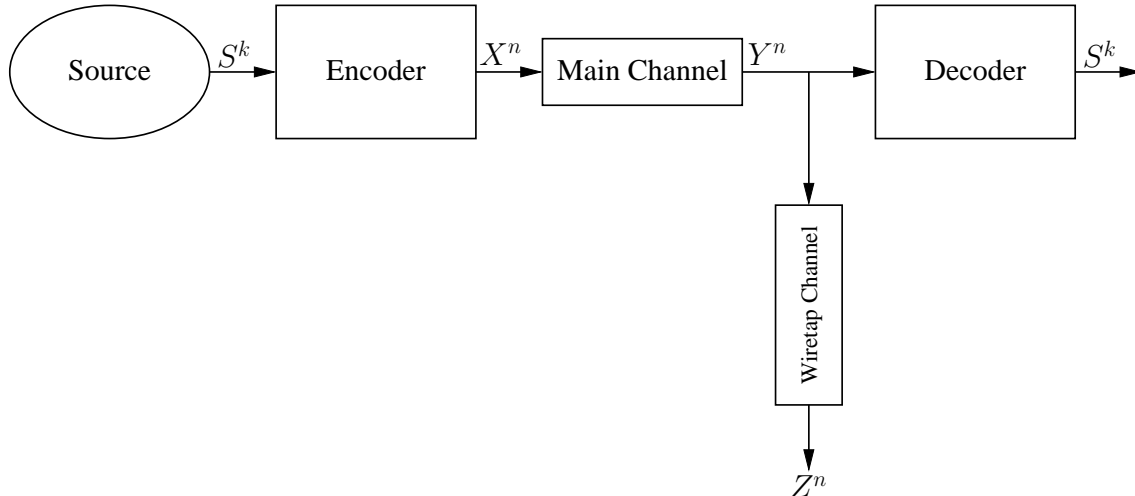


FIGURE 1. Wyner’s wiretap model.

a ‘wire-tapper,’ all terminals are aware of his presence a priori.<sup>2</sup> These results were later extended to Gaussian channels by Leung-Yan-Cheong and Hellman in [6] and also generalized to arbitrary broadcast channels  $p(y^n, z^n|x^n)$  by Csiszár and Körner in [2]. Because Wyner’s model is strictly generalized by [2], we focus on the latter work. In fact, in [2], an additional common message can be sent to both receivers.

Consider a private message  $S \in \mathcal{S}$  and a public message  $T \in \mathcal{T}$  which we wish to reliably transmit over a single-input two-output broadcast channel with general probability law  $p(y^n, z^n|x^n)$  — called a broadcast channel with confidential messages (BCC) — as depicted in Figure 2. The intended receiver for the private message observes the top channel output  $y^n$ , and the other party has access to the bottom channel output  $z^n$ . The goal is to design an encoder mapping  $f : \mathcal{S} \times \mathcal{T} \rightarrow \mathcal{X}^n$  and a pair of decoder mappings  $\phi : \mathcal{Y}^n \rightarrow \mathcal{S} \times \mathcal{T}$  and  $\psi : \mathcal{Z}^n \rightarrow \mathcal{T}$ , such that the private message can be exclusively recovered by the first terminal, and the public message can be recovered at both terminals.

Because we are concerned with secrecy, a stochastic encoder mapping may do better than a deterministic one. Accordingly, a (stochastic) encoder  $f$  of block length  $n$  is specified by a matrix of conditional probabilities  $f(x^n|s, t)$ . There is no advantage to a random decoder, so this case is not explicitly considered. Following the notation of [2], let the conditional

<sup>2</sup>A traditional wiretap is rendered useless once it is discovered by the party being monitored.

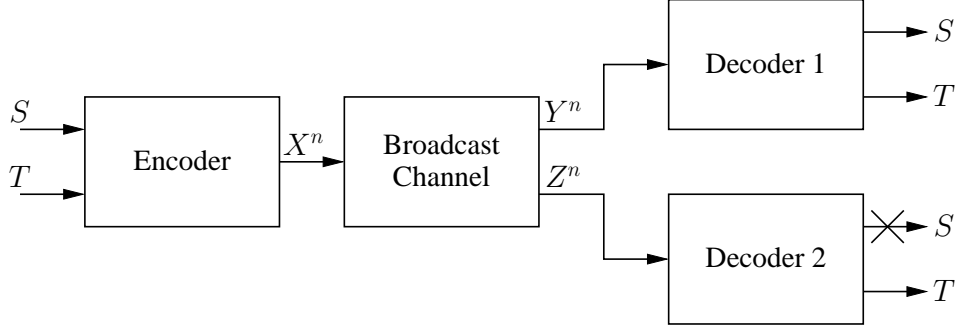


FIGURE 2. Csiszár and Körner's model as used in [2].

probability of error given the channel input  $x^n$  be denoted as  $P_{Y|X}^n[\phi(y^n) = (s, t)|x^n]$  and  $P_{Z|X}^n[\psi(z^n) = t|x^n]$  for the first and second decoders, respectively. If for every  $S \in \mathcal{S}$  and  $T \in \mathcal{T}$  the probability of error for each decoder, averaged over the stochastic encoder mapping and the channel distribution, is bounded by  $\epsilon$ , i.e.,

$$\sum_{x^n \in \mathcal{X}^n} f(x^n|s, t) P_{Y|X}^n[\phi(y^n) \neq (s, t)|x^n] \leq \epsilon \quad (1)$$

and

$$\sum_{x^n \in \mathcal{X}^n} f(x^n|s, t) P_{Z|X}^n[\psi(z^n) \neq t|x^n] \leq \epsilon, \quad (2)$$

then the encoder-decoder triple  $(f, \phi, \psi)$  is said to give rise to an  $(n, \epsilon)$ -transmission over the BCC. Furthermore, if sequences of such mappings exist along with message sets  $\mathcal{S}^n$  and  $\mathcal{T}^n$  and  $\epsilon_n \rightarrow 0$  such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|\mathcal{S}^n\| = R_1 \quad (3)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|\mathcal{T}^n\| = R'_0 \quad (4)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} H(S^n|Z^n) \geq R_e \quad (5)$$

then  $(f_n, \phi_n, \psi_n)$  is said to be an achievable rate triple for the BCC. The set of achievable rate triples  $(R_1, R_e, R_0)$  satisfying (3) is denoted as  $\mathcal{R}$ , where  $R_1$  and  $R_0$  are the private and common message rates, respectively, which are achieved at an equivocation rate  $R_e$ .

We are now in a position to state the main result of [2].

**Theorem 1.**  $\mathcal{R}$  is a closed convex set consisting of all rate triples  $(R_1, R_e, R_0)$  for which there exists auxiliary RVs  $U$  and  $V$  satisfying the Markov relationship  $U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z)$  and

$$0 \leq R_e \leq R_1 \tag{6}$$

$$R_e \leq I(V; Y|U) - I(V; Z|U) \tag{7}$$

$$R_1 + R_0 \leq I(V; Y|U) + \min [I(U; Y), I(U; Z)] \tag{8}$$

$$0 \leq R_0 \leq \min [I(U; Y), I(U; Z)]. \tag{9}$$

The proof of Theorem 1 is omitted here, but can be found in [2]. Instead we now turn our attention to some interesting special cases. If the amount of uncertainty about  $S^n$  after observing  $Z^n$ , i.e., the equivocation rate, equals  $H(S|T)$ , then  $S$  is communicated in *perfect secrecy*. If we do not transmit a common message so that  $R_0 = 0$ , then the achievable rate region reduces to

$$\mathcal{R}_{1e} = \{(R_1, R_e) : (R_1, R_e, 0) \in \mathcal{R}\} \tag{10}$$

The secrecy capacity is then defined as

$$C_s := \max_{(R_1, R_e) \in \mathcal{R}_{1e}} R_1. \tag{11}$$

Note that in (11), the maximization is over rate pairs such that  $R_e = R_1$ , i.e., rate pairs satisfying the condition for perfect secrecy.

As was first shown by Wyner [1], it is actually often possible to communicate in perfect secrecy. Surprisingly, for the more general model currently under consideration, this turns out to be true for quite a large class of channels. In fact, all that is needed to guarantee a positive secrecy capacity is that channel 2 is not ‘more noisy’ than channel 1, in the sense that

$$I(V; Y) \geq I(V; Z) \Rightarrow C_s > 0. \tag{12}$$

Wyner's results rely on the more restrictive condition of channel degradedness, but are interesting nonetheless. As a corollary to Theorem 1 we have

**Corollary 1.**  $(R_1, R_e) \in \mathcal{R}_{1e}$  if and only if there exists auxiliary RVs  $U$  and  $V$  with  $U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z)$  such that  $I(U; Y) \leq I(U; Z)$  and

$$0 \leq R_e \leq I(V; Y|U) - I(V; Z|U). \quad (13)$$

The secrecy capacity is then given by

$$C_s = \max_{V \leftrightarrow X \leftrightarrow (Y, Z)} [I(V; Y) - I(V; Z)]. \quad (14)$$

If channel 1 is 'more capable' than channel 2 such that  $I(X; Y) \geq I(X; Z)$ , then the secrecy capacity further simplifies to

$$C_s = \max [I(X; Y) - I(X; Z)]. \quad (15)$$

This final expression for the secrecy capacity is essentially equivalent to Wyner's characterization. The significance of (15) is that it can be interpreted as the difference in capacity between the main channel and the composite channel seen by the wire-tapper.

Wyner's wiretap model was extended to Gaussian channels in [6], however the requirement that the wiretap channel is a degraded version of the main channel is still imposed. Csiszár and Körner state that generalizations of their work to arbitrary alphabets and stationary ergodic source pairs is straightforward.

### 3. MULTI-TERMINAL SECRECY CAPACITY

Despite the 'wire-tapper' misnomer given to the malevolent observer in Section 2, secrecy capacity for the point-to-point wiretap channel admitted to a heuristically pleasing definition. As we will see, even for simple multi-terminal network models, the appropriate definition of secrecy capacity depends critically on the extent to which information has been compromised. Unlike in the first section, this section exclusively considers noiseless channel models, allowing

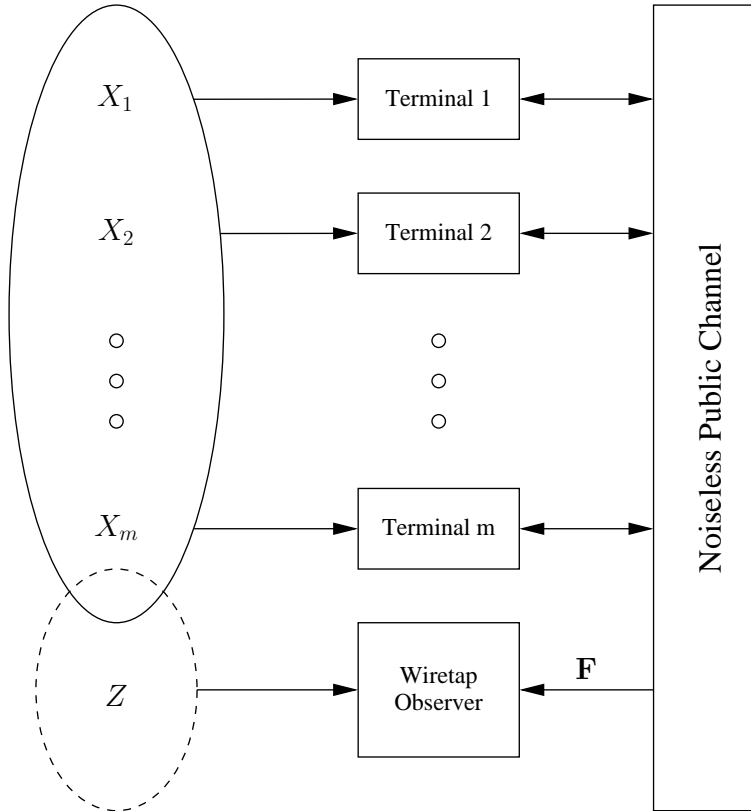


FIGURE 3. Csiszár and Narayan’s model as used in [3] for multi-terminal secret key generation with a malevolent observer.

the focus to be on generating common secrecy among terminals—which can then be used for encrypted communication. Naturally, each terminal must have access to some sort of information that can be used to generate a suitable key. In the model considered here, the way in which this information is shared is what determines the level of secrecy obtained.

Following the model of Csiszár and Narayan given in [3] and depicted in Figure 3, consider a collection of  $m$  terminals, each with access to i.i.d. repetitions of respective components of a discrete memoryless multiple source (DMMS)  $X_{\mathcal{M}} = (X_1, \dots, X_m)$ . During an observation interval of length of  $n$ , the  $i$ th terminal observes  $X_i^n, i \in \mathcal{M}$ , where  $\mathcal{M} = \{1, \dots, m\}$ . Let  $A \subset \mathcal{M}$  denote an arbitrary subset of benevolent terminals, leaving  $A^c = \mathcal{M} \setminus A$  as ‘helper terminals’ that can participate in key generation as desired, but do not ultimately care about the encryption key.

To facilitate communication among terminals they are allowed to exchange information over a noiseless public channel occurring in  $r$  consecutive rounds. The transmission protocol is described by the set of deterministic mappings  $f_\nu, \nu = 1, \dots, r \cdot m$ . Each mapping has a corresponding RV  $F_\nu = f_\nu(X_{\nu \bmod m}^n, F_1, \dots, F_{\nu-1})$ , with the vector of all such RVs denoted as  $\mathbf{F} = (F_1, \dots, F_{rm})$ . The channel carries the implicit limitation that the malevolent user is completely aware of all transmissions over the channel. As we will see, an interesting result of [3] is that interactive communication does no better than non-interactive communication in terms of secrecy capacity, i.e., only a single round of communication is necessary ( $r = 1$ ) and the mappings  $F_i = f_i(X_i^n), i \in \mathcal{M}$  suffice.

Given the above model, the ultimate goal is to exchange information among terminals in such a way that every terminal in the set  $A$  can agree upon a secret key (SK)  $K$ , while the malevolent user gains no knowledge of it, despite perfect observation of the public channel (the realization of  $\mathbf{F}$ ). To make this statement more precise and arrive at a suitable definition of secrecy capacity requires three things:

- An explicit formalization of what is meant by the terminals ‘agreeing on’ a secret key;
- An appropriate characterization of the malevolent user gaining ‘no knowledge of’ the secret key, i.e., secrecy; and
- A suitable definition of an achievable secret key rate.

The secrecy capacity then follows naturally as the supremum of achievable secret key rates.

Agreeing on a secret key is related to correctly decoding a codeword sent over a noisy discrete memoryless channel. As such, a function  $K$  of the DMMS  $X_{\mathcal{M}}^n$  is considered  *$\epsilon$ -common randomness* ( $\epsilon$ -CR) for a set of terminals  $A \subset \mathcal{M}$ , if<sup>3</sup>  $\Pr[K \neq g_i(X_i^n, \mathbf{F})] \leq \epsilon$  for some function<sup>4</sup>  $g_i$  and each  $i \in A$ .

Intuitively, a secret key should be difficult to guess both before and after observation of the public channel. Therefore, the RV  $K$  should have a somewhat uniform distribution over the

<sup>3</sup>We omit the intermediate definition of an  $\epsilon$ -recoverable RV used in [3] for brevity.

<sup>4</sup>In [3] the decoder functions are denoted as  $f$ , however we use  $g_i$  to avoid confusion with the mappings representing communication over the public channel.

set of possible keys  $\mathcal{K}$ , and the mutual information between  $K$  and  $\mathbf{F}$  should be vanishingly small. A function  $K$  of  $X_{\mathcal{M}}^n$  constitutes an  $\epsilon$ -secret key ( $\epsilon$ -SK) for a set of terminals  $A \subset \mathcal{M}$  if it is  $\epsilon$ -CR for  $A$  and satisfies both the uniformity condition

$$\frac{1}{n}H(K) \geq \frac{1}{n} \log |\mathcal{K}| - \epsilon \quad (16)$$

and the secrecy condition

$$\frac{1}{n}I(K; \mathbf{F}) \leq \epsilon. \quad (17)$$

The final prerequisite for defining the secret key capacity is a suitable definition of an achievable SK rate. Let  $H$  be an *achievable secret key rate* for a set of terminals  $A \subset \mathcal{M}$  if there exist  $\epsilon_n$ -SKs  $K^{(n)}$  such that

$$\epsilon_n \rightarrow 0 \quad \text{and} \quad \frac{1}{n}H(K^{(n)}) \rightarrow H \quad (18)$$

as  $n \rightarrow \infty$ . As stated earlier, the *secret key capacity*  $C_S(A)$  is now defined as the supremum of achievable SK rates for  $A$ .

Generalizing the model, the malevolent observer may have direct access to some form of side information about the source. To characterize this scenario the DMMS is augmented with an additional (possibly vector) RV  $Z$ , jointly distributed with  $X_{\mathcal{M}}$  and only available to the malevolent observer. The side information available to the wire-tapper is shown in Figure 3 as the dashed circle below the main DMMS. In this case, an  $\epsilon$ -SK for a set of terminals  $A \subset \mathcal{M}$  is also an  $\epsilon$ -wiretap secret key ( $\epsilon$ -WSK) if it satisfies the stronger secrecy condition

$$\frac{1}{n}I(K; \mathbf{F}, Z^n) \leq \epsilon. \quad (19)$$

Similarly, define an *achievable WSK rate* as above but with  $\epsilon_n$ -WSKs. Also define the *wireless secret key capacity*  $C_W(A)$  accordingly.

The above given definitions are actually inadequate for cryptography [7]; however, if  $\epsilon_n$  converges exponentially, the so-called *strongly achievable* rates and corresponding *strong secrecy capacity* are sufficient for this purpose.

As we will see, the secret key capacity admits to an intuitively pleasing interpretation, i.e., it is the total entropy of the source minus the amount of information that must be shared over the public channel for each terminal to gain complete knowledge of the source. An almost obvious lower bound on the SK capacity comes from performing Slepian-Wolf (SW) source coding across the terminals and sharing this information over the public channel. Although this represents a non-interactive scheme, it turns out that interactive communication can do no better. Consequently, the SW lower bound is tight and the SK capacity is achievable through Slepian-Wolf distributed compression. Of course, the difficulty lies in the proof of the converse—showing that no interactive scheme can outperform the SW strategy.

The Slepian-Wolf achievable rate region is given by

$$\mathcal{R}(A) = \left\{ R_{\mathcal{M}} : \sum_{i \in B} R_i \geq H(X_B | X_{B^c}), B \subset \mathcal{M}, A \not\subset B \right\}, \quad (20)$$

where

$$R_i = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|F_i^{(n)}\| \quad (21)$$

with  $\|F_i^{(n)}\|$  denoting the cardinality of the range of  $F_i^{(n)}$ . Now consider how much information must be exchanged such that each terminal has complete knowledge of the source  $X_{\mathcal{M}}^n$ , i.e., to obtain *common omniscience* (CO). Let  $R_{\text{CO}}(A)$  be the smallest  $R$  such that  $K = X_{\mathcal{M}}^n$  constitutes achievable  $\epsilon_n$ -CR for a set of terminals  $A \subset \mathcal{M}$  with

$$\epsilon_n \rightarrow 0 \quad \text{and} \quad \frac{1}{n} \log \|\mathbf{F}^{(n)}\| \rightarrow R, \quad (22)$$

then

$$R_{\text{CO}}(A) = \min_{R_{\mathcal{M}} \in \mathcal{R}(A)} \sum_{i=1}^m R_i. \quad (23)$$

Finally, the (strong) secret key capacity for a set of terminals  $A \subset \mathcal{M}$  is

$$C_S(A) = H(X_{\mathcal{M}}) - R_{\text{CO}}(A), \quad (24)$$

and can be achieved by non-interactive communication. The proof can be found in [3].

## REFERENCES

- [1] Aaron D. Wyner. The Wire-Tap Channel. *Bell Syst. Tech. J.*, 54(8):1355–1387, October 1975.
- [2] Imre Csiszár and János Körner. Broadcast Channels with Confidential Messages. *IEEE Trans. Inform. Theory*, 24(3):339–348, May 1978.
- [3] Imre Csiszár and Prakash Narayan. Secrecy Capacities for Multiple Terminals. *IEEE Trans. Inform. Theory*, 50(12):3047–3061, December 2004.
- [4] Rudolf F. Ahlswede and Imre Csiszár. Common Randomness in Information Theory and Cryptography—Part I: Secret Sharing. *IEEE Trans. Inform. Theory*, 39(4):1121–1132, July 1993.
- [5] Rudolf F. Ahlswede and Imre Csiszár. Common Randomness in Information Theory and Cryptography—Part II: CR Capacity. *IEEE Trans. Inform. Theory*, 44(1):225–240, January 1998.
- [6] S. K. Leung-Yan-Cheong and Martin E. Hellman. The Gaussian Wire-tap Channel. *IEEE Trans. Inform. Theory*, 24(4):451–456, July 1978.
- [7] Ueli M. Maurer. *The Strong Secret Key Rate of Discrete Random Triples*. Kluwer, Norwell, MA, 1994.