3 Secrecy capacity

In this chapter, we develop the notion of *secrecy capacity*, which plays a central role in physical-layer security. The secrecy capacity characterizes the fundamental limit of secure communications over noisy channels, and it is essentially the counterpart to the usual point-to-point channel capacity when communications are subject not only to reliability constraints but also to an information-theoretic secrecy requirement. It is inherently associated with a channel model called the *wiretap channel*, which is a broadcast channel in which one of the receivers is treated as an adversary. This adversarial receiver, which we call the *eavesdropper* to emphasize its passiveness, should remain ignorant of the messages transmitted over the channel. The mathematical tools, and especially the random-coding argument, presented in this chapter are the basis for most of the theoretical research in physical-layer security, and we use them extensively in subsequent chapters.

We start with a review of Shannon's model of secure communications (Section 3.1), and then we informally discuss the problem of secure communications over noisy channels (Section 3.2). The intuition we develop from loose arguments is useful to grasp the concepts underlying the proofs of the secrecy capacity and motivates a discussion of the choice of an information-theoretic secrecy metric (Section 3.3). We then study in detail the fundamental limits of secure communication over degraded wiretap channels (Section 3.4) and broadcast channels with confidential messages (Section 3.5). We also discuss the multiplexing of secure and non-secure messages as well as the role of feedback for securing communications (Section 3.6). Finally, we conclude the chapter with a summary of the lessons learned from the analysis of fundamental limits and a review of the explicit and implicit assumptions used in the models (Section 3.7). The Gaussian wiretap channel and its extensions to multiple-input multiple-output channels and wireless channels are considered separately in Chapter 5.

3.1 Shannon's cipher system

Shannon proposed the idea of measuring quantitatively the secrecy level of encryption systems on the basis of his mathematical theory of communication. Shannon's model of secure communications, which is often called *Shannon's cipher system*, is illustrated in Figure 3.1; it considers a situation in which a transmitter communicates with a legitimate receiver over a noiseless channel, while an eavesdropper overhears all signals sent over

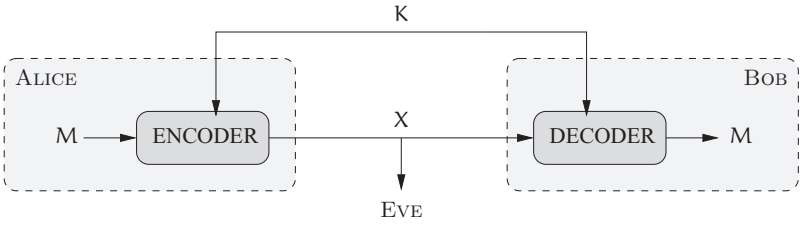

Figure 3.1 Shannon's cipher system.

the channel. To prevent the eavesdropper from retrieving information, the transmitter encodes his messages into codewords by means of a secret key, which is known to the legitimate receiver but unknown to the eavesdropper.[1] Messages, codewords, and keys are represented by the random variables $\mathsf{M} \in \mathcal{M}$, $\mathsf{X} \in \mathcal{X}$, and $\mathsf{K} \in \mathcal{K}$, respectively, and we assume that K is independent of M. The encoding function is denoted by $e : \mathcal{M} \times \mathcal{K} \to \mathcal{X}$, the decoding function is denoted by $d : \mathcal{X} \times \mathcal{K} \to \mathcal{M}$, and we refer to the pair (e, d) as a *coding scheme*. The legitimate receiver is assumed to retrieve messages without error, that is

$$\mathsf{M} = d(\mathsf{X}, \mathsf{K}) \quad \text{if} \quad \mathsf{X} = e(\mathsf{M}, \mathsf{K}).$$

Although the eavesdropper has no knowledge about the secret key, he is assumed to know the encoding function e and the decoding function d.

To measure secrecy with respect to Eve in terms of an information-theoretic quantity, it is natural to consider the conditional entropy $\mathbb{H}(\mathsf{M}|\mathsf{X})$, which we call the eavesdropper's *equivocation*. Intuitively, the equivocation represents Eve's uncertainty about the messages after intercepting the codewords. A coding scheme is said to achieve *perfect secrecy* if

$$\mathbb{H}(\mathsf{M}|\mathsf{X}) = \mathbb{H}(\mathsf{M}) \qquad \text{or, equivalently,} \qquad \mathbb{I}(\mathsf{M};\mathsf{X}) = 0.$$

We call the quantity $\mathbb{I}(\mathsf{M};\mathsf{X})$ the *leakage* of information to the eavesdropper. In other words, perfect secrecy is achieved if codewords X are statistically independent of messages M. This definition of security differs from the traditional assessment based on computational complexity not only because it provides a *quantitative* metric to measure secrecy but also because it disregards the computational power of the eavesdropper. Perfect secrecy guarantees that the eavesdropper's optimal attack is to guess the message M at random and that there exists no algorithm that could extract any information about M from X.

Proposition 3.1. *If a coding scheme for Shannon's cipher system achieves perfect secrecy, then*

$$\mathbb{H}(\mathsf{K}) \geqslant \mathbb{H}(\mathsf{M}).$$

[1] In cryptography, it is customary to call a message a *plaintext*, and a codeword a *ciphertext* or a *cryptogram*. We adopt instead the nomenclature prevalent in information theory and coding theory.

Proof. Consider a coding scheme that achieves perfect secrecy. By assumption, $\mathbb{H}(\mathsf{M}|\mathsf{X}) = \mathbb{H}(\mathsf{M})$; in addition, since messages M are decoded without errors upon observing X and K, Fano's inequality also ensures that $\mathbb{H}(\mathsf{M}|\mathsf{XK}) = 0$. Consequently,

$$\begin{aligned}\mathbb{H}(\mathsf{K}) &\overset{\text{(a)}}{\geqslant} \mathbb{H}(\mathsf{K}) - \mathbb{H}(\mathsf{K}|\mathsf{XM}) \\ &\overset{\text{(b)}}{\geqslant} \mathbb{H}(\mathsf{K}|\mathsf{X}) - \mathbb{H}(\mathsf{K}|\mathsf{XM}) \\ &= \mathbb{I}(\mathsf{K};\mathsf{M}|\mathsf{X}) \\ &= \mathbb{H}(\mathsf{M}|\mathsf{X}) - \mathbb{H}(\mathsf{M}|\mathsf{KX}) \\ &= \mathbb{H}(\mathsf{M}|\mathsf{X}) \\ &= \mathbb{H}(\mathsf{M}).\end{aligned}$$

Inequality (a) follows from $\mathbb{H}(\mathsf{K}|\mathsf{XM}) \geqslant 0$ and inequality (b) follows from $\mathbb{H}(\mathsf{K}) \geqslant \mathbb{H}(\mathsf{K}|\mathsf{X})$ because conditioning does not increase entropy. □

In other words, Proposition 3.1 states that it is necessary to use at least one secret-key bit for each message bit to achieve perfect secrecy. If the number of possible messages, keys, and codewords is the same, it is possible to obtain a more precise result and to establish necessary and sufficient conditions for communication in perfect secrecy.

Theorem 3.1. *If $|\mathcal{M}| = |\mathcal{X}| = |\mathcal{K}|$, a coding scheme for Shannon's cipher system achieves perfect secrecy if and only if*

- *for each pair $(m, x) \in \mathcal{M} \times \mathcal{X}$, there exists a unique key $k \in \mathcal{K}$ such that $x = e(m, k)$;*
- *the key K is uniformly distributed in $\mathcal{K}$.*

Proof. First, we establish that the conditions of Theorem 3.1 are necessary. Consider a coding scheme that achieves perfect secrecy with $|\mathcal{M}| = |\mathcal{X}| = |\mathcal{K}|$. Note that $p_{\mathsf{X}}(x) > 0$ for all $x \in \mathcal{X}$, otherwise some codewords would never be used and could be removed from $\mathcal{X}$, which would violate the assumption $|\mathcal{M}| = |\mathcal{X}|$. Since M and X are independent, this implies $p_{\mathsf{X}|\mathsf{M}}(x|m) = p_{\mathsf{X}}(x) > 0$ for all pairs $(m, x) \in \mathcal{M} \times \mathcal{X}$. In other words, for all messages $m \in \mathcal{M}$, the encoder can output all possible codewords in $\mathcal{X}$; therefore,

$$\forall m \in \mathcal{M} \quad \mathcal{X} = \{e(m, k) : k \in \mathcal{K}\}.$$

Because we have assumed $|\mathcal{X}| = |\mathcal{K}|$, for all pairs $(m, x) \in \mathcal{M} \times \mathcal{X}$ there must be a unique key $k \in \mathcal{K}$ such that $x = e(m, k)$. Now, fix an arbitrary codeword $x^* \in \mathcal{X}$. For every message $m \in \mathcal{M}$, let k_m be the unique key such that $x^* = e(m, k_m)$. Then $p_{\mathsf{K}}(k_m) = p_{\mathsf{X}|\mathsf{M}}(x^*|m)$ and $\mathcal{K} = \{k_m : m \in \mathcal{M}\}$. Using Bayes' rule, we obtain

$$p_{\mathsf{K}}(k_m) = p_{\mathsf{X}|\mathsf{M}}(x^*|m) = \frac{p_{\mathsf{M}|\mathsf{X}}(m|x^*)\, p_{\mathsf{X}}(x^*)}{p_{\mathsf{M}}(m)} = p_{\mathsf{X}}(x^*),$$

where the last equality follows from $p_{\mathsf{M}|\mathsf{X}}(m|x^*) = p_{\mathsf{M}}(m)$ by virtue of the independence of M and X. Therefore, $p_{\mathsf{K}}(k_m)$ takes on the same value for all $m \in \mathcal{M}$, which implies that K is uniformly distributed in $\mathcal{K}$.

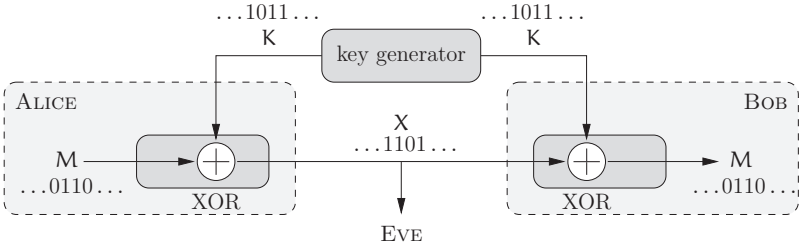

Figure 3.2 Vernam's cipher (one-time pad) illustrated for $\mathcal{M} = \{0, 1\}$.

We now show that the conditions of Theorem 3.1 are also sufficient. Since $|\mathcal{M}| = |\mathcal{X}| = |\mathcal{K}|$, we can assume without loss of generality that $\mathcal{M} = \mathcal{X} = \mathcal{K} = [\![0, |\mathcal{M}| - 1]\!]$. Consider now the coding scheme illustrated in Figure 3.2, called a *Vernam cipher* or *one-time pad*. To send a message $m \in \mathcal{M}$, Alice transmits $x = m \oplus k$, where k is the realization of a key K, which is independent of the message and with uniform distribution on $\mathcal{M}$, and $\oplus$ is the modulo-$|\mathcal{M}|$ addition. Since k is known to Bob, he can decode the message m from the codeword x without error by computing

$$x \oplus k = m \oplus k \ominus k = m,$$

where $\ominus$ is the modulo-$|\mathcal{M}|$ subtraction. In addition, this encoding procedure guarantees that, for all $x \in \mathcal{X}$,

$$p_{\mathsf{X}}(x) = \sum_{k \in \mathcal{M}} p_{\mathsf{X}|\mathsf{K}}(x|k) p_{\mathsf{K}}(k) = \sum_{k \in \mathcal{M}} p_{\mathsf{M}}(x \oplus k) \frac{1}{|\mathcal{M}|} = \frac{1}{|\mathcal{M}|},$$

and, consequently,

$$\begin{aligned} \mathbb{I}(\mathsf{M};\mathsf{X}) &= \mathbb{H}(\mathsf{X}) - \mathbb{H}(\mathsf{X}|\mathsf{M}) \\ &\overset{(a)}{=} \mathbb{H}(\mathsf{X}) - \mathbb{H}(\mathsf{K}|\mathsf{M}) \\ &\overset{(b)}{=} \mathbb{H}(\mathsf{X}) - \mathbb{H}(\mathsf{K}) \\ &= \log|\mathcal{M}| - \log|\mathcal{M}| \\ &= 0, \end{aligned}$$

where (a) follows from $\mathbb{H}(\mathsf{X}|\mathsf{M}) = \mathbb{H}(\mathsf{K}|\mathsf{M})$ because there is a one-to-one mapping between X and K given M and (b) follows from $\mathbb{H}(\mathsf{K}|\mathsf{M}) = \mathbb{H}(\mathsf{K})$ because M and K are independent. Notice that this result holds for any probability distribution of the message p_{M} for which $\forall m \in \mathcal{M} p_{\mathsf{M}}(m) > 0$. □

The fact that a one-time pad guarantees perfect secrecy is a result usually referred to as the "crypto lemma," which holds under very general conditions; in particular, the finite alphabet $\mathcal{M}$ can be replaced by a compact abelian group $\mathcal{G}$.[2]

[2] An abelian group $\mathcal{G}$ is a commutative group that need not be finite. The assumption that $\mathcal{G}$ is compact guarantees that its Haar measure is finite so that it is possible to define a uniform probability distribution over $\mathcal{G}$.

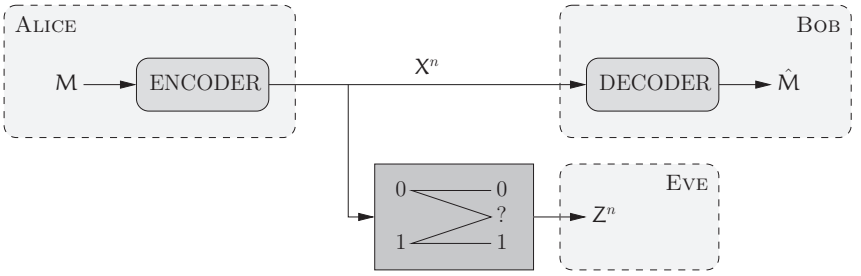

Figure 3.3 Communication over a binary erasure wiretap channel.

Lemma 3.1 (Crypto lemma, Forney)**.** *Let* $(\mathcal{G}, +)$ *be a compact abelian group with binary operation* $+$ *and let* $\mathsf{X} = \mathsf{M} + \mathsf{K}$*, where* M *and* K *are random variables over* $\mathcal{G}$ *and* K *is independent of* M *and uniform over* $\mathcal{G}$*. Then* X *is independent of* M *and uniform over* $\mathcal{G}$*.*

Although Theorem 3.1 shows the existence of coding schemes that achieve perfect secrecy, it provides an unsatisfactory result. In fact, since a one-time pad requires a new key bit for each message bit, it essentially replaces the problem of secure communication by that of secret-key distribution. Nevertheless, we show in the next sections that this disappointing result stems from the absence of noise at the physical layer in the model; in particular, Shannon's cipher system does not take into account the noise affecting the eavesdropper's observation of the codewords.

Remark 3.1. *Requiring perfect secrecy is much more stringent than preventing the eavesdropper from decoding correctly. To see this, assume for simplicity that messages are taken from the set* $[\![1, M]\!]$ *and that each of them is equally likely, in which case the eavesdropper minimizes his probability of decoding error* $\mathbf{P}_\mathrm{e}$ *by performing maximum-likelihood decoding. Since the a-priori distribution of the message* M *is uniform over* $[\![1, M]\!]$*, the condition* $\mathbb{H}(\mathsf{M}|\mathsf{X}) = \mathbb{H}(\mathsf{M})$ *ensures that* $p_{\mathsf{M}|\mathsf{X}}(m|x) = 1/M$ *for all messages* $m \in \mathcal{M}$ *and codewords* $x \in \mathcal{M}$ *or, equivalently, that the probability of error under maximum-likelihood decoding is* $\mathbf{P}_\mathrm{e} = (M-1)/M$*. In contrast, evaluating secrecy in terms of the non-decodability of the messages would merely guarantee that the probability of error under maximum-likelihood decoding is bounded away from zero, that is* $\mathbf{P}_\mathrm{e} > \epsilon$ *for some fixed* $\epsilon > 0$*.*

3.2 Secure communication over a noisy channel

Before we study secrecy capacity in detail, it is instructive to consider the effect of noise with the simple model illustrated in Figure 3.3, which is called a *binary erasure wiretap channel*. This channel is a special case of more general models that are studied in Section 3.4 and Section 3.5. Here, a transmitter communicates messages to a legitimate receiver by sending binary codewords of length n over a noiseless channel, while an eavesdropper observes a corrupted version of these codewords at the output of a binary

erasure channel with erasure probability $\epsilon \in (0, 1)$. Messages are taken from the set $[\![1, M]\!]$ uniformly at random, and are represented by the random variable M. Codewords are denoted by the random variable $\mathsf{X}^n \in \{0, 1\}^n$ and the eavesdropper's observation is denoted by $\mathsf{Z}^n \in \{0, 1, ?\}^n$. We assume that different messages are always encoded into different codewords, so that the reliable transmission rate is $(1/n)\mathbb{H}(\mathsf{M}) = (1/n)\log M$.

Rather than requiring perfect secrecy and exact statistical independence of M and X^n, we consider a more tractable condition and we say that a coding scheme is secure if it guarantees $\lim_{n\to\infty} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n) = 0$. The key difficulty is now that of how to determine the type of encoder that could enforce this condition. To obtain some insight, we consider a specific coding scheme for the model in Figure 3.3.

Example 3.1. Assume that messages are taken uniformly at random from the set $[\![1, 2]\!]$ so that $\mathbb{H}(\mathsf{M}) = 1$, and let n be arbitrary. Let $\mathcal{C}_1$ be the set of binary sequences of length n with odd parity and let $\mathcal{C}_2$ be the set of binary sequences of length n with even parity. To send a message $m \in \{1, 2\}$, the emitter transmits a sequence x^n chosen uniformly at random in $\mathcal{C}_m$. The rate of the coding scheme is simply $1/n$. Now, assume that the eavesdropper observes a sequence Z^n with k erasures. If $k > 0$, the parity of the erased bits is just as likely to be even as it is to be odd. If $k = 0$, the eavesdropper knows perfectly which codeword was sent and thus knows its parity. To analyze the eavesdropper's equivocation formally, we introduce the random variable $\mathsf{E} \in \{0, 1\}$ such that

$$\mathsf{E} = \begin{cases} 0 & \text{if } \mathsf{Z}^n \text{ contains no erasure;} \\ 1 & \text{otherwise.} \end{cases}$$

We can then lower bound the equivocation as

$$\begin{aligned} \mathbb{H}(\mathsf{M}|\mathsf{Z}^n) &\geqslant \mathbb{H}(\mathsf{M}|\mathsf{Z}^n\mathsf{E}) \\ &\overset{(a)}{=} \mathbb{H}(\mathsf{M}|\mathsf{Z}^n\mathsf{E} = 1)\,(1 - (1 - \epsilon)^n) \\ &= \mathbb{H}(\mathsf{M})\,(1 - (1 - \epsilon)^n) \\ &\overset{(b)}{=} \mathbb{H}(\mathsf{M}) - (1 - \epsilon)^n. \end{aligned}$$

Equality (a) follows from the fact that $\mathbb{H}(\mathsf{M}|\mathsf{Z}^n\mathsf{E} = 0) = 0$ and equality (b) follows from $\mathbb{H}(\mathsf{M}) = 1$. Hence, we obtain

$$\mathbb{I}(\mathsf{M}; \mathsf{Z}^n) = \mathbb{H}(\mathsf{M}) - \mathbb{H}(\mathsf{M}|\mathsf{Z}^n) \leqslant (1 - \epsilon)^n,$$

which vanishes exponentially fast with n; therefore, the coding scheme is secure.

In practice, the coding scheme of Example 3.1 is not really useful because the code rate vanishes with n as well, albeit more slowly than does $\mathbb{I}(\mathsf{M}; \mathsf{Z}^n)$; nevertheless, the example suggests that assigning multiple codewords to every message and selecting them randomly is useful to *confuse* the eavesdropper and to guarantee secrecy.

3.3 Perfect, weak, and strong secrecy

As mentioned in the previous section, the notion of perfect secrecy is too stringent and is not easily amenable to further analysis. It is convenient to replace the requirement of exact statistical independence between messages M and the eavesdropper's observations Z^n by *asymptotic* statistical independence as the codeword length n goes to infinity. In principle, this asymptotic independence can be measured in terms of any distance d defined on the set of joint probability distributions on $\mathcal{M} \times \mathcal{Z}^n$ as

$$\lim_{n\to\infty} d(p_{\mathsf{M}\mathsf{Z}^n}, p_{\mathsf{M}} p_{\mathsf{Z}^n}) = 0.$$

For instance, in the previous section we implicitly used the Kullback–Leibler divergence[3] and we required

$$\lim_{n\to\infty} \mathbb{D}(p_{\mathsf{M}\mathsf{Z}^n} \| p_{\mathsf{M}} p_{\mathsf{Z}^n}) = \lim_{n\to\infty} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n) = 0.$$

This condition, which we call the *strong secrecy condition*, requires the amount of information leaked to the eavesdropper to vanish. For technical purposes, it is also convenient to consider the condition

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n) = 0,$$

which requires only the *rate* of information leaked to the eavesdropper to vanish. This condition is weaker than strong secrecy since it is satisfied as long as $\mathbb{I}(\mathsf{M}; \mathsf{Z}^n)$ grows at most sub-linearly with n. We call it the *weak secrecy condition*.

From an information-theoretic perspective, the specific measure of asymptotic statistical independence may seem irrelevant, and we may be tempted to choose a metric solely on the basis of its mathematical tractability; unfortunately, the weak secrecy condition and the strong secrecy condition are not equivalent and, more importantly, it is possible to construct examples of coding schemes with evident security flaws that satisfy the weak secrecy condition.

Example 3.2. Let $n \geqslant 1$ and $t \triangleq \lfloor \sqrt{n} \rfloor$. Suppose that Alice encodes message bits $\mathsf{M}^n \in \{0,1\}^n$ into a codeword $\mathsf{X}^n \in \{0,1\}^n$ with $n-t$ secret-key bits $\mathsf{K}^{n-t} \in \{0,1\}^{n-t}$ as

$$\mathsf{X}_i = \begin{cases} \mathsf{M}_i \oplus \mathsf{K}_i & \text{for } i \in [\![1, n-t]\!], \\ \mathsf{M}_i & \text{for } i \in [\![n-t+1, n]\!]. \end{cases}$$

The key bits K_i for $i \in [\![1, n-t]\!]$ are assumed i.i.d. according to $\mathcal{B}(\frac{1}{2})$ and known to Bob. In other words, Alice performs a one-time pad of the first $n-t$ bits of M with the $n-t$ key bits and she appends the remaining t bits unprotected. Eve is assumed to intercept the codeword X^n directly.

[3] Strictly speaking, the Kullback–Leibler divergence is not a distance because it is not symmetric; nevertheless, $\mathbb{D}(p_{\mathsf{M}\mathsf{Z}^n} \| p_{\mathsf{M}} p_{\mathsf{Z}^n}) = 0$ if and only if M is independent of Z^n, and we ignore this subtlety.

Using the crypto lemma, we obtain

$$\forall n \geqslant 1 \qquad \mathbb{H}(\mathsf{M}|\mathsf{X}^n) = n - t = \mathbb{H}(\mathsf{M}) - t.$$

Therefore, $\mathbb{I}(\mathsf{M};\mathsf{X}^n) = t = \lfloor\sqrt{n}\rfloor$ and this scheme does not satisfy the strong secrecy criterion; even worse, the information leaked to the eavesdropper grows unbounded with n. However, notice that

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathsf{M};\mathsf{X}^n) = \lim_{n\to\infty} \frac{\lfloor\sqrt{n}\rfloor}{n} = 0.$$

Hence, this scheme satisfies the weak secrecy criterion.

Example 3.3. Suppose that Alice encodes messages $\mathsf{M} = (\mathsf{M}_1 \ldots \mathsf{M}_n)$ uniformly distributed on $\{0,1\}^n$ into codewords $\mathsf{X}^n \in \{0,1\}^n$ with secret keys $\mathsf{K}^n \in \{0,1\}^n$ as

$$\mathsf{X}_i = \mathsf{M}_i \oplus \mathsf{K}_i \quad \text{for } i \in [\![1, n]\!].$$

The secret key K^n, which we assume is known to Bob, is such that the all-zero n-bit sequence $\bar{0}^n$ has probability $1/n$ and all non-zero sequences are equally likely. Formally, the probability distribution of the secret key is

$$p_{\mathsf{K}^n}(k^n) = \begin{cases} \dfrac{1}{n} & \text{if } k^n = \bar{0}^n, \\ \dfrac{1 - 1/n}{2^n - 1} & \text{if } k^n \neq \bar{0}^n. \end{cases}$$

Since K^n is not uniformly distributed, this encryption scheme no longer guarantees perfect secrecy. As in the previous example, we assume that Eve directly intercepts X^n.

We first prove that this scheme satisfies the weak secrecy criterion. We introduce a random variable J such that

$$\mathsf{J} \triangleq \begin{cases} 0 & \text{if } \mathsf{K}^n = \bar{0}^n, \\ 1 & \text{otherwise.} \end{cases}$$

Since conditioning does not increase entropy, we can write

$$\begin{aligned} \mathbb{H}(\mathsf{M}|\mathsf{X}^n) &\geqslant \mathbb{H}(\mathsf{M}|\mathsf{X}^n\mathsf{J}) \\ &= \mathbb{H}(\mathsf{M}|\mathsf{X}^n, \mathsf{J}=0)p_{\mathsf{J}}(0) + \mathbb{H}(\mathsf{M}|\mathsf{X}^n, \mathsf{J}=1)p_{\mathsf{J}}(1). \end{aligned} \tag{3.1}$$

By definition, $\mathsf{K}^n = \bar{0}^n$ if $\mathsf{J} = 0$; hence, $\mathbb{H}(\mathsf{M}|\mathsf{X}^n, \mathsf{J}=0) = 0$ and we can restrict our attention to the term

$$\mathbb{H}(\mathsf{M}|\mathsf{X}^n, \mathsf{J}=1)p_{\mathsf{J}}(1) = -\sum_{m,x^n} p(m, x^n, j=1)\log p(m|x^n, j=1).$$

For any $m \in \{0,1\}^n$ and $x^n \in \{0,1\}^n$, the joint probability $p(m, x^n, j=1)$ can be written as

$$p(x^n, m, j=1) = p(m|x^n, j=1)p(x^n|j=1)p(j=1)$$

with

$$p(m|x^n, j=1) = \begin{cases} 0 & \text{if } m = x^n, \\ 1/(2^n-1) & \text{otherwise,} \end{cases}$$

$$p(x^n|j=1) = \frac{1}{2^n},$$

$$p(j=1) = 1 - \frac{1}{n}.$$

On substituting these values into (3.1), we obtain

$$\begin{aligned} \mathbb{H}(\mathsf{M}|\mathsf{X}^n) &\geqslant -\sum_{x^n} \sum_{m \neq x^n} \frac{1}{2^n - 1} \frac{1}{2^n} \left(1 - \frac{1}{n}\right) \log\left(\frac{1}{2^n - 1}\right) \\ &= -\left(1 - \frac{1}{n}\right) \log\left(\frac{1}{2^n - 1}\right) \\ &= \log(2^n - 1) - \frac{\log(2^n - 1)}{n} \\ &\geqslant \log(2^n - 1) - 1. \end{aligned}$$

Since $\mathbb{H}(\mathsf{M}) = n$, we obtain

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \mathbb{I}(\mathsf{M}; \mathsf{X}^n) &= 1 - \lim_{n\to\infty} \frac{1}{n} \mathbb{H}(\mathsf{M}|\mathsf{X}^n) \\ &\leqslant 1 - \lim_{n\to\infty} \frac{\log(2^n - 1) - 1}{n} \\ &= 0. \end{aligned}$$

Hence, this scheme satisfies the weak secrecy criterion. However,

$$\begin{aligned} \mathbb{H}(\mathsf{M}|\mathsf{X}^n) &= \mathbb{H}(\mathsf{X}^n \oplus \mathsf{K}|\mathsf{X}^n) \\ &= \mathbb{H}(\mathsf{K}|\mathsf{X}^n) \\ &\leqslant \mathbb{H}(\mathsf{K}) \\ &= -\frac{1}{n} \log\left(\frac{1}{n}\right) - (2^n - 1) \cdot \frac{1 - 1/n}{2^n - 1} \log\left(\frac{1 - 1/n}{2^n - 1}\right) \\ &= \mathbb{H}_\mathrm{b}(1/n) + (1 - 1/n)\log(2^n - 1) \\ &< \mathbb{H}_\mathrm{b}(1/n) + n - 1 \\ &< n - 0.5 \quad \text{for } n \text{ large enough.} \end{aligned}$$

Therefore, $\lim_{n\to\infty} \mathbb{I}(\mathsf{M}; \mathsf{X}^n) > 0.5$, and this scheme does not satisfy the strong secrecy criterion.

One could argue that Example 3.2 and Example 3.3 have been constructed *ad hoc* to exhibit flaws. In Example 3.2, the eavesdropper always obtains a fraction of the

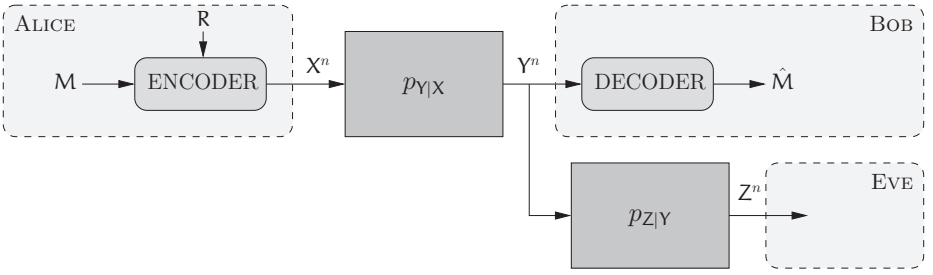

Figure 3.4 Communication over a DWTC. M represents the message to be transmitted securely, whereas R represents randomness used to randomize the encoder.

message bits without errors; in Example 3.3, the distribution of the key is skewed in such a way that the all-zero key is with overwhelming probability more likely than any other. Therefore, these examples do not imply that all weakly secure schemes are useless, but merely suggest that not all measures of asymptotic statistical independence are meaningful from a cryptographic perspective. In particular, this is a good indication that the weak secrecy criterion is likely not appropriate and, consequently, we should strive to prove all results with a strong secrecy criterion.

3.4 Wyner's wiretap channel

Secrecy capacity was originally introduced by Wyner for a channel model called a *degraded wiretap channel* (DWTC for short). Although this model is a special case of the broadcast channel with confidential messages studied in Section 3.5, it allows us to introduce many of the mathematical tools of information-theoretic security without the additional complexity of fully general models. As illustrated in Figure 3.4, a DWTC models a situation in which a sender (Alice) tries to communicate with a legitimate receiver (Bob) over a noisy channel, while an eavesdropper (Eve) observes a *degraded* version of the signal obtained by the legitimate receiver.

Formally, a discrete memoryless DWTC $\left(\mathcal{X}, p_{Z|Y}\, p_{Y|X}, \mathcal{Y}, \mathcal{Z}\right)$ consists of a finite input alphabet $\mathcal{X}$, two finite output alphabets $\mathcal{Y}$ and $\mathcal{Z}$, and transition probabilities $p_{Y|X}$ and $p_{Z|Y}$ such that

$$\forall n \geqslant 1 \quad \forall (x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$$

$$p_{Y^n Z^n|X^n}(y^n, z^n|x^n) = \prod_{i=1}^{n} p_{Y|X}(y_i|x_i) p_{Z|Y}(z_i|y_i). \quad (3.2)$$

The DMC $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$ characterized by the marginal transition probabilities $p_{Y|X}$ is referred to as the *main channel*, while the DMC $\left(\mathcal{X}, p_{Z|X}, \mathcal{Z}\right)$ characterized by the marginal transition probabilities $p_{Z|X}$ is referred to as the *eavesdropper's channel*. The eavesdropper is sometimes called the *wiretapper*, and, accordingly, its channel is called the *wiretapper's channel*, but we avoid this terminology because it makes limited

sense for the wireless channels discussed in Chapter 5. Throughout this book, we always assume that the transmitter, the receiver, and the eavesdropper know the channel statistics ahead of time.

As hinted in Section 3.2, randomness in the encoding process is what enables secure communications. It is convenient to represent this randomness by the realization of a DMS $(\mathcal{R}, p_{\mathrm{R}})$, which is independent of the channel and of the messages to be transmitted. Because the source is available to Alice but not to Bob and Eve, we call it a source of *local randomness*.

Definition 3.1. *A $\left(2^{nR}, n\right)$ code $\mathcal{C}_n$ for a DWTC consists of*

- *a message set $\mathcal{M} = [\![1, 2^{nR}]\!]$;*
- *a source of local randomness at the encoder $(\mathcal{R}, p_{\mathrm{R}})$;*
- *an encoding function $f : \mathcal{M} \times \mathcal{R} \to \mathcal{X}^n$, which maps a message m and a realization of the local randomness r to a codeword x^n;*
- *a decoding function $g : \mathcal{Y}^n \to \mathcal{M} \cup \{\mathbf{?}\}$, which maps each channel observation y^n to a message $\hat{m} \in \mathcal{M}$ or an error message $\mathbf{?}$.*

Note that the DMS $(\mathcal{R}, p_{\mathrm{R}})$ is included in the definition because we implicitly assume that it can be optimized as part of the code design. The $(2^{nR}, n)$ code $\mathcal{C}_n$ is assumed known by Alice, Bob, and Eve, and this knowledge includes the statistics of the DMS $(\mathcal{R}, p_{\mathrm{R}})$; however, the realizations of the DMS used for encoding are accessible only to Alice. We also assume that the message M is uniformly distributed in $\mathcal{M}$, so that the code rate is $(1/n)\mathbb{H}(\mathsf{M}) = R + \delta(n)$. The reliability performance of $\mathcal{C}_n$ is measured in terms of its average probability of error

$$\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) \triangleq \mathbb{P}\left[\hat{\mathsf{M}} \neq \mathsf{M} \mid \mathcal{C}_n\right],$$

while its secrecy performance is measured in terms of the equivocation

$$\mathbf{E}(\mathcal{C}_n) \triangleq \mathbb{H}(\mathsf{M}|\mathsf{Z}^n\mathcal{C}_n).$$

We emphasize that equivocation is conditioned on the code $\mathcal{C}_n$ because the eavesdropper knows the code ahead of time. Equivalently, the secrecy performance of the code $\mathcal{C}_n$ can be measured in terms of the leakage

$$\mathbf{L}(\mathcal{C}_n) \triangleq \mathbb{I}(\mathsf{M}; \mathsf{Z}^n|\mathcal{C}_n),$$

which measures the information leaked to the eavesdropper instead of the uncertainty of the eavesdropper.

Remark 3.2. *In the literature, it is common to introduce the local randomness implicitly by considering a* stochastic *encoder $f : \mathcal{M} \to \mathcal{X}^n$, which maps a message $m \in \mathcal{M}$ to a codeword $x^n \in \mathcal{X}^n$ according to transition probabilities $p_{\mathsf{X}^n|\mathsf{M}}$.*

Remark 3.3. *Stochastic encoding is crucial to enable secure communications but there is no point in considering a stochastic decoder. To see this, consider a stochastic decoder that maps each channel observation $y^n \in \mathcal{Y}^n$ to a symbol $v \in \mathcal{V}$ according*

to transition probabilities $p_{\mathsf{V}|\mathsf{Y}^n}$, where $\mathcal{V}$ is an arbitrary alphabet. From the data-processing inequality, we have $\mathbb{I}(\mathsf{M};\mathsf{V}) \leqslant \mathbb{I}(\mathsf{M};\mathsf{Y}^n)$; therefore, according to the channel coding theorem, stochastic decoding can only reduce the rate of reliable transmission over the main channel, while having no effect on the eavesdropper's equivocation.

Definition 3.2. *A weak rate–equivocation pair (R, R_{e}) is achievable for the DWTC if there exists a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that*

$$\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = 0 \qquad \textit{(reliability condition)}, \tag{3.3}$$

$$\varliminf_{n\to\infty} \frac{1}{n}\mathbf{E}(\mathcal{C}_n) \geqslant R_{\mathrm{e}} \qquad \textit{(weak secrecy condition)}. \tag{3.4}$$

The weak rate–equivocation region of a DWTC is

$$\mathcal{R}^{\mathrm{DWTC}} \triangleq \mathrm{cl}(\{(R, R_{\mathrm{e}}) : (R, R_{\mathrm{e}}) \text{ is achievable}\}),$$

and the weak secrecy capacity of a DWTC is

$$C_{\mathrm{s}}^{\mathrm{DWTC}} \triangleq \sup_{R}\big\{R : (R, R) \in \mathcal{R}^{\mathrm{DWTC}}\big\}.$$

Remark 3.4. *According to our definition, if a rate–equivocation pair (R, R_{e}) is achievable, then any pair (R, R'_{e}) with $R'_{\mathrm{e}} \leqslant R_e$ is achievable as well. In particular, note that $(R, 0)$ is always achievable.*

The rate–equivocation region $\mathcal{R}^{\mathrm{DWTC}}$ encompasses rate–equivocation pairs for which R_{e} is not equal to R; it characterizes the equivocation rate that can be guaranteed for an arbitrary rate R. If a pair (R, R_{e}) with $R_{\mathrm{e}} = R$ is achievable, we say that R is a *full secrecy* rate. In this case, notice that a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ that achieves a full secrecy rate satisfies

$$\lim_{n\to\infty} \frac{1}{n}\mathbf{L}(\mathcal{C}_n) = 0.$$

Full secrecy is of practical importance because the messages transmitted are then entirely hidden from the eavesdropper. In the literature, full secrecy is sometimes called "perfect secrecy." In this book, the term "perfect secrecy" is restricted to Shannon's definition of information-theoretic security, which requires *exact* statistical independence.

The secrecy condition (3.4) is weak because it is based on the equivocation rate $(1/n)\mathbf{E}(\mathcal{C}_n)$. As discussed in Section 3.3, it would be preferable to use a stronger condition and to rely on the following definition.

Definition 3.3. *A strong rate–equivocation pair (R, R_{e}) is achievable for the DWTC if there exists a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that*

$$\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = 0 \qquad \textit{(reliability condition)}, \tag{3.5}$$

$$\varliminf_{n\to\infty} \big(\mathbf{E}(\mathcal{C}_n) - nR_{\mathrm{e}}\big) \geqslant 0 \qquad \textit{(strong secrecy condition)}. \tag{3.6}$$

The strong rate–equivocation region of a DWTC is

$$\overline{\mathcal{R}}^{\mathrm{DWTC}} \triangleq \mathrm{cl}(\{(R, R_{\mathrm{e}}) : (R, R_{\mathrm{e}}) \textit{ is achievable}\}),$$

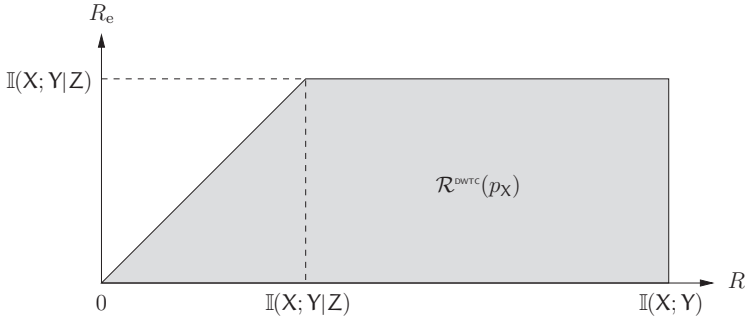

Figure 3.5 Typical shape of the rate–equivocation region $\mathcal{R}^{\text{DWTC}}(p_X)$.

and the strong secrecy capacity of a DWTC is

$$\overline{C}_s^{\text{DWTC}} \triangleq \sup_R \left\{ R : (R, R) \in \overline{\mathcal{R}}^{\text{DWTC}} \right\}.$$

Unfortunately, directly dealing with the stronger condition (3.6) is more arduous than dealing with the weak secrecy condition (3.4). Moreover, we will show in Section 4.5 that $\mathcal{R}^{\text{DWTC}} = \overline{\mathcal{R}}^{\text{DWTC}}$ and $C_s^{\text{DWTC}} = \overline{C}_s^{\text{DWTC}}$; therefore, we will content ourselves with (3.4) for now, but the reader should keep in mind that this is mainly for mathematical tractability.

It is not a priori obvious whether the reliability condition (3.3) and the secrecy condition (3.4) can be satisfied simultaneously. On the one hand, reliability calls for the introduction of redundancy to mitigate the effect of channel noise; on the other hand, creating too much redundancy is likely to jeopardize secrecy. Perhaps surprisingly, the balance between reliability and secrecy can be precisely controlled with appropriate coding schemes and the rate–equivocation region can be characterized exactly.

Theorem 3.2 (Wyner)**.** *Consider a DWTC* $\left(\mathcal{X}, p_{Z|Y}\, p_{Y|X}, \mathcal{Y}, \mathcal{Z}\right)$*. For any distribution* p_X *on* $\mathcal{X}$*, define the set* $\mathcal{R}^{\text{DWTC}}(p_X)$ *as*

$$\mathcal{R}^{\text{DWTC}}(p_X) \triangleq \left\{ (R, R_e) : \begin{array}{l} 0 \leqslant R_e \leqslant R \leqslant \mathbb{I}(X;Y) \\ 0 \leqslant R_e \leqslant \mathbb{I}(X;Y|Z) \end{array} \right\},$$

where the joint distribution of X, Y*, and* Z *factorizes as* $p_X p_{Y|X} p_{Z|Y}$*. Then, the rate–equivocation region of the DWTC is the convex region*

$$\mathcal{R}^{\text{DWTC}} = \bigcup_{p_X} \mathcal{R}^{\text{DWTC}}(p_X). \tag{3.7}$$

The typical shape of $\mathcal{R}^{\text{DWTC}}(p_X)$ is illustrated in Figure 3.5. At transmission rates below $\mathbb{I}(X;Y|Z)$, it is always possible to find codes achieving full secrecy rates. It is also possible to transmit at rates above $\mathbb{I}(X;Y|Z)$, but the equivocation rate saturates at $R_e = \mathbb{I}(X;Y|Z)$, and there is no secrecy guaranteed for the remaining fraction of the rate.

Remark 3.5. *In Wyner's original work, the equivocation rate is defined per source symbol as* $\Delta \triangleq (1/k)\mathbb{H}(M|Z^n C_n)$ *with* $k = \log\lceil 2^{nR} \rceil$*. Since* $\Delta = R_e/R$*, the rate–equivocation region* (R, Δ) *can be obtained from the rate–equivocation region* (R, R_e)*, but, in general, the region* (R, Δ) *is not convex.*

Theorem 3.2 is proved in Section 3.4.1 and Section 3.4.2. Before getting into the details of the proof, it is instructive to consider some of its implications. First, by specializing Theorem 3.2 to full secrecy rates for which $R_e = R$, we obtain the secrecy capacity of the degraded wiretap channel.

Corollary 3.1. *The secrecy capacity of a DWTC* $\left(\mathcal{X}, p_{Z|Y}\, p_{Y|X}, \mathcal{Y}, \mathcal{Z}\right)$ *is*

$$C_s^{\text{DWTC}} = \max_{p_X} \mathbb{I}(X;Y|Z) = \max_{p_X} (\mathbb{I}(X;Y) - \mathbb{I}(X;Z)). \tag{3.8}$$

If $Y = Z$, that is the eavesdropper obtains the same observation as the legitimate receiver, then $\mathbb{I}(X;Y|Z) = 0$ and thus $C_s^{\text{DWTC}} = 0$. This result is consistent with the analysis of Shannon's cipher system in Section 3.1 and the idea that information-theoretic security cannot be achieved over noiseless channels without secret keys.

Remark 3.6. *Theorem 3.2 and Corollary 3.1 also hold for a vector channel on replacing random variables by random vectors where appropriate.*

Corollary 3.1 is quite appealing because the secrecy capacity is expressed as the difference between an information rate conveyed to the legitimate receiver and an information rate leaked to the eavesdropper. To obtain an even simpler and more intuitive characterization, it is also useful to relate the secrecy capacity to the main channel capacity $C_m \triangleq \max_{p_X} \mathbb{I}(X;Y)$ and to the eavesdropper's channel capacity $C_e \triangleq \max_{p_X} \mathbb{I}(X;Z)$. For a generic DWTC $\left(\mathcal{X}, p_{Z|Y}\, p_{Y|X}, \mathcal{Y}, \mathcal{Z}\right)$, we have

$$\begin{aligned} C_s^{\text{DWTC}} &= \max_{p_X} (\mathbb{I}(X;Y) - \mathbb{I}(X;Z)) \\ &\geqslant \max_{p_X} \mathbb{I}(X;Y) - \max_{p_X} \mathbb{I}(X;Z) \\ &= C_m - C_e; \end{aligned}$$

that is, the secrecy capacity is *at least* as large as the difference between the main channel capacity and the eavesdropper's channel capacity. The inequality can be strict, as illustrated by the following example.

Example 3.4. Consider the DWTC illustrated in Figure 3.6, in which the main channel is a "Z-channel" with parameter p, while the eavesdropper's channel is a binary symmetric channel with cross-over probability p. One can check that

$$C_m = \max_{q \in [0,1]} (\mathbb{H}_b(q(1-p)) - q\mathbb{H}_b(p)),$$

$$C_e = 1 - \mathbb{H}_b(p),$$

$$C_s^{\text{DWTC}} = \max_{q \in [0,1]} (\mathbb{H}_b(q(1-p)) + (1-q)\mathbb{H}_b(p) - \mathbb{H}_b(p+q-2pq)).$$

For $p = 0.1$, we obtain numerically $C_m - C_e \approx 0.232$ bits, whereas $C_s^{\text{DWTC}} \approx 0.246$ bits.

Nevertheless, there are DWTCs for which the lower bound $C_m - C_e$ turns out to be exactly the secrecy capacity.

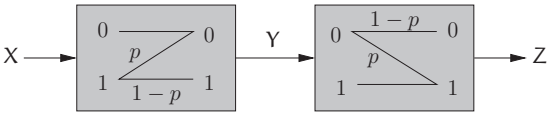

Figure 3.6 Example of a DWTC with non-symmetric channels.

Definition 3.4 (Weakly symmetric channels)**.** *A DMC* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ *is weakly symmetric if the rows of the channel transition-probability matrix are permutations of each other and the column sums* $\sum_{x\in\mathcal{X}} p_{\mathsf{Y}|\mathsf{X}}(y|x)$ *are independent of* y*.*

An important characteristic of weakly symmetric channels is captured by the following lemma.

Lemma 3.2. *The capacity-achieving input distribution of a weakly symmetric channel* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ *is the uniform distribution over* $\mathcal{X}$*.*

Proof. For an input distribution p_{X}, the mutual information $\mathbb{I}(\mathsf{X};\mathsf{Y})$ is

$$\mathbb{I}(\mathsf{X};\mathsf{Y}) = \mathbb{H}(\mathsf{Y}) - \mathbb{H}(\mathsf{Y}|\mathsf{X}) = \mathbb{H}(\mathsf{Y}) - \sum_{x\in\mathcal{X}} \mathbb{H}(\mathsf{Y}|\mathsf{X}=x)p_{\mathsf{X}}(x).$$

Notice that $\mathbb{H}(\mathsf{Y}|\mathsf{X}=x)$ is a constant, say H, that is independent of x because the rows of the channel transition-probability matrix are permutations of each other. Thus,

$$\mathbb{I}(\mathsf{X};\mathsf{Y}) = \mathbb{H}(\mathsf{Y}) - H \leqslant \log|\mathcal{Y}| - H,$$

with equality if Y is uniform. We show that choosing $p_{\mathsf{X}}(x) = 1/|\mathcal{X}|$ for all $x \in \mathcal{X}$ induces a uniform distribution for Y. In fact,

$$p_{\mathsf{Y}}(y) = \sum_{x\in\mathcal{X}} p_{\mathsf{Y}|\mathsf{X}}(y|x)p_{\mathsf{X}}(x) = \frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}} p_{\mathsf{Y}|\mathsf{X}}(y|x).$$

Since $\sum_x p_{\mathsf{Y}|\mathsf{X}}(y|x)$ is independent of y by assumption, $p_{\mathsf{Y}}(y)$ is a constant. By the law of total probability, it must hold that $p_{\mathsf{Y}}(y) = 1/|\mathcal{Y}|$ for all $y \in \mathcal{Y}$. □

Proposition 3.2 (Leung-Yan-Cheong)**.** *If the main channel and the eavesdropper's channel of a DWTC* $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{Y}}\, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z}\right)$ *are both weakly symmetric, then*

$$C_{\mathrm{s}}^{\mathrm{DWTC}} = C_{\mathrm{m}} - C_{\mathrm{e}}, \tag{3.9}$$

where C_{m} *is the capacity of the main channel and* C_{e} *is that of the eavesdropper's channel.*

The proof of Proposition 3.2 hinges on a general concavity property of the conditional mutual information $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$, which we establish in the following lemma.

Lemma 3.3. *Let* $\mathsf{X} \in \mathcal{X}$*,* $\mathsf{Y} \in \mathcal{Y}$*, and* $\mathsf{Z} \in \mathcal{Z}$ *be three random variables with joint probability distribution* p_{XYZ}*. Then,* $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$ *is a concave function of* p_{X} *for fixed* $p_{\mathsf{YZ}|\mathsf{X}}$*.*

Proof. For fixed transition probabilities $p_{\mathsf{YZ}|\mathsf{X}}$, we interpret $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$ as a function of p_{X} and we write $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}) \triangleq f(p_{\mathsf{X}})$. Let X_1, Y_1, and Z_1 be random variables such that

$$\forall (x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \quad p_{\mathsf{X}_1\mathsf{Y}_1\mathsf{Z}_1}(x,y,z) = p_{\mathsf{YZ}|\mathsf{X}}(y,z|x)p_{\mathsf{X}_1}(x).$$

Similarly, let X_2, Y_2, and Z_2 be random variables such that

$$\forall (x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \quad p_{\mathsf{X}_2\mathsf{Y}_2\mathsf{Z}_2}(x,y,z) = p_{\mathsf{YZ}|\mathsf{X}}(y,z|x)p_{\mathsf{X}_2}(x).$$

We introduce the random variable $\mathsf{Q} \in \{1,2\}$ which is independent of all others such that

$$\mathsf{Q} \triangleq \begin{cases} 1 & \text{with probability } \lambda, \\ 2 & \text{with probability } 1-\lambda, \end{cases}$$

and we define the random variables

$$\mathsf{X} \triangleq \mathsf{X}_{\mathsf{Q}}, \quad \mathsf{Y} \triangleq \mathsf{Y}_{\mathsf{Q}}, \quad \text{and} \quad \mathsf{Z} \triangleq \mathsf{Z}_{\mathsf{Q}}.$$

Note that $\mathsf{Q} \to \mathsf{X} \to \mathsf{YZ}$ forms a Markov chain and that, for all $x \in \mathcal{X}$, $p_{\mathsf{X}}(x) = \lambda p_{\mathsf{X}_1}(x) + (1-\lambda)p_{\mathsf{X}_2}(x)$. Then,

$$\begin{aligned} \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}) &= \mathbb{H}(\mathsf{Y}|\mathsf{Z}) - \mathbb{H}(\mathsf{Y}|\mathsf{XZ}) \\ &\geqslant \mathbb{H}(\mathsf{Y}|\mathsf{ZQ}) - \mathbb{H}(\mathsf{Y}|\mathsf{XZQ}), \end{aligned}$$

where the inequality follows from $\mathbb{H}(\mathsf{Y}|\mathsf{Z}) \geqslant \mathbb{H}(\mathsf{Y}|\mathsf{ZQ})$, since conditioning does not increase entropy, and $\mathbb{H}(\mathsf{Y}|\mathsf{ZX}) = \mathbb{H}(\mathsf{Y}|\mathsf{ZXQ})$, since Q is independent of Y given X. Therefore,

$$\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}) \geqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{ZQ}) = \lambda \mathbb{I}(\mathsf{X}_1;\mathsf{Y}_1|\mathsf{Z}_1) + (1-\lambda)\mathbb{I}(\mathsf{X}_2;\mathsf{Y}_2|\mathsf{Z}_2),$$

or, equivalently,

$$f\left(\lambda p_{\mathsf{X}_1} + (1-\lambda)p_{\mathsf{X}_2}\right) \geqslant \lambda f(p_{\mathsf{X}_1}) + (1-\lambda)f(p_{\mathsf{X}_2}),$$

which is the desired result. □

Note that Lemma 3.3 holds for *any* transition probabilities $p_{\mathsf{YZ}|\mathsf{X}}$, not just those corresponding to a degraded channel.

Proof of Proposition 3.2. The DMCs $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ and $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ are weakly symmetric; therefore, by Lemma 3.2, $\mathbb{I}(\mathsf{X};\mathsf{Y})$ and $\mathbb{I}(\mathsf{X};\mathsf{Z})$ are both maximized if X is uniformly distributed over $\mathcal{X}$. For a degraded channel, $\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) = \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$, which is a concave function of p_{X} by Lemma 3.3. Therefore, $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$ is also maximized if X is uniformly distributed and

$$C_{\mathrm{s}}^{\mathrm{DWTC}} = \max_{p_{\mathsf{X}}} \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}) = \max_{p_{\mathsf{X}}} \mathbb{I}(\mathsf{X};\mathsf{Y}) - \max_{p_{\mathsf{X}}} \mathbb{I}(\mathsf{X};\mathsf{Z}) = C_{\mathrm{m}} - C_{\mathrm{e}}.$$ □

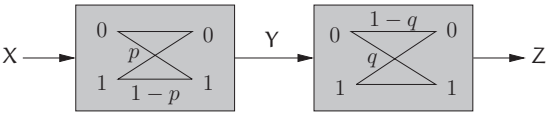

Figure 3.7 Example of a DWTC with weakly symmetric channels.

Remark 3.7. *From the proof of Proposition 3.2, we see that a sufficient condition to obtain* $C_{\mathrm{s}}^{\mathrm{DWTC}} = C_{\mathrm{m}} - C_{\mathrm{e}}$ *is that* $\mathbb{I}(\mathsf{X};\mathsf{Y})$ *and* $\mathbb{I}(\mathsf{X};\mathsf{Z})$ *are maximized for the same input distribution* p_{X}*. Nevertheless, checking that the channels* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ *and* $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ *are weakly symmetric is, in general, a much simpler task.*

Proposition 3.2 is useful because many channels of interest (binary symmetric channels) are indeed weakly symmetric and their secrecy capacity then follows easily.

Example 3.5. Consider the DWTC illustrated in Figure 3.7, which is obtained by cascading two binary symmetric channels BSC(p) and BSC(q). The main channel is symmetric by construction, and the eavesdropper's channel is a BSC($p + q - 2pq$), which is also symmetric. Therefore, by Proposition 3.2,

$$\begin{aligned} C_{\mathrm{s}}^{\mathrm{DWTC}} &= C_{\mathrm{m}} - C_{\mathrm{e}} \\ &= 1 - \mathbb{H}_{\mathrm{b}}(p) - (1 - \mathbb{H}_{\mathrm{b}}(p + q - 2pq)) \\ &= \mathbb{H}_{\mathrm{b}}(p + q - 2pq) - \mathbb{H}_{\mathrm{b}}(p). \end{aligned}$$

3.4.1 Achievability proof for the degraded wiretap channel

In this section, we prove that the rate pairs in $\mathcal{R}^{\mathrm{DWTC}}$ given by Theorem 3.2 are achievable. As is usual in information theory, we use a random-coding argument, and we show the existence of codes for the DWTC without constructing them explicitly; nevertheless, before we can start the proof, it is still necessary to identify a generic code structure that can guarantee secrecy and reliability simultaneously. In the next paragraphs, we do so by developing two desirable properties that wiretap codes should satisfy to guarantee full secrecy.

The discussion and example in Section 3.2 suggest that several codewords should represent the same message and that the choice of which codeword to transmit should be random, to "confuse" the eavesdropper. This statement can be made somewhat more precise; we can argue that, in general, a wiretap code *must* possess this property. To see this, assume that we use a wiretap code $\mathcal{C}_n$ that guarantees communication with full secrecy. It is reasonable to assume that messages are determined uniquely by codewords, that is $\mathbb{H}(\mathsf{M}|\mathsf{X}^n\mathcal{C}_n) = 0$. In addition, assume that the encoding function is a one-to-one

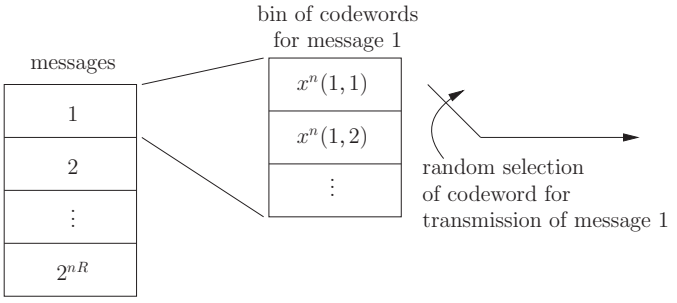

Figure 3.8 Binning structure and encoding process for a wiretap code.

mapping, that is $\mathbb{H}(X^n|M\mathcal{C}_n) = 0$. We can write the leakage $(1/n)\mathbf{L}(\mathcal{C}_n)$ as

$$\begin{aligned}\frac{1}{n}\mathbf{L}(\mathcal{C}_n) &= \frac{1}{n}\mathbb{I}(M; Z^n|\mathcal{C}_n)\\ &= \frac{1}{n}\mathbb{I}(MX^n; Z^n|\mathcal{C}_n) - \frac{1}{n}\mathbb{I}(X^n; Z^n|M\mathcal{C}_n)\\ &= \frac{1}{n}\mathbb{I}(X^n; Z^n|\mathcal{C}_n) + \frac{1}{n}\mathbb{I}(M; Z^n|X^n\mathcal{C}_n) - \frac{1}{n}\mathbb{I}(X^n; Z^n|M\mathcal{C}_n).\end{aligned}$$

Since we have assumed that $\mathbb{H}(M|X^n\mathcal{C}_n) = 0$ and $\mathbb{H}(X^n|M\mathcal{C}_n) = 0$, we have also $\mathbb{I}(M; Z^n|X^n\mathcal{C}_n) = 0$ and $\mathbb{I}(X^n; Z^n|M\mathcal{C}_n) = 0$; therefore,

$$\frac{1}{n}\mathbb{I}(X^n; Z^n|\mathcal{C}_n) = \frac{1}{n}\mathbf{L}(\mathcal{C}_n),$$

which means that the information leaked to the eavesdropper about codewords is equal to the information leaked to the eavesdropper about messages. If $\mathcal{C}_n$ allows communication in full secrecy, then for some small $\epsilon > 0$ we have $(1/n)\mathbf{L}(\mathcal{C}_n) \leqslant \epsilon$ and, consequently, $(1/n)\mathbb{I}(X^n; Z^n|\mathcal{C}_n) \leqslant \epsilon$ as well. Notice that the relation between X^n and Z^n is determined in part by the channel, over which the transmitter does not have full control. For a DMC, guaranteeing that $(1/n)\mathbb{I}(X^n; Z^n|\mathcal{C}_n) \leqslant \epsilon$ is in general[4] possible only if $(1/n)\mathbb{H}(X^n|\mathcal{C}_n) \leqslant \delta(\epsilon)$; that is, the transmission rate must be negligible. Therefore, to transmit at a non-negligible rate, we need $(1/n)\mathbb{H}(X^n|M\mathcal{C}_n)$ to be non-zero. In other words, the encoder should select a codeword at random among a set of codewords representing the same message. As illustrated in Figure 3.8, we can think of such a set as a sub-codebook or as a "bin" of codewords within the codebook; hence, we will say that a wiretap code should possess a *binning structure*.

Remark 3.8. *Despite the similarity between Figure 3.8 and Figure 2.10, the binning structure of a wiretap code is different from the superposition coding structure introduced for the broadcast channel in Section 2.3.3. A wiretap code consists of a* single *codebook partitioned into bins, whereas a superposition codebook for the broadcast channel consists of* several *codebooks that are superposed.*

[4] It is possible to prove that, with high probability, the good random codes identified by random-coding arguments leak an information rate that grows linearly with n over a DMC.

The second desirable property of wiretap codes concerns the local randomness R used in the encoding process. Since codewords are a function of messages and local randomness, note that $\mathbb{I}(X^n; Z^n|\mathcal{C}_n) = \mathbb{I}(MR; Z^n|\mathcal{C}_n)$. In addition, since the local randomness R is independent of M, we have $\mathbb{H}(R|M\mathcal{C}_n) = \mathbb{H}(R|\mathcal{C}_n)$. Therefore,

$$\begin{aligned}\frac{1}{n}\mathbf{L}(\mathcal{C}_n) &= \frac{1}{n}\mathbb{I}(M; Z^n|\mathcal{C}_n)\\ &= \frac{1}{n}\mathbb{I}(MR; Z^n|\mathcal{C}_n) - \frac{1}{n}\mathbb{I}(R; Z^n|M\mathcal{C}_n)\\ &= \frac{1}{n}\mathbb{I}(X^n; Z^n|\mathcal{C}_n) - \frac{1}{n}\mathbb{H}(R|M\mathcal{C}_n) + \frac{1}{n}\mathbb{H}(R|Z^nM\mathcal{C}_n)\\ &= \frac{1}{n}\mathbb{I}(X^n; Z^n|\mathcal{C}_n) - \frac{1}{n}\mathbb{H}(R|\mathcal{C}_n) + \frac{1}{n}\mathbb{H}(R|Z^nM\mathcal{C}_n). \qquad (3.10)\end{aligned}$$

If the code $\mathcal{C}_n$ allows communication in full secrecy, then it must be that $(1/n)\mathbf{L}(\mathcal{C}_n) \leqslant \epsilon$ for some $\epsilon > 0$. Note that (3.10) suggests that this is indeed possible, because the confusion introduced by the source of local randomness is represented by the term $(1/n)\mathbb{H}(R|\mathcal{C}_n)$, which compensates in part for the information rate leaked to the eavesdropper $(1/n)\mathbb{I}(X^n; Z^n|\mathcal{C}_n)$. However, to ensure that the confusion *cancels out* the information rate leaked, it seems desirable to design a code such that $(1/n)\mathbb{H}(R|Z^nM\mathcal{C}_n)$ is small.

Remark 3.9. *Now that generic properties of wiretap codes have been identified, it would be tempting to start a random-coding argument and to analyze both the error probability and the equivocation rate for a random-code ensemble. However, there is a subtle but critical detail to which we must pay attention. Let* C_n *be the random variable that denotes the choice of a code* $\mathcal{C}_n$ *in the code ensemble. The probability of error averaged over the ensemble is then*

$$\mathbb{P}\left[M \neq \hat{M}\right] = \mathbb{E}_{C_n}\left[\mathbb{P}\left[M \neq \hat{M}|C_n\right]\right] = \sum_{\mathcal{C}_n} p_{C_n}(\mathcal{C}_n)\mathbf{P}_e(\mathcal{C}_n).$$

In other words, the probability of error averaged over the ensemble is equal *to the average of the probability of error of individual codes. Consequently, if the probability of error averaged over the ensemble is smaller than some* $\epsilon > 0$*, there must exist at least one* specific *code* $\mathcal{C}_n$ *with* $\mathbf{P}_e(\mathcal{C}_n) \leqslant \epsilon$*. In contrast, for the equivocation of the code ensemble* $\mathbb{H}(M|Z^n)$*, we have*

$$\frac{1}{n}\mathbb{H}(M|Z^n) \geqslant \frac{1}{n}\mathbb{H}(M|Z^nC_n) = \sum_{\mathcal{C}_n} p_{C_n}(\mathcal{C}_n)\,\frac{1}{n}\mathbf{E}(\mathcal{C}_n).$$

The equivocation of the code ensemble is greater *than the average of equivocations of individual codes. Therefore, even if* $(1/n)\mathbb{H}(M|Z^n)$ *is greater than some value* R_e*, this does not ensure the existence of a specific code* $\mathcal{C}_n$ *such that* $(1/n)\mathbf{E}(\mathcal{C}_n) \geqslant R_e$*.*

Consequently, our proof must somehow analyze $\mathbb{H}(M|Z^nC_n)$ *or* $\mathbb{H}(M|Z^n\mathcal{C}_n)$ *directly. Wyner's original approach was to study the equivocation* $\mathbb{H}(M|Z^n\mathcal{C}_n)$ *of a*

well-constructed code $\mathcal{C}_n$; in this book we choose to study $\mathbb{H}(\mathsf{M}|\mathsf{Z}^n\mathsf{C}_n)$ directly with a random-coding argument.

For ease of reading, we carry out the proof in three distinct steps.

1. For a fixed distribution p_X on $\mathcal{X}$ and $R < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$, we use a random-coding argument and show the existence of a sequence of $\left(2^{nR}, n\right)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ that possess the binning structure illustrated in Figure 3.8 and are such that
$$\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = 0, \quad \lim_{n\to\infty} \frac{1}{n}\mathbb{H}(\mathsf{R}|\mathsf{Z}^n\mathsf{M}\mathcal{C}_n) = 0, \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \epsilon$$
for some arbitrary $\epsilon > 0$. This shows the existence of wiretap codes with "close" to full secrecy and
$$\mathcal{R}'(p_X) \triangleq \left\{(R, R_{\mathrm{e}}): \begin{array}{l} 0 \leqslant R \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}) \\ 0 \leqslant R_{\mathrm{e}} \leqslant R \end{array}\right\} \subseteq \mathcal{R}^{\text{DWTC}}.$$
The region $\mathcal{R}'(p_X)$ contains the full secrecy rate $R < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$, but is, in general, strictly smaller than $\mathcal{R}^{\text{DWTC}}(p_X)$ defined in Theorem 3.2.
2. We show that $\mathcal{R}^{\text{DWTC}}(p_X) \subseteq \mathcal{R}^{\text{DWTC}}$ with a minor modification of the codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ analyzed in Step 1.
3. We show that $\mathcal{R}^{\text{DWTC}}$ is convex.

Step 1. Random-coding argument

We prove the existence of a sequence of $\left(2^{nR}, n\right)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ for the DWTC with a binning structure as in Figure 3.8 such that
$$\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = 0, \quad \lim_{n\to\infty} \frac{1}{n}\mathbb{H}(\mathsf{R}|\mathsf{Z}^n\mathsf{M}\mathcal{C}_n) = 0, \tag{3.11}$$
$$\lim_{n\to\infty} \frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon). \tag{3.12}$$
The existence of these codes is established by choosing a specific source of local randomness and by combining the two constraints in (3.11) into a single reliability constraint for the *enhanced* DWTC illustrated in Figure 3.9. This channel enhances the original DWTC by

- introducing a *virtual receiver*, hereafter named Charlie, who observes the same channel output Z^n as Eve in the original DWTC, but who also has access to M through an error-free side channel;
- using a message M_{d} with uniform distribution over $[\![1, 2^{nR_{\mathrm{d}}}]\!]$ in place of the source of local randomness $(\mathcal{R}, p_R)$, and by requiring M_{d} to be reliably decoded by both Bob and Charlie.

Formally, a code for the enhanced channel is defined as follows.

Definition 3.5. *A $\left(2^{nR}, 2^{nR_{\mathrm{d}}}, n\right)$ code $\mathcal{C}_n$ for the enhanced DWTC consists of*

- *two message sets, $\mathcal{M} = [\![1, 2^{nR}]\!]$ and $\mathcal{M}_{\mathrm{d}} = [\![1, 2^{nR_{\mathrm{d}}}]\!]$;*

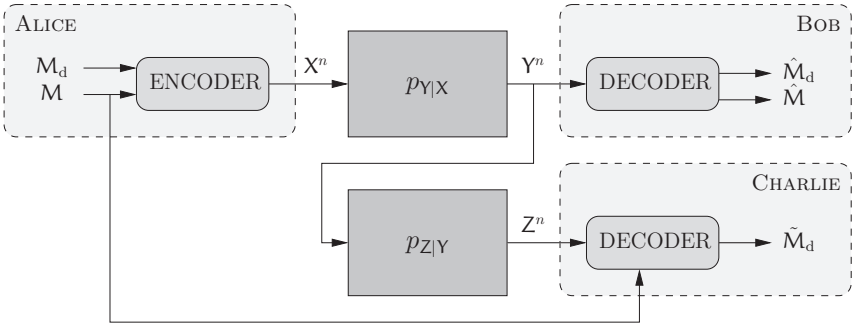

Figure 3.9 Enhanced DWTC.

- *an encoding function* $f : \mathcal{M} \times \mathcal{M}_d \to \mathcal{X}^n$*, which maps each message pair* (m, m_d) *to a codeword* x^n*;*
- *a decoding function* $g : \mathcal{Y}^n \to (\mathcal{M} \times \mathcal{M}_d) \cup \{?\}$*, which maps each channel observation* y^n *to a message pair* $(\hat{m}, \hat{m}_d) \in \mathcal{M} \times \mathcal{M}_d$ *or an error message* **?***;*
- *a decoding function* $h : \mathcal{Z}^n \times \mathcal{M} \to \mathcal{M}_d \cup \{?\}$*, which maps each channel observation* z^n *and its corresponding message* m *to a message* $\tilde{m}_d \in \mathcal{M}_d$ *or an error message* **?**.

We assume that M and M_d are uniformly distributed in their respective sets. The reliability performance of a $(2^{nR}, 2^{nR_d}, n)$ code $\mathcal{C}_n$ is measured in terms of its average probability of error

$$\mathbf{P}_e(\mathcal{C}_n) \triangleq \mathbb{P}\left[(\hat{\mathsf{M}}, \hat{\mathsf{M}}_d) \neq (\mathsf{M}, \mathsf{M}_d) \text{ or } \tilde{\mathsf{M}}_d \neq \mathsf{M}_d \middle| \mathcal{C}_n\right].$$

Because the message M_d is a *dummy message* that corresponds to a specific choice for the source of local randomness $(\mathcal{R}, p_R)$ in the original DWTC, a $\left(2^{nR}, 2^{nR_d}, n\right)$ code $\mathcal{C}_n$ for the enhanced channel is also a $\left(2^{nR}, n\right)$ code $\mathcal{C}_n$ for the original DWTC. By construction, the probability of error for the DWTC does not exceed the probability of error for the enhanced DWTC, since

$$\mathbb{P}\left[\hat{\mathsf{M}} \neq \mathsf{M} \middle| \mathcal{C}_n\right] \leqslant \mathbb{P}\left[(\hat{\mathsf{M}}, \hat{\mathsf{M}}_d) \neq (\mathsf{M}, \mathsf{M}_d) \text{ or } \tilde{\mathsf{M}}_d \neq \mathsf{M}_d \middle| \mathcal{C}_n\right] = \mathbf{P}_e(\mathcal{C}_n).$$

In addition, from Fano's inequality, we have

$$\frac{1}{n}\mathbb{H}(\mathsf{M}_d | Z^n \mathsf{M} \mathcal{C}_n) \leqslant \delta(\mathbf{P}_e(\mathcal{C}_n)).$$

Therefore, if $\lim_{n\to\infty} \mathbf{P}_e(\mathcal{C}_n) = 0$, the constraints of (3.11) are automatically satisfied with M_d in place of R.

The leakage guaranteed by the codes $\mathcal{C}_n$ is formally calculated in the next paragraphs; nevertheless, it is useful to understand intuitively why the structure of $\mathcal{C}_n$ makes this calculation possible. By using the condition $(1/n)\mathbb{H}(\mathsf{M}_d | Z^n \mathsf{M} \mathcal{C}_n) \approx 0$ in (3.10), we

write the leakage $(1/n)\mathbf{L}(\mathcal{C}_n)$ as

$$\begin{aligned}\frac{1}{n}\mathbf{L}(\mathcal{C}_n) &= \frac{1}{n}\mathbb{I}(\mathsf{M};\mathsf{Z}^n\mathcal{C}_n)\\ &= \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n\mathcal{C}_n) - \frac{1}{n}\mathbb{H}(\mathsf{M}_\mathrm{d}|\mathcal{C}_n) + \frac{1}{n}\mathbb{H}(\mathsf{M}_\mathrm{d}|\mathsf{Z}^n\mathsf{M}\mathcal{C}_n)\\ &\approx \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathcal{C}_n) - R_\mathrm{d}.\end{aligned}$$

Notice that the dummy-message rate R_d counterbalances the information rate $(1/n)\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathcal{C}_n)$ about codewords leaked to the eavesdropper. In what follows, we design $\mathcal{C}_n$ carefully so that the dummy-message rate almost *cancels out* the information leaked to the eavesdropper.

Remark 3.10. *Consider a DWTC such that for any distribution p_X on $\mathcal{X}$*

$$\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) = 0 \quad \textit{or} \quad \mathbb{I}(\mathsf{X};\mathsf{Z}) = 0.$$

If $\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) = 0$, the equivocation bound in (3.7) *reduces to $R_\mathrm{e} = 0$, which is always achieved, as has already been discussed in Remark 3.4. If $\mathbb{I}(\mathsf{X};\mathsf{Z}) = 0$, the eavesdropper's observation is independent of the channel input, which automatically ensures full secrecy $R_\mathrm{e} = R$. In both cases, the achievability proof reduces to that of the channel coding theorem for the DMC $\big(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\big)$.*

We now go back to the construction of codes for the enhanced DWTC. We begin by choosing a distribution p_X on $\mathcal{X}$ and, following the discussion in Remark 3.10, we assume without losing generality that p_X is such that

$$\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) > 0 \quad \text{and} \quad \mathbb{I}(\mathsf{X};\mathsf{Z}) > 0.$$

Let $0 < \epsilon < \mu_{\mathsf{XYZ}}$, where

$$\mu_{\mathsf{XYZ}} \triangleq \min_{(x,y,z)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}} p_{\mathsf{XYZ}}(x,y,z),$$

and let $n \in \mathbb{N}^*$. Let $R > 0$ and $R_\mathrm{d} > 0$ be rates to be specified later. We construct a $(2^{nR}, 2^{nR_\mathrm{d}}, n)$ code for the enhanced DWTC as follows.

- *Codebook construction.* Construct a codebook $\mathcal{C}_n$ with $\lceil 2^{nR}\rceil\lceil 2^{nR_\mathrm{d}}\rceil$ codewords labeled $x^n(m, m_\mathrm{d})$ with $m \in [\![1, 2^{nR}]\!]$ and $m_\mathrm{d} \in [\![1, 2^{nR_\mathrm{d}}]\!]$, by generating the symbols $x_i(m, m_\mathrm{d})$ for $i \in [\![1, n]\!]$, $m \in [\![1, 2^{nR}]\!]$, and $m_\mathrm{d} \in [\![1, 2^{nR_\mathrm{d}}]\!]$ independently according to p_X. In terms of the binning structure of Figure 3.8, m_d indexes the codewords within the bin corresponding to message m. The codebook is revealed to Alice, Bob, and Charlie.
- *Alice's encoder f.* Given (m, m_d), transmit $x^n(m, m_\mathrm{d})$.
- *Bob's decoder g.* Given y^n, output $(\hat{m}, \hat{m}_\mathrm{d})$ if it is the unique message pair such that $(x^n(\hat{m}, \hat{m}_\mathrm{d}), y^n) \in \mathcal{T}_\epsilon^n(\mathsf{XY})$; otherwise, output an error **?**.
- *Charlie's decoder h.* Given z^n and m, output $\tilde{m}_\mathrm{d}$ if it is the unique message such that $(x^n(m, \tilde{m}_\mathrm{d}), z^n) \in \mathcal{T}_\epsilon^n(\mathsf{XZ})$; otherwise, output an error **?**.

The random variable that represents the randomly generated codebook $\mathcal{C}_n$ is denoted by C_n. First, we develop an upper bound for $\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathrm{C}_n)]$ as in the proof of the channel coding theorem. Note that, from the symmetry of the random-coding construction, we have

$$\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathrm{C}_n)] = \mathbb{E}_{\mathrm{C}_n}\left[\mathbb{P}\left[\hat{\mathrm{M}} \neq \mathrm{M} | \mathrm{C}_n, \mathrm{M} = 1\right]\right].$$

Thus, without loss of generality, we can assume that $\mathrm{M} = 1$ and $\mathrm{M}_{\mathrm{d}} = 1$, and we can express $\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathrm{C}_n)]$ in terms of the events

$$\begin{aligned}\mathcal{E}_{ij} &= \left\{(\mathrm{X}^n(i,j), \mathrm{Y}^n) \in \mathcal{T}_\epsilon^n(\mathrm{XY})\right\} \text{ for } (i,j) \in [\![1, 2^R]\!] \times [\![1, 2^{nR_d}]\!], \\ \mathcal{F}_i &= \left\{(\mathrm{X}^n(1,i), \mathrm{Z}^n) \in \mathcal{T}_\epsilon^n(\mathrm{XZ})\right\} \text{ for } i \in [\![1, 2^{nR_d}]\!]\end{aligned}$$

as

$$\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathrm{C}_n)] = \mathbb{P}\left[\mathcal{E}_{11}^{\mathrm{c}} \cup \bigcup_{(i,j)\neq(1,1)} \mathcal{E}_{ij} \cup \mathcal{F}_1^{\mathrm{c}} \cup \bigcup_{i\neq 1} \mathcal{F}_i\right].$$

By the union bound,

$$\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathrm{C}_n)] \leqslant \mathbb{P}\left[\mathcal{E}_{11}^{\mathrm{c}}\right] + \sum_{(i,j)\neq(1,1)} \mathbb{P}\left[\mathcal{E}_{ij}\right] + \mathbb{P}\left[\mathcal{F}_1^{\mathrm{c}}\right] + \sum_{i\neq 1} \mathbb{P}[\mathcal{F}_i]. \tag{3.13}$$

By the AEP, we know that

$$\mathbb{P}\left[\mathcal{E}_{11}^{\mathrm{c}}\right] \leqslant \delta_\epsilon(n) \quad \text{and} \quad \mathbb{P}\left[\mathcal{F}_1^{\mathrm{c}}\right] \leqslant \delta_\epsilon(n). \tag{3.14}$$

For $(i,j) \neq (1,1)$, $\mathrm{X}^n(i,j)$ is independent of Y^n; hence, Corollary 2.2 applies and

$$\mathbb{P}\left[\mathcal{E}_{ij}\right] \leqslant 2^{-n(\mathbb{I}(\mathrm{X};\mathrm{Y})-\delta(\epsilon))} \quad \text{for } (i,j) \neq (1,1). \tag{3.15}$$

Similarly, for $i \neq 1$, $\mathrm{X}^n(1,i)$ is independent of Z^n and, by Corollary 2.2,

$$\mathbb{P}[\mathcal{F}_i] \leqslant 2^{-n(\mathbb{I}(\mathrm{X};\mathrm{Z})-\delta(\epsilon))} \quad \text{for } i \neq 1. \tag{3.16}$$

On substituting (3.14), (3.15), and (3.16) into (3.13), we obtain

$$\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathrm{C}_n)] \leqslant \delta_\epsilon(n) + \lceil 2^{nR} \rceil \lceil 2^{nR_{\mathrm{d}}} \rceil 2^{-n(\mathbb{I}(\mathrm{X};\mathrm{Y})-\delta(\epsilon))} + \lceil 2^{nR_{\mathrm{d}}} \rceil 2^{-n(\mathbb{I}(\mathrm{X};\mathrm{Z})-\delta(\epsilon))}. \tag{3.17}$$

Hence, if we choose the rates R and R_{d} to satisfy

$$R + R_{\mathrm{d}} < \mathbb{I}(\mathrm{X};\mathrm{Y}) - \delta(\epsilon) \quad \text{and} \quad R_{\mathrm{d}} < \mathbb{I}(\mathrm{X};\mathrm{Z}) - \delta(\epsilon), \tag{3.18}$$

then (3.17) implies that

$$\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathrm{C}_n)] \leqslant \delta_\epsilon(n). \tag{3.19}$$

Next, we compute an upper bound for $(1/n)\mathbb{E}[\mathbf{L}(\mathrm{C}_n)]$. Following the same steps as in (3.10), we obtain

$$\begin{aligned}\frac{1}{n}\mathbb{E}[\mathbf{L}(\mathrm{C}_n)] &= \frac{1}{n}\mathbb{I}(\mathrm{M};\mathrm{Z}^n|\mathrm{C}_n) \\ &= \frac{1}{n}\mathbb{I}(\mathrm{X}^n;\mathrm{Z}^n|\mathrm{C}_n) - \frac{1}{n}\mathbb{H}(\mathrm{M}_{\mathrm{d}}|\mathrm{C}_n) + \frac{1}{n}\mathbb{H}(\mathrm{M}_{\mathrm{d}}|\mathrm{Z}^n\mathrm{M}\mathrm{C}_n).\end{aligned} \tag{3.20}$$

We proceed to lower bound each of the three terms on the right-hand side of (3.20) separately. First, notice that, by construction, all randomly generated codes contain the same number of codewords; in addition, all these codewords are used with equal probability. Therefore,

$$\begin{aligned}\frac{1}{n}\mathbb{H}(\mathsf{M}_{\mathrm{d}}|\mathsf{C}_n) &= \sum_{\mathcal{C}_n} p_{\mathsf{C}_n}(\mathcal{C}_n)\,\frac{1}{n}\mathbb{H}(\mathsf{M}_{\mathrm{d}}|\mathcal{C}_n)\\ &= \frac{1}{n}\log(\lceil 2^{nR_{\mathrm{d}}}\rceil)\\ &\geqslant R_{\mathrm{d}}. \end{aligned} \tag{3.21}$$

Next, by Fano's inequality,

$$\begin{aligned}\frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{M}\mathsf{Z}^n\mathsf{C}_n) &= \sum_{\mathcal{C}_n} p_{\mathsf{C}_n}(\mathcal{C}_n)\frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{M}\mathsf{Z}^n\mathcal{C}_n)\\ &\leqslant \sum_{\mathcal{C}_n} p_{\mathsf{C}_n}(\mathcal{C}_n)\left(\frac{1}{n}+\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n)\frac{1}{n}\log\lceil 2^{nR_{\mathrm{d}}}\rceil\right)\\ &= \delta(n)+\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathsf{C}_n)](R_{\mathrm{d}}+\delta(n))\\ &= \delta_\epsilon(n), \end{aligned} \tag{3.22}$$

where the last inequality follows from (3.19). Finally, note that $\mathsf{C}_n \to \mathsf{X}^n \to \mathsf{Z}^n$ forms a Markov chain. Therefore,

$$\frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{C}_n) \leqslant \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n) = \mathbb{I}(\mathsf{X};\mathsf{Z}), \tag{3.23}$$

since $(\mathsf{X}^n, \mathsf{Z}^n)$ is i.i.d. according to $p_{\mathsf{X}\mathsf{Z}}$. On substituting (3.21), (3.22), and (3.23) into (3.20), we obtain

$$\mathbb{E}\left[\frac{1}{n}\mathbf{L}(\mathsf{C}_n)\right] \leqslant \mathbb{I}(\mathsf{X};\mathsf{Z}) - R_{\mathrm{d}} + \delta_\epsilon(n).$$

In particular, for the specific choice

$$R < \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) \quad \text{and} \quad R_{\mathrm{d}} = \mathbb{I}(\mathsf{X};\mathsf{Z}) - \delta(\epsilon), \tag{3.24}$$

which is compatible with the conditions in (3.18), we obtain

$$\mathbb{E}\left[\frac{1}{n}\mathbf{L}(\mathsf{C}_n)\right] \leqslant \delta(\epsilon) + \delta_\epsilon(n).$$

Finally, by applying the selection lemma to the random variable C_n and the functions $\mathbf{P}_{\mathrm{e}}$ and $\mathbf{L}$, we conclude that there exists a specific code $\mathcal{C}_n$ such that $\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) \leqslant \delta_\epsilon(n)$ and $(1/n)\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon) + \delta_\epsilon(n)$; consequently, there exists a sequence of $(2^{nR}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon),$$

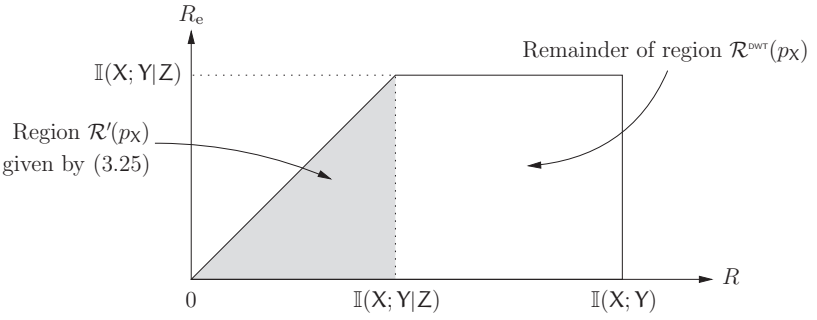

Figure 3.10 Rate–equivocation regions $\mathcal{R}'(p_X)$ given by (3.25) and $\mathcal{R}^{\text{DWTC}}(p_X)$ given in Theorem 3.2.

and the rate–equivocation pair $(R, R - \delta(\epsilon))$ is achievable. Since ϵ can be chosen arbitrarily small, since R satisfies the conditions in (3.24), and since $\mathbb{I}(X;Y) - \mathbb{I}(X;Z) = \mathbb{I}(X;Y|Z)$ for a DWTC, we conclude that

$$\mathcal{R}'(p_X) \triangleq \left\{(R, R_e): \begin{array}{l} 0 \leqslant R \leqslant \mathbb{I}(X;Y|Z) \\ 0 \leqslant R_e \leqslant R \end{array}\right\} \subseteq \mathcal{R}^{\text{DWTC}}. \tag{3.25}$$

Remark 3.11. *The fact that the channel is degraded has not been used to obtain* (3.25)*; therefore,*

$$\left\{(R, R_e): \begin{array}{l} 0 \leqslant R \leqslant \mathbb{I}(X;Y) - \mathbb{I}(X;Z) \\ 0 \leqslant R_e \leqslant R \end{array}\right\} \subseteq \mathcal{R}^{\text{DWTC}}$$

for any DMC $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$*, not just for those of the form* $\left(\mathcal{X}, p_{Z|Y}\, p_{Y|X}, \mathcal{Y}, \mathcal{Z}\right)$.

Step 2. Achieving the entire region $\mathcal{R}^{\text{DWTC}}$

As illustrated in Figure 3.10, the region $\mathcal{R}'(p_X)$ given in (3.25) is only a *subset* of the region $\mathcal{R}^{\text{DWTC}}(p_X)$. The key idea to achieve the full region is to modify the $(2^{nR}, 2^{nR_d}, n)$ codes identified in Step 1 and to exploit part of the dummy message M_d to transmit additional information. However, we have to be careful because the analysis of the probability of error and leakage for the $(2^{nR}, 2^{nR_d}, n)$ codes assumed that M and M_d were uniformly distributed; hence, we must check that our modifications do not affect the results. In the following paragraphs, we prove that this is indeed the case, but in the remainder of the book, we overlook this subtlety.

Consider a $(2^{nR}, 2^{nR_d}, n)$ code $\mathcal{C}_n$ identified in Step 1 with $R_d = \mathbb{I}(X;Z) - \delta(\epsilon)$ and $R < \mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ and such that $\mathbf{P}_e(\mathcal{C}_n) \leqslant \delta_\epsilon(n)$ and $(1/n)\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon) + \delta_\epsilon(n)$ provided that M and M_d are uniformly distributed. For $R' \leqslant R_d$, note that $\lceil 2^{nR'} \rceil$ might not divide $\lceil 2^{nR_d} \rceil$, and, by Euclidean division,

$$\lceil 2^{nR_d} \rceil = q \lceil 2^{nR'} \rceil + r,$$

for some integer $q > 0$ and some integer $0 \leqslant r < \lceil 2^{nR'} \rceil$. As illustrated in Figure 3.11, for each $m \in [\![1, 2^{nR}]\!]$, we distribute the codewords $x^n(m, m_d)$ with $m_d \in [\![1, 2^{nR_d}]\!]$, into sub-bins $\mathcal{B}_m(i)$ with $i \in [\![1, 2^{nR'}]\!]$, such that r of the sub-bins have size $q + 1$

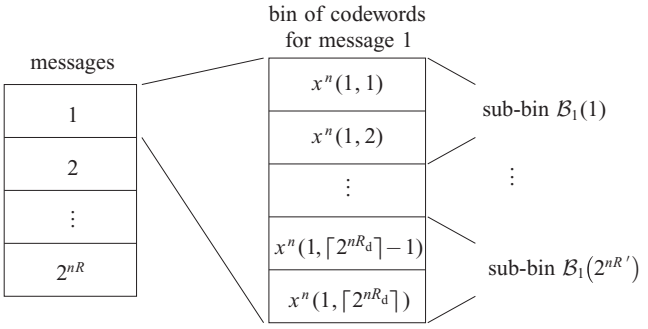

Figure 3.11 Sub-binning of codewords in $\mathcal{C}_n$.

while the remaining $\lceil 2^{nR'}\rceil - r$ have size q. We relabel the codewords $x^n(m, m', k)$ with $m \in [\![1, 2^{nR}]\!]$, $m' \in [\![1, 2^{nR'}]\!]$, and $k \in [\![1, q+1]\!]$ or $k \in [\![1, q]\!]$. The sub-binning is revealed to all parties, and we consider the following encoding/decoding procedure.

- *Encoder.* Given m and m', transmit a codeword $x^n(m, m', k) \in \mathcal{C}_n$ chosen uniformly at random in $\mathcal{B}_m(m')$.
- *Decoder.* Given y^n, use the decoding procedure of $\mathcal{C}_n$.

The sub-binning together with the encoding defines a $(2^{nR''}, n)$ code $\tilde{\mathcal{C}}_n$ for the DWTC, with $R'' = R + R' + \delta(n)$. We assume that M and M' are uniformly distributed but, because the sub-bins have different sizes, the distribution of the codewords is now slightly non-uniform. In fact, with our encoding scheme, some codewords x^n are selected with probability $p_{X^n|\mathcal{C}_n}(x^n) = 1/(\lceil 2^{nR}\rceil\lceil 2^{nR'}\rceil q)$ while others are selected with probability $p_{X^n|\mathcal{C}_n}(x^n) = 1/[\lceil 2^{nR}\rceil\lceil 2^{nR'}\rceil(q+1)]$. Nevertheless, the reader can check that

$$\sum_{x^n \in \mathcal{C}_n} \left| p_{X^n|\mathcal{C}_n}(x^n) - \frac{1}{\lceil 2^{nR}\rceil\lceil 2^{nR_d}\rceil} \right| \leqslant \delta(n);$$

that is, the variational distance between the distribution of codewords $p_{X^n|\mathcal{C}_n}$ and the uniform distribution over $\mathcal{C}_n$ vanishes for n large enough. Consequently, the probability of decoding $\mathbf{P}_e(\tilde{\mathcal{C}}_n)$ satisfies

$$\mathbf{P}_e(\tilde{\mathcal{C}}_n) \leqslant \mathbf{P}_e(\mathcal{C}_n)(1 + \delta(n)) \leqslant \delta_\epsilon(n).$$

In addition, the equivocation $(1/n)\mathbf{E}(\tilde{\mathcal{C}}_n)$ satisfies

$$\begin{aligned}\frac{1}{n}\mathbf{E}(\tilde{\mathcal{C}}_n) &= \frac{1}{n}\mathbb{H}\big(\mathsf{M}\mathsf{M}'|Z^n\tilde{\mathcal{C}}_n\big)\\ &\geqslant \frac{1}{n}\mathbb{H}\big(\mathsf{M}|Z^n\tilde{\mathcal{C}}_n\big)\\ &= \frac{1}{n}\mathbb{H}\big(\mathsf{M}|\tilde{\mathcal{C}}_n\big) - \frac{1}{n}\mathbb{I}\big(\mathsf{M}; Z^n|\tilde{\mathcal{C}}_n\big).\end{aligned}$$

Since M is chosen uniformly at random, $\mathbb{H}\big(\mathsf{M}|\tilde{\mathcal{C}}_n\big) = \mathbb{H}(\mathsf{M}|\mathcal{C}_n)$. Also, $\mathbb{I}\big(\mathsf{M}; Z^n|\tilde{\mathcal{C}}_n\big)$ is a continuous function of p_{X^n}; therefore,

$$\mathbb{I}\big(\mathsf{M}; Z^n|\tilde{\mathcal{C}}_n\big) \leqslant \mathbb{I}(\mathsf{M}; Z^n|\mathcal{C}_n) + \delta(n) = \mathbf{L}(\mathcal{C}_n) + \delta(n).$$

Hence,

$$\begin{aligned}\frac{1}{n}\mathbf{E}(\tilde{\mathcal{C}}_n) &\geqslant \frac{1}{n}\mathbb{H}(\mathsf{M}|\mathcal{C}_n) - \frac{1}{n}\mathbf{L}(\mathcal{C}_n) - \delta(n)\\ &\geqslant R - \delta(\epsilon) - \delta_\epsilon(n).\end{aligned}$$

Therefore, the rate–equivocation pair $(R + R', R - \delta(\epsilon))$ is achievable. Since R' can be chosen as large as $\mathbb{I}(\mathsf{X};\mathsf{Z}) - \delta(\epsilon)$ and since ϵ can be chosen arbitrarily small, we conclude that

$$\mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}}) = \left\{(R, R_{\text{e}}): \quad \begin{matrix} 0 \leqslant R_{\text{e}} \leqslant R \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}) \\ 0 \leqslant R_{\text{e}} \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}) \end{matrix}\right\} \subseteq \mathcal{R}^{\text{DWTC}}. \tag{3.26}$$

Step 3. Convexity of the rate–equivocation region

We show that $\mathcal{R}^{\text{DWTC}}$ is convex by proving that, for any distributions p_{X_1} and p_{X_2} on $\mathcal{X}$, the convex hull of $\mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_1}) \cup \mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_2})$ is in $\mathcal{R}^{\text{DWTC}}$.

Let $(R_1, R_{\text{e}1}) \in \mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_1})$ be a rate–equivocation pair satisfying the inequalities in (3.26) for some random variables $\mathsf{X}_1, \mathsf{Y}_1$, and Z_1 whose joint distribution is such that

$$\forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \quad p_{\mathsf{X}_1\mathsf{Y}_1\mathsf{Z}_1}(x, y, z) = p_{\mathsf{X}_1}(x)p_{\mathsf{Y}|\mathsf{X}}(y|x)p_{\mathsf{Z}|\mathsf{Y}}(z|y).$$

Similarly, let $(R_2, R_{\text{e}2}) \in \mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_2})$ be a rate–equivocation pair satisfying the inequalities in (3.26) for some random variables $\mathsf{X}_2, \mathsf{Y}_2$, and Z_2 whose joint distribution is such that

$$\forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \quad p_{\mathsf{X}_2\mathsf{Y}_2\mathsf{Z}_2}(x, y, z) = p_{\mathsf{X}_2}(x)p_{\mathsf{Y}|\mathsf{X}}(y|x)p_{\mathsf{Z}|\mathsf{Y}}(z|y).$$

Our objective is to show that, for any $\lambda \in [0, 1]$, there exists a distribution p_{X_λ} on $\mathcal{X}$ such that

$$(\lambda R_1 + (1 - \lambda)R_2, \lambda R_{\text{e}1} + (1 - \lambda)R_{\text{e}2}) \in \mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_\lambda}).$$

We define a random variable $\mathsf{Q} \in \{1, 2\}$ that is independent of all others such that

$$\mathsf{Q} \triangleq \begin{cases} 1 & \text{with probability } \lambda, \\ 2 & \text{with probability } 1 - \lambda. \end{cases}$$

By construction, $\mathsf{Q} \to \mathsf{X}_\mathsf{Q} \to \mathsf{Y}_\mathsf{Q} \to \mathsf{Z}_\mathsf{Q}$ forms a Markov chain and the joint distribution of $\mathsf{X}_\mathsf{Q}, \mathsf{Y}_\mathsf{Q}$, and Z_Q satisfies

$$\forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \quad p_{\mathsf{X}_\mathsf{Q}\mathsf{Y}_\mathsf{Q}\mathsf{Z}_\mathsf{Q}}(x, y, z) = p_{\mathsf{X}_\mathsf{Q}}(x)p_{\mathsf{Y}|\mathsf{X}}(y|x)p_{\mathsf{Z}|\mathsf{Y}}(z|y).$$

We set $\mathsf{X}_\lambda \triangleq \mathsf{X}_\mathsf{Q}$, $\mathsf{Y}_\lambda \triangleq \mathsf{Y}_\mathsf{Q}$, and $\mathsf{Z}_\lambda \triangleq \mathsf{Z}_\mathsf{Q}$. Then,

$$\begin{aligned}\mathbb{I}(\mathsf{X}_\lambda, \mathsf{Y}_\lambda) &= \mathbb{I}\big(\mathsf{X}_\mathsf{Q}; \mathsf{Y}_\mathsf{Q}\big)\\ &\geqslant \mathbb{I}\big(\mathsf{X}_\mathsf{Q}; \mathsf{Y}_\mathsf{Q}|\mathsf{Q}\big)\\ &= \lambda\,\mathbb{I}(\mathsf{X}_1; \mathsf{Y}_1) + (1 - \lambda)\mathbb{I}(\mathsf{X}_2; \mathsf{Y}_2)\\ &\geqslant \lambda R_1 + (1 - \lambda)R_2,\end{aligned}$$

and, similarly,

$$\begin{aligned}\mathbb{I}(\mathsf{X}_\lambda;\mathsf{Y}_\lambda|\mathsf{Z}_\lambda) &= \mathbb{I}\big(\mathsf{X}_\mathsf{Q};\mathsf{Y}_\mathsf{Q}|\mathsf{Z}_\mathsf{Q}\big)\\ &\geqslant \mathbb{I}\big(\mathsf{X}_\mathsf{Q};\mathsf{Y}_\mathsf{Q}|\mathsf{Z}_\mathsf{Q}\mathsf{Q}\big)\\ &= \lambda\,\mathbb{I}(\mathsf{X}_1;\mathsf{Y}_1|\mathsf{Z}_1)+(1-\lambda)\mathbb{I}(\mathsf{X}_2;\mathsf{Y}_2|\mathsf{Z}_2)\\ &\geqslant \lambda R_{\mathrm{e}1}+(1-\lambda)R_{\mathrm{e}2}.\end{aligned}$$

Hence, for any $\lambda\in[0,1]$, there exists X_λ such that

$$(\lambda R_1+(1-\lambda)R_2,\lambda R_{\mathrm{e}1}+(1-\lambda)R_{\mathrm{e}2})\in\mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_\lambda})\subseteq\mathcal{R}^{\text{DWTC}}.$$

Therefore, the convex hull of $\mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_1})\cup\mathcal{R}^{\text{DWTC}}(p_{\mathsf{X}_2})$ is included in $\mathcal{R}^{\text{DWTC}}$ and $\mathcal{R}^{\text{DWTC}}$ is convex.

3.4.2 Converse proof for the degraded wiretap channel

Let $(R,R_{\mathrm{e}})\in\mathcal{R}^{\text{DWTC}}$ be an achievable rate–equivocation pair and let $\epsilon>0$. For n sufficiently large, there exists a $(2^{nR},n)$ code $\mathcal{C}_n$ such that

$$\frac{1}{n}\mathbb{H}(\mathsf{M}|\mathcal{C}_n)\geqslant R,\quad \frac{1}{n}\mathbf{E}(\mathcal{C}_n)\geqslant R_{\mathrm{e}}-\delta(\epsilon),\quad \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n)\leqslant\delta(\epsilon).$$

In the remainder of this section, we drop the conditioning on $\mathcal{C}_n$ to simplify the notation. By Fano's inequality, we have

$$\frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{Y}^n\mathsf{Z}^n)\leqslant\frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{Y}^n)\leqslant\delta(\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n))=\delta(\epsilon).$$

Therefore,

$$\begin{aligned}R &\leqslant \frac{1}{n}\mathbb{H}(\mathsf{M})\\ &= \frac{1}{n}\mathbb{I}(\mathsf{M};\mathsf{Y}^n)+\frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{Y}^n)\\ &\overset{\text{(a)}}{\leqslant} \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Y}^n)+\delta(\epsilon)\\ &= \frac{1}{n}\mathbb{H}(\mathsf{Y}^n)-\frac{1}{n}\mathbb{H}(\mathsf{Y}^n|\mathsf{X}^n)+\delta(\epsilon)\\ &\overset{\text{(b)}}{=} \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{H}\big(\mathsf{Y}_i|\mathsf{Y}^{i-1}\big)-\frac{1}{n}\mathbb{H}\big(\mathsf{Y}_i|\mathsf{X}_i\mathsf{Y}^{i-1}\big)\right)+\delta(\epsilon)\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(\mathsf{X}_i;\mathsf{Y}_i|\mathsf{Y}^{i-1}\big)+\delta(\epsilon),\end{aligned}\tag{3.27}$$

where (a) follows from the data-processing inequality applied to the Markov chain $\mathsf{M}\to\mathsf{X}^n\to\mathsf{Y}^n$ and the bound $(1/n)\mathbb{H}(\mathsf{M}|\mathsf{Y}^n)\leqslant\delta(\epsilon)$, and (b) follows because the

channel is memoryless. Similarly, we bound the equivocation rate R_e as

$$\begin{aligned}
R_e &\leqslant \frac{1}{n}\mathbb{H}(M|Z^n) + \delta(\epsilon) \\
&= \frac{1}{n}\mathbb{I}(M;Y^n|Z^n) + \frac{1}{n}\mathbb{H}(M|Y^nZ^n) + \delta(\epsilon) \\
&\overset{(a)}{\leqslant} \frac{1}{n}\mathbb{I}(M;Y^n|Z^n) + \delta(\epsilon) \\
&\overset{(b)}{\leqslant} \frac{1}{n}\mathbb{I}(X^n;Y^n|Z^n) + \delta(\epsilon) \\
&= \frac{1}{n}\mathbb{H}(Y^n|Z^n) - \frac{1}{n}\mathbb{H}(Y^n|X^nZ^n) + \delta(\epsilon) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{H}\left(Y_i|Y^{i-1}Z^n\right) - \mathbb{H}\left(Y_i|Y^{i-1}X^nZ^n\right)\right) + \delta(\epsilon) \\
&\overset{(c)}{\leqslant} \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{H}\left(Y_i|Y^{i-1}Z_i\right) - \mathbb{H}\left(Y_i|Y^{i-1}X_iZ_i\right)\right) + \delta(\epsilon) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\left(X_i;Y_i|Z_iY^{i-1}\right) + \delta(\epsilon),
\end{aligned} \tag{3.28}$$

where (a) follows from the bound $(1/n)\mathbb{H}(M|Y^nZ^n) \leqslant \delta(\epsilon)$, (b) follows from the data-processing inequality applied to the Markov chain $M \to X^n \to Y^n \to Z^n$, and (c) follows from $\mathbb{H}\left(Y_i|Y^{i-1}X^nZ^n\right) = \mathbb{H}\left(Y_i|Y^{i-1}X_iZ_i\right)$ because the channel is memoryless and $\mathbb{H}\left(Y_i|Y^{i-1}Z^n\right) \leqslant \mathbb{H}\left(Y_i|Y^{i-1}Z_i\right)$ since conditioning does not increase entropy.

We now introduce a random variable Q, which is independent of all other random variables and uniformly distributed in $[\![1, n]\!]$, so that we can rewrite (3.27) and (3.28) as

$$\begin{aligned}
R &\leqslant \sum_{i=1}^{n}\frac{1}{n}\mathbb{I}\left(X_i;Y_i|Y^{i-1}\right) + \delta(\epsilon) = \mathbb{I}\left(X_Q;Y_Q|Y^{Q-1}Q\right) + \delta(\epsilon), \\
R_e &\leqslant \sum_{i=1}^{n}\frac{1}{n}\mathbb{I}\left(X_i;Y_i|Z_iY^{i-1}\right) + \delta(\epsilon) = \mathbb{I}\left(X_Q;Y_Q|Z_QY^{Q-1}Q\right) + \delta(\epsilon).
\end{aligned} \tag{3.29}$$

Finally, we define the random variables

$$X \triangleq X_Q, \quad Y \triangleq Y_Q, \quad Z \triangleq Z_Q, \quad \text{and} \quad U \triangleq Y^{Q-1}Q. \tag{3.30}$$

Note that $U \to X \to Y \to Z$ forms a Markov chain and that the transition probabilities $p_{Z|Y}$ and $p_{Y|X}$ are the same as those of the original DWTC. On substituting (3.30) into (3.29), we obtain the conditions

$$\begin{aligned}
0 \leqslant R_e \leqslant R \leqslant \mathbb{I}(X;Y|U) + \delta(\epsilon), \\
0 \leqslant R_e \leqslant \mathbb{I}(X;Y|ZU) + \delta(\epsilon).
\end{aligned}$$

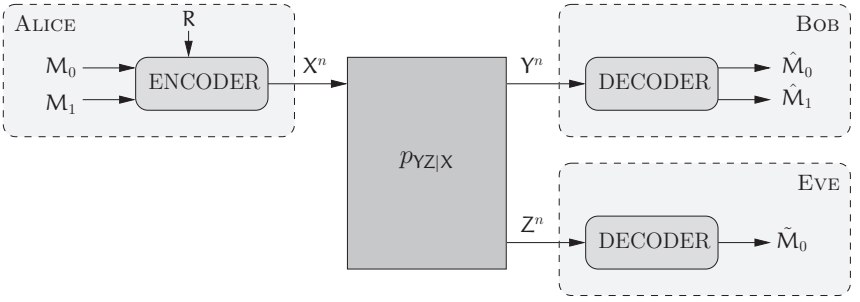

Figure 3.12 Communication over a broadcast channel with confidential messages. M_0 represents a common message for both Bob and Eve. M_1 represents an individual message for Bob, which should be kept secret from Eve. R represents local randomness used in Alice's encoder.

Since $U \to X \to Y \to Z$ forms a Markov chain, $\mathbb{I}(X;Y|U) \leqslant \mathbb{I}(X;Y)$ and $\mathbb{I}(X;Y|ZU) \leqslant \mathbb{I}(X;Y|Z)$; therefore,

$$0 \leqslant R_e \leqslant R \leqslant \mathbb{I}(X;Y) + \delta(\epsilon),$$
$$0 \leqslant R_e \leqslant \mathbb{I}(X;Y|Z) + \delta(\epsilon).$$

Since ϵ can be chosen arbitrarily small, we conclude that

$$\mathcal{R}^{\text{DWTC}} \subseteq \bigcup_{p_X} \left\{ (R, R_e): \begin{array}{l} 0 \leqslant R_e \leqslant R \leqslant \mathbb{I}(X;Y) \\ 0 \leqslant R_e \leqslant \mathbb{I}(X;Y|Z) \end{array} \right\} = \bigcup_{p_X} \mathcal{R}^{\text{DWTC}}(p_X).$$

3.5 Broadcast channel with confidential messages

The DWTC model is not entirely satisfactory because it explicitly puts the eavesdropper at a disadvantage. Although the achievability proof of Section 3.4.1 does not exploit the degraded nature of the channel, the converse proof does and it is not obvious whether the specific stochastic encoding used in Section 3.4.1 is still optimal for non-degraded channels. In addition, it is useful to characterize the trade-off between reliability and security more precisely; in particular, we would like to investigate whether one could transmit reliable messages to the eavesdropper and conceal other messages simultaneously.

These issues are resolved by analyzing a more general model than the DWTC. As illustrated in Figure 3.12, we consider a broadcast channel with two receivers for which a sender attempts to send two messages simultaneously: a *common message*, which is intended for both receivers, and an *individual secret message*, which is intended for only one receiver, treating the other receiver as an eavesdropper. This channel model was termed the *broadcast channel with confidential messages* (BCC for short) by Csiszár and Körner. In the absence of a common message, the channel is called a *wiretap channel* (WTC for short).

Formally, a *discrete memoryless* BCC $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$ consists of a finite input alphabet $\mathcal{X}$, two finite output alphabets $\mathcal{Y}$ and $\mathcal{Z}$, and transition probabilities $p_{YZ|X}$ such

that

$$\forall n \geqslant 1 \quad \forall (x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$$

$$p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(y^n, z^n|x^n) = \prod_{i=1}^{n} p_{\mathsf{YZ}|\mathsf{X}}(y_i, z_i|x_i).$$

The marginal probabilities $p_{\mathsf{Y}|\mathsf{X}}$ and $p_{\mathsf{Z}|\mathsf{X}}$ define two DMCs. By convention, the DMC $\big(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\big)$ is the main channel and the DMC $\big(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\big)$ is the eavesdropper's channel.

Definition 3.6. *A $\big(2^{nR_0}, 2^{nR_1}, n\big)$ code $\mathcal{C}_n$ for the BCC consists of*

- *a common message set $\mathcal{M}_0 = [\![1, 2^{nR_0}]\!]$ and an individual message set $\mathcal{M}_1 = [\![1, 2^{nR_1}]\!]$;*
- *a source of local randomness $(\mathcal{R}, p_{\mathsf{R}})$;*
- *an encoding function $f : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{R} \to \mathcal{X}^n$, which maps a message pair (m_0, m_1) and a realization of local randomness r to a codeword x^n;*
- *a decoding function $g : \mathcal{Y}^n \to (\mathcal{M}_0 \times \mathcal{M}_1) \cup \{\mathbf{?}\}$, which maps each channel observation y^n to a message pair $(\hat{m}_0, \hat{m}_1) \in \mathcal{M}_0 \times \mathcal{M}_1$ or an error message $\{\mathbf{?}\}$;*
- *a decoding function $h : \mathcal{Z}^n \to \mathcal{M}_0 \cup \{\mathbf{?}\}$, which maps each channel observation z^n to a message $\tilde{m}_0 \in \mathcal{M}_0$ or an error message $\{\mathbf{?}\}$.*

The $(2^{nR_0}, 2^{nR_1}, n)$ code $\mathcal{C}_n$ is known to Alice, Bob, and Eve, and we assume that messages M_0 and M_1 are chosen uniformly at random. The reliability performance of the code $\mathcal{C}_n$ is measured in terms of its average probability of error

$$\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) \triangleq \mathbb{P}\Big[(\hat{\mathsf{M}}_0, \hat{\mathsf{M}}_1) \neq (\mathsf{M}_0, \mathsf{M}_1) \text{ or } \tilde{\mathsf{M}}_0 \neq \mathsf{M}_0 \big| \mathcal{C}_n\Big],$$

while its secrecy performance is measured in terms of the equivocation

$$\mathbf{E}(\mathcal{C}_n) \triangleq \mathbb{H}(\mathsf{M}_1|\mathsf{Z}^n\mathcal{C}_n).$$

Definition 3.7. *A rate tuple $(R_0, R_1, R_{\mathrm{e}})$ is achievable for the BCC if there exists a sequence of $\big(2^{nR_0}, 2^{nR_1}, n\big)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that*

$$\lim_{n\to\infty} \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = 0 \qquad \textit{(reliability condition)}, \tag{3.31}$$

$$\varliminf_{n\to\infty} \frac{1}{n}\mathbf{E}(\mathcal{C}_n) \geqslant R_{\mathrm{e}} \qquad \textit{(weak secrecy condition)}. \tag{3.32}$$

The rate–equivocation region of the BCC is

$$\mathcal{R}^{\mathrm{BCC}} \triangleq \mathrm{cl}(\{(R_0, R_1, R_{\mathrm{e}}) : (R_0, R_1, R_{\mathrm{e}}) \text{ is achievable}\}).$$

The secrecy-capacity region is

$$\mathcal{C}^{\mathrm{BCC}} \triangleq \mathrm{cl}\big(\{(R_0, R_1) : (R_0, R_1, R_1) \in \mathcal{R}^{\mathrm{BCC}}\}\big),$$

the rate–equivocation region of the wiretap channel is

$$\mathcal{R}^{\mathrm{WT}} \triangleq \mathrm{cl}\big(\{(R_1, R_e) : (0, R_1, R_e) \in \mathcal{R}^{\mathrm{BCC}}\}\big),$$

and the secrecy capacity is

$$C_s^{\mathrm{WT}} \triangleq \sup_R \{R : (0, R, R) \in \mathcal{R}^{\mathrm{BCC}}\}.$$

The regions $\mathcal{C}^{\mathrm{BCC}}$ and $\mathcal{R}^{\mathrm{WT}}$ are specializations of the region $\mathcal{R}^{\mathrm{BCC}}$ that highlight different characteristics of a BCC. The region $\mathcal{C}^{\mathrm{BCC}}$ captures the fundamental trade-off between reliable communication with both Bob and Eve and communication in full secrecy with Bob, while the region $\mathcal{R}^{\mathrm{WT}}$ is just the generalization of the rate–equivocation region $\mathcal{R}^{\mathrm{DWTC}}$ defined in Section 3.4 for DWTCs. By replacing the weak secrecy condition (3.32) by the stronger requirement $\underline{\lim}_{n\to\infty}(\mathbf{E}(\mathcal{C}_n) - nR_e) \geqslant 0$, we obtain the strong rate–equivocation region $\overline{\mathcal{R}}^{\mathrm{BCC}}$, the strong secrecy-capacity region $\overline{\mathcal{C}}^{\mathrm{BCC}}$, and strong secrecy capacity $\overline{C}_s^{\mathrm{WT}}$.

Theorem 3.3 (Csiszár and Körner)**.** *Consider a BCC $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$. For any joint distribution $p_{\mathsf{U}\mathsf{V}\mathsf{X}}$ on $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$ that factorizes as $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}$, define the set $\mathcal{R}^{\mathrm{BCC}}(p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}})$ as*

$$\mathcal{R}^{\mathrm{BCC}}(p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}) \triangleq \left\{ (R_0, R_1, R_e) : \begin{array}{l} 0 \leqslant R_e \leqslant R_1 \\ 0 \leqslant R_e \leqslant \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) \\ 0 \leqslant R_0 \leqslant \min\left(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})\right) \\ 0 \leqslant R_1 + R_0 \leqslant \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \end{array} \right\},$$

where the joint distribution of U, V, X, Y, and Z factorizes as $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}} p_{\mathsf{YZ}|\mathsf{X}}$. Then, the rate–equivocation region of the BCC is the convex region

$$\mathcal{R}^{\mathrm{BCC}} = \bigcup_{p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}} \mathcal{R}^{\mathrm{BCC}}(p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}).$$

In addition, the cardinality of the sets $\mathcal{U}$ and $\mathcal{V}$ can be limited to

$$|\mathcal{U}| \leqslant |\mathcal{X}| + 3 \quad \textit{and} \quad |\mathcal{V}| \leqslant |\mathcal{X}|^2 + 4|\mathcal{X}| + 3.$$

The typical shape of $\mathcal{R}^{\mathrm{BCC}}(p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}})$ is illustrated in Figure 3.13. Note that the upper bound for the equivocation R_e is similar to that obtained in Theorem 3.2 for the DWTC and involves the difference between two information rates. However, the expression includes the *auxiliary random variables* U and V. Exactly how and why U and V appear in the expressions will become clear when we discuss the details of the proof in Section 3.5.2 and Section 3.5.3; at this point, suffice it to say that U, which appears as a conditioning random variable, represents the common message decodable by both the legitimate receiver and the eavesdropper while V accounts for additional randomization in the encoder.

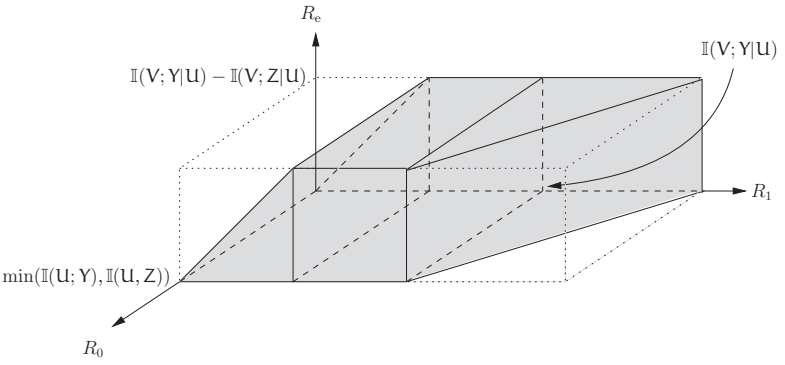

Figure 3.13 Typical shape of $\mathcal{R}^{\text{BCC}}(p_U p_{V|U} p_{X|V})$.

Theorem 3.3 leads to the following characterizations of $\mathcal{C}^{\text{BCC}}$, $\mathcal{R}^{\text{WT}}$, and $\mathcal{C}^{\text{WT}}$.

Corollary 3.2. *Consider a BCC* $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$. *For any joint distribution* p_{UVX} *on* $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$ *that factorizes as* $p_U p_{V|U} p_{X|V}$, *define the set* $\mathcal{C}^{\text{BCC}}(p_U p_{V|U} p_{X|V})$ *as*

$$\mathcal{C}^{\text{BCC}}(p_U p_{V|U} p_{X|V}) \triangleq \left\{ (R_0, R_1): \begin{array}{l} 0 \leqslant R_1 \leqslant \mathbb{I}(V;Y|U) - \mathbb{I}(V;Z|U) \\ 0 \leqslant R_0 \leqslant \min(\mathbb{I}(U;Y), \mathbb{I}(U;Z)) \end{array} \right\},$$

where the joint distribution of U, V, X, Y, *and* Z *factorizes as* $p_U p_{V|U} p_{X|V} p_{YZ|X}$. *Then, the secrecy-capacity region of the BCC is the convex set*

$$\mathcal{C}^{\text{BCC}} = \bigcup_{p_U p_{V|U} p_{X|V}} \mathcal{C}^{\text{BCC}}(p_U p_{V|U} p_{X|V}).$$

Corollary 3.3. *Consider a WTC* $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$. *For any joint distribution* p_{UVX} *on* $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$ *that factorizes as* $p_U p_{V|U} p_{X|V}$, *define the set* $\mathcal{R}^{\text{WT}}\left(p_U p_{V|U} p_{X|V}\right)$ *as*

$$\mathcal{R}^{\text{WT}}(p_U p_{V|U} p_{X|V}) \triangleq \left\{ (R, R_e): \begin{array}{l} 0 \leqslant R_e \leqslant R \leqslant \mathbb{I}(V;Y) \\ 0 \leqslant R_e \leqslant \mathbb{I}(V;Y|U) - \mathbb{I}(V;Z|U) \end{array} \right\}.$$

Then, the rate–equivocation region of the WTC is the convex set

$$\mathcal{R}^{\text{WT}} = \bigcup_{p_U p_{V|U} p_{X|V}} \mathcal{R}^{\text{WT}}\left(p_U p_{V|U} p_{X|V}\right).$$

In addition, the distributions $p_U p_{V|U} p_{X|V}$ *can be limited to those such that* $\mathbb{I}(U;Y) \leqslant \mathbb{I}(U;Z)$.

Remark 3.12. *The additional condition* $\mathbb{I}(U;Y) \leqslant \mathbb{I}(U;Z)$ *in Corollary 3.3 may seem unnecessary since* $\mathcal{R}^{\text{WT}}$ *is already completely characterized by taking the union over all distributions factorizing as* $p_U p_{V|U} p_{X|V}$; *however, if we were to evaluate* $\mathcal{R}^{\text{WT}}$ *numerically, we could speed up computations tremendously by focusing on the subset of distributions satisfying the condition* $\mathbb{I}(U;Y) \leqslant \mathbb{I}(U;Z)$.

Proof of Corollary 3.3. Substituting $R_0 = 0$ into Theorem 3.3 shows that

$$\mathcal{R}^{\mathrm{WT}} = \bigcup_{p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}} \left\{ (R, R_{\mathrm{e}}) : \begin{array}{l} 0 \leqslant R_{\mathrm{e}} \leqslant R_1 \leqslant \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \\ 0 \leqslant R_{\mathrm{e}} \leqslant \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) \end{array} \right\}.$$

We need to show that, without loss of generality, the upper bound $\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z}))$ can be replaced by $\mathbb{I}(\mathsf{V};\mathsf{Y})$.

This is always true if $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}} p_{\mathsf{YZ}|\mathsf{X}}$ is such that $\mathbb{I}(\mathsf{U};\mathsf{Y}) \leqslant \mathbb{I}(\mathsf{U};\mathsf{Z})$ because, in this case,

$$\begin{aligned} \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) &= \mathbb{I}(\mathsf{UV};\mathsf{Y}) \\ &= \mathbb{I}(\mathsf{V};\mathsf{Y}) + \mathbb{I}(\mathsf{U};\mathsf{Y}|\mathsf{V}) \\ &= \mathbb{I}(\mathsf{V};\mathsf{Y}), \end{aligned}$$

where the last equality follows from $\mathbb{I}(\mathsf{U};\mathsf{Y}|\mathsf{V}) = 0$ since $\mathsf{U} \to \mathsf{V} \to \mathsf{Y}$ forms a Markov chain. In particular, note that the distributions $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}} p_{\mathsf{YZ}|\mathsf{X}}$ for which U is a constant satisfy $\mathbb{I}(\mathsf{U};\mathsf{Y}) = \mathbb{I}(\mathsf{U};\mathsf{Z}) = 0$ and thus $\mathbb{I}(\mathsf{U};\mathsf{Y}) \leqslant \mathbb{I}(\mathsf{U};\mathsf{Z})$.

Consider now a distribution $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}} p_{\mathsf{YZ}|\mathsf{X}}$ such that $\mathbb{I}(\mathsf{U};\mathsf{Y}) > \mathbb{I}(\mathsf{U};\mathsf{Z})$. Then,

$$\begin{aligned} \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) &= \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \mathbb{I}(\mathsf{U};\mathsf{Z}) \\ &< \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \mathbb{I}(\mathsf{U};\mathsf{Y}) \\ &= \mathbb{I}(\mathsf{V};\mathsf{Y}), \end{aligned}$$

and, similarly,

$$\begin{aligned} \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) &= \sum_{u \in \mathcal{U}} p_{\mathsf{U}}(u)(\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}=u) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}=u)) \\ &\leqslant \max_{u \in \mathcal{U}} (\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}=u) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}=u)). \end{aligned}$$

Therefore, the rates (R_1, R_{e}) obtained with a distribution $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}} p_{\mathsf{YZ}|\mathsf{X}}$ such that $\mathbb{I}(\mathsf{U};\mathsf{Y}) > \mathbb{I}(\mathsf{U};\mathsf{Z})$ are upper bounded by the rates obtained with a distribution $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}} p_{\mathsf{YZ}|\mathsf{X}}$ in which U is a constant. Therefore, without loss of generality, we can obtain the entire region $\mathcal{R}^{\mathrm{WT}}$ by restricting the union to the distributions $p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}} p_{\mathsf{YZ}|\mathsf{X}}$ that satisfy $\mathbb{I}(\mathsf{U};\mathsf{Y}) \leqslant \mathbb{I}(\mathsf{U};\mathsf{Z})$ and we can replace the upper bound $\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z}))$ by $\mathbb{I}(\mathsf{V};\mathsf{Y})$. □

Corollary 3.4. *The secrecy capacity of a WTC* $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$ *is*

$$C_{\mathrm{s}}^{\mathrm{WT}} = \max_{p_{\mathsf{V}\mathsf{X}}} (\mathbb{I}(\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{V};\mathsf{Z})).$$

Proof. By Corollary 3.3,

$$C_{\mathrm{s}}^{\mathrm{WT}} = \max_{p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}} (\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U})),$$

which can be expanded as

$$\begin{aligned} C_{\mathrm{s}}^{\mathrm{WT}} &= \max_{p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}} (\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U})) \\ &= \max_{p_{\mathsf{U}} p_{\mathsf{V}|\mathsf{U}} p_{\mathsf{X}|\mathsf{V}}} \left(\sum_{u \in \mathcal{U}} p_{\mathsf{U}}(u) \left(\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}=u) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}=u)\right)\right) \\ &= \max_{p_{\mathsf{V}\mathsf{X}}} (\mathbb{I}(\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{V};\mathsf{Z})). \end{aligned}$$

□

3.5.1 Channel comparison

Although Corollary 3.4 provides an exact characterization of the secrecy capacity, the auxiliary random variable V makes the evaluation of $C_{\mathrm{s}}^{\mathrm{WT}}$ arduous and prevents us from developing much intuition about the possibility of secure communication. Nevertheless, it is possible to establish the following general result.

Lemma 3.4 (Liang *et al.*)**.** *The secrecy capacity of a WTC* $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$ *depends on the transition probabilities* $p_{\mathsf{YZ}|\mathsf{X}}$ *only through the marginal transition probabilities* $p_{\mathsf{Y}|\mathsf{X}}$ *and* $p_{\mathsf{Z}|\mathsf{X}}$.

Proof. Consider a code $\mathcal{C}_n$ designed for a WTC $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$. By definition, the average error probability $\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = \mathbb{P}\left[\mathsf{M} \neq \hat{\mathsf{M}} \mid \mathcal{C}_n\right]$ is determined by the distribution $p_{\mathsf{M}\mathsf{X}^n\mathsf{Y}^n}$ and hence depends on the transition probabilities $p_{\mathsf{Y}|\mathsf{X}}$ but not on the transition probabilities $p_{\mathsf{Z}|\mathsf{X}}$. Similarly, by definition, the equivocation $\mathbf{E}(\mathcal{C}_n) = \mathbb{H}(\mathsf{M}|\mathsf{Z}^n\mathcal{C}_n)$ is determined by the distribution $p_{\mathsf{M}\mathsf{X}^n\mathsf{Z}^n}$ and hence depends on the transition probabilities $p_{\mathsf{Z}|\mathsf{X}}$ but not on the transition probabilities $p_{\mathsf{Y}|\mathsf{X}}$. Consequently, whether a rate is achievable or not depends only on the marginal transition probabilities $p_{\mathsf{Y}|\mathsf{X}}$ and $p_{\mathsf{Z}|\mathsf{X}}$. □

Intuitively, Lemma 3.4 states that we can understand whether secure communication is possible or not if we can somehow compare the main channel $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ with the eavesdropper's channel $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$.

We have already studied a specific relation between the main channel and the eavesdropper's channel when we analyzed the DWTC in Section 3.4. In fact, the transition probabilities $p_{\mathsf{YZ}|\mathsf{X}}$ factorize as $p_{\mathsf{Z}|\mathsf{Y}} p_{\mathsf{Y}|\mathsf{X}}$ for a DWTC. We formalize this relation between the eavesdropper's channel and the main channel by introducing the notion of *physically degraded* channels.

Definition 3.8 (Physically degraded channel)**.** *We say that* $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ *is physically degraded with respect to* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ *if*

$$\forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \quad p_{\mathsf{YZ}|\mathsf{X}}(y, z|x) = p_{\mathsf{Z}|\mathsf{Y}}(z|y) p_{\mathsf{Y}|\mathsf{X}}(y|x)$$

for some transition probabilities $p_{\mathsf{Z}|\mathsf{Y}}$. *In other words,* $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ *is physically degraded with respect to* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ *if* $\mathsf{X} \to \mathsf{Y} \to \mathsf{Z}$ *forms a Markov chain.*

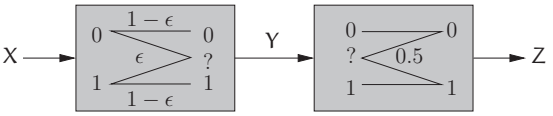

Figure 3.14 Example of a physically degraded channel.

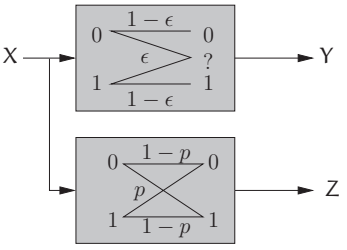

Figure 3.15 Example of a non-physically degraded channel.

Therefore, a DWTC is simply a WTC in which the eavesdropper's channel is physically degraded with respect to the main channel (see (3.2)).

Example 3.6. Consider the concatenated channels illustrated in Figure 3.14, which are such that $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ is a binary erasure channel BEC(ϵ) and $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is a binary symmetric channel BSC($\epsilon/2$). By construction, $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is physically degraded with respect to $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$.

Since physical degradedness is a stringent constraint, it is useful to consider weaker relations between the channels $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ and $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ obtained from the marginals of a broadcast channel $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$.

Definition 3.9 (Stochastically degraded channel)**.** *We say that $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is stochastically degraded with respect to $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ if there exists a channel $(\mathcal{Y}, p_{Z|Y}, \mathcal{Z})$ such that*

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z} \quad p_{Z|X}(z|x) = \sum_{y \in \mathcal{Y}} p_{Z|Y}(z|y) p_{Y|X}(y|x).$$

In other words, $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ has the same marginal as a channel that is physically degraded with respect to $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$.

Example 3.7. Consider the broadcast channel illustrated in Figure 3.15 in which $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ is a binary erasure channel BEC(ϵ) and $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is a binary symmetric channel BSC(p) with $p \in [0, \frac{1}{2}]$. If $0 \leqslant \epsilon \leqslant 2p$, then $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is stochastically degraded with respect to $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$. This fact is the consequence of a more general result that we derive in Proposition 6.4.

Note that there is no real difference between stochastically degraded channels and physically degraded channels because, by Lemma 3.4, the secrecy capacity depends only on the marginal transition probabilities of the WTC. For stochastically degraded channels, it is possible to relax the assumptions made about the eavesdropper's channel as follows.

Definition 3.10 (Class of stochastically degraded channels)**.** *A class of channels is said to be stochastically degraded with respect to a worst channel* $\left(\mathcal{X}, p_{\mathrm{Y}_0|\mathrm{X}}, \mathcal{Y}_0\right)$ *if and only if every channel* $\left(\mathcal{X}, p_{\mathrm{Y}|\mathrm{X}}, \mathcal{Y}\right)$ *in the class is stochastically degraded with respect to the worst channel.*

Proposition 3.3 (Robustness of worst-case design)**.** *Given a class of stochastically degraded eavesdropper's channels, a wiretap code ensuring equivocation* R_e *for the worst channel guarantees at least the same equivocation for any eavesdropper's channel in the class.*

Proof. The result is a direct consequence of the data-processing inequality. Let M be the message sent, let Y_0^n denote the output of the worst channel, and let Y^n denote the output of the eavesdropper's channel. Note that Y^n is statistically indistinguishable from the output of a physically degraded channel for which $\mathrm{M} \to \mathrm{Y}_0^n \to \mathrm{Y}^n$; therefore, the data-processing inequality ensures that

$$\frac{1}{n}\mathbb{H}(\mathrm{M}|\mathrm{Y}^n) \geqslant \frac{1}{n}\mathbb{H}\left(\mathrm{M}|\mathrm{Y}_0^n\right) \geqslant R_\mathrm{e}. \qquad \square$$

Despite is simplicity, Proposition 3.3 has useful applications, as illustrated by the following examples.

Example 3.8. Consider a binary erasure wiretap channel BEC(ϵ) as in Figure 3.3, for which only a lower bound ϵ^* of the eavesdropper's erasure probability is known. It is easy to verify that the set of erasure channels with erasure probability $\epsilon \geqslant \epsilon^*$ is a class of stochastically degraded channels, for which the worst channel is the one with erasure probability ϵ^*. Proposition 3.3 ensures that a wiretap code designed for the worst channel guarantees secrecy no matter what the actual erasure probability is.

Example 3.9. Another application of Proposition 3.3 is the situation in which we do not know how to design a code for a specific channel. For instance, we will see in Chapter 6 that designing wiretap codes for channels other than erasure channels is challenging. If we can show that the channel is stochastically degraded with respect to another "simpler" channel C^*, we can design a wiretap code for C^*. For instance, consider a wiretap channel with a noiseless main channel and a binary symmetric channel BSC(p) ($p < \frac{1}{2}$) as the eavesdropper's channel. The binary symmetric channel is degraded with respect to a binary erasure channel BEC($2p$), hence any wiretap code designed to operate for the latter channel also provides secrecy for the former.

Definition 3.11 (Noisier channel). $\left(\mathcal{X}, p_{Z|X}, \mathcal{Z}\right)$ *is noisier than* $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$ *if for every random variable* V *such that* $V \to X \to YZ$ *we have* $\mathbb{I}(V;Y) \geqslant \mathbb{I}(V;Z)$.

Recall from Corollary 3.4 that $C_s^{\text{WT}} = \max_{p_{VX}}(\mathbb{I}(V;Y) - \mathbb{I}(V;Z))$; therefore, $C_s = 0$ if and only if $\mathbb{I}(V;Y) \leqslant \mathbb{I}(V;Z)$ for all Markov chains $V \to X \to YZ$, which is exactly the definition of the eavesdropper's channel being noisier than the main channel. This result is summarized in the following proposition.

Proposition 3.4. *The secrecy capacity of a WTC* $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$ *is zero if and only if the main channel* $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$ *is noisier than the eavesdropper's channel* $\left(\mathcal{X}, p_{Z|X}, \mathcal{Z}\right)$.

Because the definition of "being noisier" involves an auxiliary random variable V, it may be much harder to verify that a channel is noisier than to verify that it is physically or stochastically degraded. Fortunately, the property "being noisier" admits the following characterization, which is sometimes simpler to check.

Proposition 3.5 (Van Dijk). $\left(\mathcal{X}, p_{Z|X}, \mathcal{Z}\right)$ *is noisier than* $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$ *if and only if* $\mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ *is a concave function of the input probability distribution* p_X.

Proof. Suppose $V \to X \to Y$ forms a Markov chain. Using the conditional independence of V and Y given X we have

$$\begin{aligned}\mathbb{I}(V;Y) &= \mathbb{I}(VX;Y) - \mathbb{I}(X;Y|V)\\ &= \mathbb{I}(X;Y) + \mathbb{I}(V;Y|X) - \mathbb{I}(X;Y|V)\\ &= \mathbb{I}(X;Y) - \mathbb{I}(X;Y|V).\end{aligned}$$

The same equality also holds if we replace Y by Z; therefore,

$$\begin{aligned}\mathbb{I}(V;Z) \leqslant \mathbb{I}(V;Y) &\Leftrightarrow \mathbb{I}(X;Z) - \mathbb{I}(X;Z|V) \leqslant \mathbb{I}(X;Y) - \mathbb{I}(X;Y|V)\\ &\Leftrightarrow \mathbb{I}(X;Y|V) - \mathbb{I}(X;Z|V) \leqslant \mathbb{I}(X;Y) - \mathbb{I}(X;Z). \qquad (3.33)\end{aligned}$$

For any $v \in \mathcal{V}$, we define the random variable X_v, whose distribution satisfies

$$\forall x \in \mathcal{X} \quad p_{X_v}(x) \triangleq p_{X|V}(x|v).$$

Since $V \to X \to YZ$ forms a Markov chain, note that

$$\begin{aligned}\mathbb{I}(X;Y|V) &= \sum_{v\in\mathcal{V}} p_V(v)\mathbb{I}(X;Y|V=v) = \sum_{v\in\mathcal{V}} p_V(v)\mathbb{I}(X_v;Y),\\ \mathbb{I}(X;Z|V) &= \sum_{v\in\mathcal{V}} p_V(v)\mathbb{I}(X;Z|V=v) = \sum_{v\in\mathcal{V}} p_V(v)\mathbb{I}(X_v;Z);\end{aligned}$$

therefore, we can rewrite (3.33) as

$$\mathbb{I}(V;Z) \leqslant \mathbb{I}(V;Y) \Leftrightarrow \sum_{v\in\mathcal{V}} p_V(v)\left(\mathbb{I}(X_v;Y) - \mathbb{I}(X_v;Z)\right) \leqslant \mathbb{I}(X;Y) - \mathbb{I}(X;Z).$$

Noting that

$$\forall x \in \mathcal{X} \quad p_X(x) = \sum_{v\in\mathcal{V}} p_{X|V}(x|v)p_V(v) = \sum_{v\in\mathcal{V}} p_{X_v}(x)p_V(v),$$

and treating $\mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ as a function of the input distribution p_X, the condition $\sum_{v\in\mathcal{V}} p_V(v)\,(\mathbb{I}(X_v;Y) - \mathbb{I}(X_v;Z)) \leqslant \mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ means that $\mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ is a concave function of the input distribution p_X. □

Example 3.10. Let $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ be a BEC(ϵ) and let $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ be a BSC(p) with $p \in [0, \frac{1}{2}]$ as in Figure 3.15. We show that $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is noisier than $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ if and only if

$$0 \leqslant \epsilon \leqslant 4p(1-p).$$

Proof. The result holds trivially if $\epsilon = 1$ or $p = \frac{1}{2}$. Hence, we assume $\epsilon > 0$ and $p < \frac{1}{2}$. Let $X \sim \mathcal{B}(q)$ for some $q \in [0,1]$. Then, $\mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ is a differentiable function of q given by

$$f : q \mapsto (1-\epsilon)\mathbb{H}_b(q) - \mathbb{H}_b(p + q(1-2p)) + \mathbb{H}_b(p).$$

By Proposition 3.5, it suffices to determine conditions for f to be concave. After some algebra, one obtains

$$\frac{\mathrm{d}^2 f}{\mathrm{d}q^2}(q) \leqslant 0 \Leftrightarrow -\epsilon q^2 + \epsilon q - \frac{(1-\epsilon)p(1-p)}{(1-2p)^2} \leqslant 0.$$

The quadratic polynomial in q

$$-\epsilon q^2 + \epsilon q - \frac{(1-\epsilon)(1-p)}{(1-2p)^2}$$

is negative if and only if its discriminant Δ is negative; one can check that

$$\Delta \leqslant 0 \Leftrightarrow \epsilon \leqslant 4p(1-p).$$ □

Notice that, for $2p < \epsilon \leqslant 4p(1-p)$, $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is noisier than but not stochastically degraded with respect to $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$.

Definition 3.12 (Less capable channel)**.** *$(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is less capable than $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ if for every input X we have $\mathbb{I}(X;Y) \geqslant \mathbb{I}(X;Z)$.*

Example 3.11. Let $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ be a BEC(ϵ) and let $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ be a BSC(p) with $p \in [0, \frac{1}{2}]$ as in Figure 3.15. We show that $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$ is less capable than $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ if

$$0 \leqslant \epsilon \leqslant \mathbb{H}_b(p).$$

Proof. The result holds trivially if $p = \frac{1}{2}$, hence without loss of generality we assume $p < \frac{1}{2}$. Assume $\epsilon \leqslant \mathbb{H}_b(p)$, let $X \sim \mathcal{B}(q)$ for some $q \in [0,1]$, and let f be defined as in Example 3.10. Since $f(q) = f(1-q)$, the function f is symmetric around $q = \frac{1}{2}$. In addition, because $f(0) = 0$ and $f(\frac{1}{2}) = \mathbb{H}_b(p) - \epsilon \geqslant 0$, we prove that $f(q) \geqslant 0$ for $q \in [0,1]$ by showing that $\mathrm{d}f/\mathrm{d}q(0) \geqslant 0$ and that $\mathrm{d}f/\mathrm{d}q$ changes sign at most once in

the interval $[0, \frac{1}{2}]$. After some algebra, one obtains

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}q}(q) \geqslant 0 &\Leftrightarrow (1-\epsilon)\log\left(\frac{1-q}{q}\right) - (1-2p)\log\left(\frac{1-p-q(1-2p)}{p+q(1-2p)}\right) \geqslant 0\\ &\Leftrightarrow a(p+q(1-2p))(1-q) - ((1-p)-q(1-2p))q \geqslant 0\\ &\Leftrightarrow -(a-1)(1-2p)q^2 + q(a(1-3p)-(1-p)) + ap \geqslant 0,\end{aligned}$$

with

$$a \triangleq \exp\left(\frac{1-\epsilon}{1-2p}\right) \geqslant 1.$$

The quadratic polynomial in q

$$P(q) \triangleq -(a-1)(1-2p)q^2 + q(a(1-3p)-(1-p)) + ap$$

is such that $P(0) = ap \geqslant 0$; therefore, $\mathrm{d}f/\mathrm{d}q(0) \geqslant 0$. In addition, $P(q)$ has at most one root in the interval $[0, \frac{1}{2}]$; therefore, $\mathrm{d}f/\mathrm{d}q$ changes sign at most once, which establishes the result. □

Notice that, for $4p(1-p) < \epsilon \leqslant \mathbb{H}_\mathrm{b}(p)$, $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ is less capable than but not noisier than $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$.

The following proposition shows how Definitions 3.8–3.12 relate to one another.

Proposition 3.6. *Let $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$ and $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ be two DMCs and consider the following statements:*

(1) $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ *is physically degraded w.r.t.* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$
(2) $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ *is stochastically degraded w.r.t.* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$
(3) $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ *is noisier than* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$
(4) $\left(\mathcal{X}, p_{\mathsf{Z}|\mathsf{X}}, \mathcal{Z}\right)$ *is less capable than* $\left(\mathcal{X}, p_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}\right)$.

Then,

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4).$$

Examples 3.6–3.11 show that the implications of Proposition 3.6 are strict.

Corollary 3.5. *The secrecy capacity of a WTC $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$ in which the eavesdropper's channel is less capable than the main channel is*

$$C_\mathrm{s}^{\mathsf{WT}} = \max_{p_\mathsf{X}} (\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z})).$$

Proof. The achievability of rates below $C_\mathrm{s}^{\mathsf{WT}}$ follows directly from Corollary 3.3 for general non-degraded channels on choosing $\mathsf{V} = \mathsf{X}$. To obtain the converse, note that

$$\begin{aligned}\mathbb{I}(\mathsf{V};\mathsf{Y}) &= \mathbb{I}(\mathsf{VX};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{V})\\ &= \mathbb{I}(\mathsf{X};\mathsf{Y}) + \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{X}) - \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{V})\\ &= \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{V})\end{aligned}$$

since $V \to X \to Y$ forms a Markov chain. Similarly,

$$\mathbb{I}(V;Z) = \mathbb{I}(X;Z) - \mathbb{I}(X;Z|V).$$

Therefore,

$$\mathbb{I}(V;Y) - \mathbb{I}(V;Z) = \mathbb{I}(X;Y) - \mathbb{I}(X;Z) + \mathbb{I}(X;Z|V) - \mathbb{I}(X;Y|V).$$

Now, the difference $\mathbb{I}(X;Z|V) - \mathbb{I}(X;Y|V)$ can be upper bounded as

$$\begin{aligned}
\mathbb{I}(X;Z|V) - \mathbb{I}(X;Y|V) &\leqslant \max_{p_{VX}} \left(\mathbb{I}(X;Z|V) - \mathbb{I}(X;Y|V)\right) \\
&= \max_{p_{VX}} \left(\sum_{v \in \mathcal{V}} p_V(v) \left(\mathbb{I}(X;Z|V=v) - \mathbb{I}(X;Y|V=v)\right) \right) \\
&= \max_{p_X} \left(\mathbb{I}(X;Z) - \mathbb{I}(X;Y)\right) \\
&\leqslant 0,
\end{aligned}$$

where the last inequality follows from the assumption that the eavesdropper's channel is less capable than the main channel. Consequently,

$$\mathbb{I}(V;Y) - \mathbb{I}(V;Z) \leqslant \mathbb{I}(X;Y) - \mathbb{I}(X;Z),$$

which proves that the choice $V = X$ in Corollary 3.3 is optimal. □

From a practical standpoint, the fact that the choice $V = X$ is optimum means that achieving the secrecy capacity does not require additional randomization at the encoder, which is convenient because we do not have an explicit characterization of this randomization. Because of Proposition 3.6, the expression for the secrecy capacity given in Corollary 3.5 also holds if the eavesdropper's channel is noisier than, stochastically degraded with respect to, or physically degraded with respect to the main channel. In retrospect, it might be surprising that this expression remains the same on weakening the advantage of the main channel over the eavesdropper's channel, but this suggests that additional randomization in the encoder is unnecessary for a large class of channels.

If the eavesdropper's channel is noisier than the main channel and both channels are weakly symmetric, then the secrecy capacity is the difference between the main channel capacity and the eavesdropper's channel capacity. This result generalizes Proposition 3.2, which was established for DWTCs.

Proposition 3.7 (Van Dijk)**.** *For a WTC* $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$*, if the eavesdropper's channel is noisier than the main channel and both channels are weakly symmetric, then*

$$C_s^{WT} = C_m - C_e,$$

where C_m *is the channel capacity of the main channel and* C_e *that of the eavesdropper's channel.*

Proof. From Corollary 3.5, $C_s^{WT} = \max_{p_X}(\mathbb{I}(X;Y) - \mathbb{I}(X;Z))$ and, from Proposition 3.5, $\mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ is a concave function of p_X; therefore, we can reiterate the proof

of Proposition 3.2. We note that the result also holds if the channels are not weakly symmetric, provided that $\mathbb{I}(X;Y)$ and $\mathbb{I}(X;Z)$ are maximized by the same input distribution p_X. □

As illustrated by Example 3.12 below, Proposition 3.7 does *not* hold if we replace "noisier" by "less capable."

Example 3.12. Let $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$ be a BEC(ϵ) and let $\left(\mathcal{X}, p_{Z|X}, \mathcal{Z}\right)$ be a BSC(p) as in Figure 3.15. For $\epsilon = \mathbb{H}_b(p)$, we know from Example 3.11 that $\left(\mathcal{X}, p_{Z|X}, \mathcal{Z}\right)$ is less capable than but not noisier than $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$. Notice that $C_m = 1 - \epsilon = 1 - \mathbb{H}_b(p) = C_e$; therefore, $C_m - C_e = 0$ bits. On the other hand, one can check numerically that $C_s^{WT} \approx 0.026$ bits.

To conclude this section, we illustrate the usefulness of Corollary 3.5 by computing the secrecy capacities of several wiretap channels.

Example 3.13. Consider a broadcast channel in which the main channel is a BSC(p) and the eavesdropper's channel is a BSC(r) with $r > p$. Then, the eavesdropper's channel is stochastically degraded with respect to the main channel and the secrecy capacity is $C_s^{WT} = \mathbb{H}_b(r) - \mathbb{H}_b(p)$.

Example 3.14. Consider a broadcast channel in which the main channel is a BEC(ϵ_1) and the eavesdropper's channel is a BEC(ϵ_2) with $\epsilon_2 > \epsilon_1$. Although the two channels could be *correlated* erasure channels, the secrecy capacity is $C_s^{WT} = \epsilon_2 - \epsilon_1$ because the BEC(ϵ_2) is stochastically degraded with respect to the BEC(ϵ_1).

3.5.2 Achievability proof for the broadcast channel with confidential messages

The idea of the proof is similar to Section 3.4.1. The presence of a common message and absence of physical degradedness require the introduction of two auxiliary random variables and make the proof slightly more technical, but the intuition developed earlier still applies. We carry out the proof in four steps.

1. For a fixed distribution p_{UX} on $\mathcal{U} \times \mathcal{X}$, some arbitrary $\epsilon > 0$, and rates (R_0, R_1) such that

$$R_0 < \min(\mathbb{I}(U;Y), \mathbb{I}(U;Z)), \quad \text{and} \quad R_1 < \mathbb{I}(X;Y|U) - \mathbb{I}(X;Z|U),$$

we use a random-coding argument to show the existence of a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \mathbf{P}_e(\mathcal{C}_n) = 0, \quad \lim_{n\to\infty} \mathbb{H}(W|Z^n M_1 M_0 \mathcal{C}_n) = 0, \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon).$$

The proof combines the superposition coding technique introduced in Section 2.3.3 for broadcast channels with the binning structure of wiretap codes identified in Section 3.4.1. This shows the existence of wiretap codes with rate close to full secrecy and guarantees that $\mathcal{R}_1(p_{\mathsf{UX}}) \subseteq \mathcal{R}^{\text{BCC}}$, where

$$\mathcal{R}_1(p_{\mathsf{UX}}) = \left\{ (R_0, R_1, R_e): \begin{array}{l} 0 \leqslant R_0 \leqslant \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \\ 0 \leqslant R_e \leqslant R_1 \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) \end{array} \right\}.$$

2. With a minor modification of the codes identified in Step 1, which can be thought of as an outer code construction, we show that $\mathcal{R}_2(p_{\mathsf{UX}}) \subseteq \mathcal{R}^{\text{BCC}}$, where

$$\mathcal{R}_2(p_{\mathsf{UX}}) = \left\{ (R_0, R_1, R_e): \begin{array}{l} 0 \leqslant R_e \leqslant R_1 \\ R_e \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) \\ R_1 + R_0 \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \\ 0 \leqslant R_0 \leqslant \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \end{array} \right\}.$$

3. We show that the region $\bigcup_{p_{\mathsf{UX}}} \mathcal{R}_2(p_{\mathsf{UX}})$ is convex.
4. We introduce a "prefix channel" $\left(\mathcal{V}, p_{\mathsf{X}|\mathsf{V}}, \mathcal{X}\right)$ before the BCC $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$ to create a BCC $(\mathcal{V}, p_{\mathsf{YZ}|\mathsf{V}}, \mathcal{Y}, \mathcal{Z})$. This prefix channel introduces the auxiliary random variable V and accounts for more sophisticated encoders than those used in Step 1. This shows the achievability of the entire rate–equivocation region $\mathcal{R}^{\text{BCC}}$.

Step 1. Random coding argument

We prove the existence of a sequence of $\left(2^{nR_0}, 2^{nR_1}, n\right)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty} \mathbf{P}_e(\mathcal{C}_n) = 0, \quad \lim_{n\to\infty} \mathbb{H}(\mathsf{R}|Z^n \mathsf{M}_0 \mathsf{M}_1 \mathcal{C}_n) = 0, \tag{3.34}$$

$$\lim_{n\to\infty} \frac{1}{n} \mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon). \tag{3.35}$$

The proof combines the technique of superposition coding introduced in Section 2.3.3 with the technique of wiretap coding discussed in Section 3.4.1. As in Section 3.4.1, we start by combining the two constraints in (3.34) into a single reliability constraint by

- introducing a virtual receiver, hereafter named Charlie, who observes the same channel output Z^n as Eve in the original BCC, but who also has access to M_1 and M_0 through an error-free side channel;
- using a message M_d in place of the source of local randomness $(\mathcal{R}, p_R)$ and requiring M_d to be decoded by both Bob and Charlie.

A code for this "enhanced" BCC is then defined as follows.

Definition 3.13. *A $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_d}, n\right)$ code $\mathcal{C}_n$ for the enhanced channel consists of the following:*

- *three message sets, $\mathcal{M}_0 = [\![1, 2^{nR_0}]\!]$, $\mathcal{M}_1 = [\![1, 2^{nR_1}]\!]$, and $\mathcal{M}_d = [\![1, 2^{nR_d}]\!]$;*
- *an encoding function $f : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_d \to \mathcal{X}^n$, which maps each message triple (m_0, m_1, m_d) to a codeword x^n;*

- *a decoding function* $g : \mathcal{Y}^n \to (\mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_\mathrm{d}) \cup \{\mathbf{?}\}$*, which maps each channel observation* y^n *to a message triple* $(\hat{m}_0, \hat{m}_1, \hat{m}_\mathrm{d}) \in \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_\mathrm{d}$ *or an error message* **?***;*
- *a decoding function* $h : \mathcal{Z}^n \to \mathcal{M}_0 \cup \{\mathbf{?}\}$*, which maps each channel observation* z^n *to a message* $\tilde{m}_0 \in \mathcal{M}_0$ *or an error message* **?***;*
- *a decoding function* $k : \mathcal{Z}^n \times \mathcal{M}_0 \times \mathcal{M}_1 \to \mathcal{M}_\mathrm{d} \cup \{\mathbf{?}\}$*, which maps each message pair* (m_0, m_1) *and the corresponding channel observation* z^n *to a message* $\tilde{m}_\mathrm{d} \in \mathcal{M}_\mathrm{d}$ *or an error message* **?**.

We assume that messages M_0, M_1, and M_d are uniformly distributed. The reliability performance of a $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_\mathrm{d}}, n\right)$ code $\mathcal{C}_n$ is measured in terms of its average probability of error

$$\mathbf{P}_\mathrm{e}(\mathcal{C}_n) \triangleq \mathbb{P}\left[(\hat{\mathsf{M}}_0, \hat{\mathsf{M}}_1, \hat{\mathsf{M}}_\mathrm{d}) \neq (\mathsf{M}_0, \mathsf{M}_1, \mathsf{M}_\mathrm{d}) \text{ or } \tilde{\mathsf{M}}_0 \neq \mathsf{M}_0 \text{ or } \tilde{\mathsf{M}}_\mathrm{d} \neq \mathsf{M}_\mathrm{d} \middle| \mathcal{C}_n\right],$$

while its secrecy performance is still measured in terms of the leakage

$$\mathbf{L}(\mathcal{C}_n) \triangleq \mathbb{I}(\mathsf{M}_1; \mathsf{Z}^n | \mathcal{C}_n).$$

Because M_d is a dummy message that corresponds to a specific choice for the source of local randomness $(\mathcal{R}, p_\mathsf{R})$, a $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_\mathrm{d}}, n\right)$ code $\mathcal{C}_n$ for the enhanced BCC is also a $\left(2^{nR_0}, 2^{nR_1}, n\right)$ code for the original BCC. By construction, the probability of error over the original BCC is at most $\mathbf{P}_\mathrm{e}(\mathcal{C}_n)$, since

$$\mathbb{P}\left[(\hat{\mathsf{M}}_0, \hat{\mathsf{M}}_1) \neq (\mathsf{M}_0, \mathsf{M}_1) \text{ or } \tilde{\mathsf{M}}_0 \neq \mathsf{M}_0 \middle| \mathcal{C}_n\right] \leqslant \mathbf{P}_\mathrm{e}(\mathcal{C}_n).$$

In addition, using Fano's inequality, we have

$$\frac{1}{n}\mathbb{H}(\mathsf{M}_\mathrm{d} | \mathsf{Z}^n \mathsf{M}_0 \mathsf{M}_1 \mathcal{C}_n) \leqslant \delta(\mathbf{P}_\mathrm{e}(\mathcal{C}_n)).$$

Therefore, if $\lim_{n\to\infty} \mathbf{P}_\mathrm{e}(\mathcal{C}_n) = 0$, the constraints (3.34) are automatically satisfied.

We begin by choosing a joint distribution p_{UX} on $\mathcal{U} \times \mathcal{X}$ and we assume, without loss of generality, that

$$\mathbb{I}(\mathsf{X}; \mathsf{Y} | \mathsf{U}) - \mathbb{I}(\mathsf{X}; \mathsf{Z} | \mathsf{U}) > 0 \qquad \text{and} \qquad \mathbb{I}(\mathsf{X}; \mathsf{Z} | \mathsf{U}) > 0,$$

otherwise the result follows from the channel coding theorem as discussed in Remark 3.10. Let $0 < \epsilon < \mu_{\mathsf{UXYZ}}$, where

$$\mu_{\mathsf{UXYZ}} \triangleq \min_{(u,x,y,z) \in \mathcal{U}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}} p_{\mathsf{UXYZ}}(u, x, y, z),$$

and let $n \in \mathbb{N}^*$. Let $R_0 > 0$, $R_1 > 0$, and $R_\mathrm{d} > 0$ be rates to be specified later. We construct a $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_\mathrm{d}}, n\right)$ code for the enhanced BCC by combining superposition coding and binning as follows.

- *Codebook construction.* Construct codewords $u^n(m_0)$ for $m_0 \in [\![1, 2^{nR_0}]\!]$, by generating symbols $u_i(m_0)$ with $i \in [\![1, n]\!]$ and $m_0 \in [\![1, 2^{nR_0}]\!]$ independently according to p_U. Then, for every $u^n(m_0)$, generate codewords $x^n(m_0, m_1, m_\mathrm{d})$ for $m_1 \in [\![1, 2^{nR_1}]\!]$ *and* $m_\mathrm{d} \in [\![1, 2^{nR_d}]\!]$ by generating symbols $x_i(m_0, m_1, m_\mathrm{d})$ with $i \in [\![1, n]\!]$, $m_1 \in [\![1, 2^{nR_1}]\!]$, and $m_\mathrm{d} \in [\![1, 2^{nR_d}]\!]$ independently at random according to $p_{\mathsf{X}|\mathsf{U}=u_i(m_0)}$.

- *Alice's encoder f.* Given (m_0, m_1, m_d), transmit $x^n(m_0, m_1, m_d)$.
- *Bob's decoder g.* Given y^n, output $(\hat{m}_0, \hat{m}_1, \hat{m}_d)$ if it is the unique triple such that $(u^n(\hat{m}_0), x^n(\hat{m}_0, \hat{m}_1, \hat{m}_d), y^n) \in \mathcal{T}_\epsilon^n(\mathsf{UXY})$. Otherwise, output an error **?**.
- *Eve's decoder h.* Given z^n, output $\tilde{m}_0$ if it is the unique message such that $(u^n(\tilde{m}_0), z^n) \in \mathcal{T}_\epsilon^n(\mathsf{UZ})$. Otherwise, output an error **?**.
- *Charlie's decoder k.* Given z^n, m_0, and m_1, output $\tilde{m}_d$ if it is the unique message such that $(u^n(m_0), x^n(m_0, m_1, \tilde{m}_d), z^n) \in \mathcal{T}_\epsilon^n(\mathsf{UXZ})$. Otherwise, output an error **?**.

The random variable that represents the randomly generated codebook $\mathcal{C}_n$ is denoted by C_n. By combining the analysis of superposition coding in Section 2.3.3 with the analysis of wiretap coding in Section 3.4.1, we can prove that, if

$$\begin{aligned} R_0 &< \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) - \delta(\epsilon), \\ R_1 + R_d &< \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \delta(\epsilon), \\ R_d &< \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) - \delta(\epsilon), \end{aligned} \tag{3.36}$$

then

$$\mathbb{E}[\mathbf{P}_e(\mathsf{C}_n)] \leqslant \delta_\epsilon(n). \tag{3.37}$$

Next, we compute an upper bound for $\mathbb{E}[(1/n)\mathbf{L}(\mathsf{C}_n)]$. Note that

$$\begin{aligned} \mathbb{E}\left[\frac{1}{n}\mathbf{L}(\mathsf{C}_n)\right] &= \frac{1}{n}\mathbb{I}(\mathsf{M}_1;\mathsf{Z}^n|\mathsf{C}_n) \\ &\leqslant \frac{1}{n}\mathbb{I}(\mathsf{M}_1;\mathsf{Z}^n\mathsf{M}_0|\mathsf{C}_n) \\ &= \frac{1}{n}\mathbb{I}(\mathsf{M}_1\mathsf{X}^n;\mathsf{Z}^n\mathsf{M}_0|\mathsf{C}_n) - \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n\mathsf{M}_0|\mathsf{M}_1\mathsf{C}_n) \\ &= \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n\mathsf{M}_0|\mathsf{C}_n) + \frac{1}{n}\mathbb{I}(\mathsf{M}_1;\mathsf{Z}^n\mathsf{M}_0|\mathsf{X}^n\mathsf{C}_n) - \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n\mathsf{M}_0|\mathsf{M}_1\mathsf{C}_n) \\ &\overset{(a)}{=} \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{M}_0|\mathsf{C}_n) + \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{M}_0\mathsf{C}_n) - \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n\mathsf{M}_0|\mathsf{M}_1\mathsf{C}_n) \\ &\overset{(b)}{=} \frac{1}{n}\mathbb{H}(\mathsf{M}_0|\mathsf{C}_n) + \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{M}_0\mathsf{C}_n) - \frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{M}_1\mathsf{C}_n) \\ &\quad + \frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{Z}^n\mathsf{M}_0\mathsf{M}_1\mathsf{C}_n) \\ &\overset{(c)}{=} \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{M}_0\mathsf{C}_n) - \frac{1}{n}\mathbb{H}(\mathsf{M}_d|\mathsf{C}_n) + \frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{Z}^n\mathsf{M}_0\mathsf{M}_1\mathsf{C}_n), \end{aligned} \tag{3.38}$$

where (a) follows from $\mathbb{I}(\mathsf{M}_1;\mathsf{Z}^n\mathsf{M}_0|\mathsf{X}^n\mathsf{C}_n) = 0$, (b) follows from $\mathbb{I}(\mathsf{X}^n;\mathsf{M}_0|\mathsf{C}_n) = \mathbb{H}(\mathsf{M}_0|\mathsf{C}_n)$, and (c) follows from $\mathbb{H}(\mathsf{X}^n|\mathsf{M}_1\mathsf{C}_n) = \mathbb{H}(\mathsf{M}_0|\mathsf{C}_n) + \mathbb{H}(\mathsf{M}_d|\mathsf{C}_n)$. We now bound each of the terms on the right-hand side of (3.38) individually. First, notice that the code construction ensures that

$$\frac{1}{n}\mathbb{H}(\mathsf{M}_d|\mathsf{C}_n) = \sum_{\mathcal{C}_n} p_{\mathsf{C}_n}(\mathcal{C}_n)\frac{1}{n}\mathbb{H}(\mathsf{M}_d|\mathcal{C}_n) \geqslant R_d. \tag{3.39}$$

Next, using Fano's inequality,

$$\begin{aligned}\frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{Z}^n\mathsf{M}_0\mathsf{M}_1\mathsf{C}_n) &= \sum_{\mathcal{C}_n} p_{\mathsf{C}_n}(\mathcal{C}_n)\,\frac{1}{n}\mathbb{H}(\mathsf{X}^n|\mathsf{Z}^n\mathsf{M}_0\mathsf{M}_1\mathcal{C}_n)\\ &\leqslant \sum_{\mathcal{C}_n} p_{\mathsf{C}_n}(\mathcal{C}_n)\left(\frac{1}{n}+\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n)\frac{1}{n}\log\lceil 2^{nR_d}\rceil\right)\\ &= \delta(n)+\mathbb{E}[\mathbf{P}_{\mathrm{e}}(\mathsf{C}_n)](R_d+\delta(n))\\ &\leqslant \delta_\epsilon(n),\end{aligned} \tag{3.40}$$

where the last inequality follows from (3.37). Finally, note that, given a code C_n, there is a one-to-one mapping between the message M_0 and the codeword U^n. In addition, $\mathsf{C}_n\mathsf{U}^n \to \mathsf{X}^n \to \mathsf{Z}^n$ forms a Markov chain; therefore,

$$\begin{aligned}\frac{1}{n}\mathbb{I}(\mathsf{Z}^n;\mathsf{X}^n|\mathsf{M}_0\mathsf{C}_n) &= \frac{1}{n}\mathbb{I}(\mathsf{Z}^n;\mathsf{X}^n|\mathsf{U}^n\mathsf{C}_n)\\ &= \frac{1}{n}\mathbb{H}(\mathsf{Z}^n|\mathsf{U}^n\mathsf{C}_n)-\frac{1}{n}\mathbb{H}(\mathsf{Z}^n|\mathsf{X}^n\mathsf{U}^n)\\ &\leqslant \frac{1}{n}\mathbb{H}(\mathsf{Z}^n|\mathsf{U}^n)-\frac{1}{n}\mathbb{H}(\mathsf{Z}^n|\mathsf{X}^n\mathsf{U}^n)\\ &= \frac{1}{n}\mathbb{I}(\mathsf{X}^n;\mathsf{Z}^n|\mathsf{U}^n)\\ &= \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}),\end{aligned} \tag{3.41}$$

where we have used the fact that $(\mathsf{U}^n, \mathsf{X}^n, \mathsf{Z}^n)$ is i.i.d. according to p_{UXZ}. On substituting (3.39), (3.40), and (3.41) into (3.38), we obtain

$$\mathbb{E}\left[\frac{1}{n}\mathbf{L}(\mathsf{C}_n)\right] \leqslant \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) - R_{\mathrm{d}} + \delta_\epsilon(n).$$

In particular, if we choose R_1 and R_{d} such that

$$R_1 < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) \quad \text{and} \quad R_{\mathrm{d}} = \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) - \delta(\epsilon), \tag{3.42}$$

which is compatible with the constraints in (3.36), then R_{d} almost cancels out the information rate leaked to the eavesdropper and

$$\mathbb{E}\left[\frac{1}{n}\mathbf{L}(\mathsf{C}_n)\right] \leqslant \delta(\epsilon)+\delta_\epsilon(n). \tag{3.43}$$

From (3.37) and (3.43) and by applying the selection lemma to the random variable C_n and the functions $\mathbf{P}_{\mathrm{e}}$ and $\mathbf{L}$, we conclude that there exists a specific code $\mathcal{C}_n$, such that

$$\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) \leqslant \delta_\epsilon(n) \quad \text{and} \quad \frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon)+\delta_\epsilon(n).$$

Consequently, there exists a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ codes $\{\mathcal{C}_n\}_{n\geqslant 1}$ such that

$$\lim_{n\to\infty}\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) = 0 \quad \text{and} \quad \lim_{n\to\infty}\frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon),$$

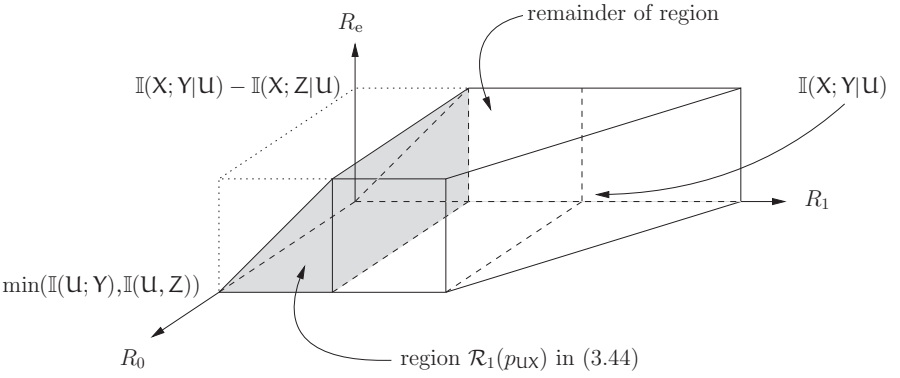

Figure 3.16 Region $\mathcal{R}_1(p_{UX})$.

which proves that $(R_0, R_1, R_1 - \delta(\epsilon))$ is achievable. Since R_0 and R_1 must satisfy the constraints imposed by (3.36) and (3.42), and since ϵ can be chosen arbitrarily small, we conclude that $\mathcal{R}_1(p_{UX}) \subseteq \mathcal{R}^{\text{BCC}}$, where

$$\mathcal{R}_1(p_{UX}) \triangleq \left\{(R_0, R_1, R_e): \begin{array}{l} 0 \leqslant R_0 \leqslant \min(\mathbb{I}(U;Y), \mathbb{I}(U;Z)) \\ 0 \leqslant R_e \leqslant R_1 \leqslant \mathbb{I}(X;Y|U) - \mathbb{I}(X;Z|U) \end{array}\right\}. \tag{3.44}$$

By construction, the joint distribution of the random variables U, X, Y, and Z in $\mathcal{R}_1(p_{UX})$ factorizes as $p_{UX}p_{YZ|X}$.

Step 2. Outer code construction

As illustrated in Figure 3.16, the rate region $\mathcal{R}_1(p_{UX})$ in (3.44) represents only a subset of $\mathcal{R}^{\text{BCC}}(p_{UVX})$; nevertheless, the entire region can be obtained with minor modifications of the $(2^{nR_0}, 2^{nR_1}, 2^{nR_d})$ codes $\mathcal{C}_n$ identified in Step 1. The key idea is to exploit part of the dummy message M_d and part of the common message M_0 as individual messages. This can be performed by introducing sub-bins for M_0 and M_d as done in Section 3.4.1, and we provide a sketch of the proof only.

- By using a fraction of the dummy message rate $R_d = \mathbb{I}(X;Z|U) - \delta(\epsilon)$, we can increase the individual-message rate without changing the equivocation and without changing the common-message rate. Hence, the region $\mathcal{R}'_1(p_{UX})$ defined as

$$\mathcal{R}'_1(p_{UX}) \triangleq \left\{(R_0, R_1, R_e): \begin{array}{l} 0 \leqslant R_e \leqslant R_1 \\ 0 \leqslant R_e \leqslant \mathbb{I}(X;Y|U) - \mathbb{I}(X;Z|U) \\ 0 \leqslant R_0 \leqslant \min(\mathbb{I}(U;Y), \mathbb{I}(U;Z)) \\ 0 \leqslant R_1 \leqslant \mathbb{I}(X;Y|U) \end{array}\right\}$$

  satisfies $\mathcal{R}'_1(p_{UX}) \subseteq \mathcal{R}^{\text{BCC}}$.
- By sacrificing a fraction of the common-message rate R_0, we can further increase the individual-message rate without changing the equivocation; however, because of the constraints (3.36), the trade-off between the individual-message rate and the common-message rate is limited by

$$R_0 + R_1 \leqslant \min(\mathbb{I}(U;Y); \mathbb{I}(U;Y)) + \mathbb{I}(X;Y|U).$$

Hence, the region $\mathcal{R}_2(p_{\mathsf{U}\mathsf{X}})$ defined as

$$\mathcal{R}_2(p_{\mathsf{U}\mathsf{X}}) \triangleq \left\{ (R_0, R_1, R_e): \begin{array}{l} 0 \leqslant R_e \leqslant R_1 \\ 0 \leqslant R_e \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{X};\mathsf{Z}|\mathsf{U}) \\ 0 \leqslant R_0 + R_1 \leqslant \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) + \min\left(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})\right) \end{array} \right\} \tag{3.45}$$

satisfies $\mathcal{R}_2(p_{\mathsf{U}\mathsf{X}}) \subseteq \mathcal{R}^{\text{BCC}}$.

Finally, since the distribution $p_{\mathsf{U}\mathsf{X}}$ is arbitrary, we obtain

$$\bigcup_{p_{\mathsf{U}\mathsf{X}}} \mathcal{R}_2(p_{\mathsf{U}\mathsf{X}}) \subseteq \mathcal{R}^{\text{BCC}}.$$

Step 3. Convexity of $\bigcup_{p_{\mathsf{U}\mathsf{X}}} \mathcal{R}_2(p_{\mathsf{U}\mathsf{X}})$

We show that $\bigcup_{p_{\mathsf{U}\mathsf{X}}} \mathcal{R}_2(p_{\mathsf{U}\mathsf{X}})$ is convex by proving that, for any distributions $p_{\mathsf{U}_1\mathsf{X}_1}$ and $p_{\mathsf{U}_2\mathsf{X}_2}$ on $\mathcal{U} \times \mathcal{X}$, the convex hull of $\mathcal{R}_2(p_{\mathsf{U}_1\mathsf{X}_1}) \cup \mathcal{R}_2(p_{\mathsf{U}_2\mathsf{X}_2})$ is included in $\bigcup_{p_{\mathsf{U}\mathsf{X}}} \mathcal{R}_2(p_{\mathsf{U}\mathsf{X}})$.

Let $(R_{0,1}, R_{1,1}, R_{e,1}) \in \mathcal{R}_2(p_{\mathsf{U}_1\mathsf{X}_1})$ be a rate triple satisfying the inequalities in (3.45) for some random variables U_1, X_1, Y_1, and Z_1 whose joint distribution satisfies

$$\forall (u, x, y, z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \qquad p_{\mathsf{U}_1\mathsf{X}_1\mathsf{Y}_1\mathsf{Z}_1}(u, x, y, z) = p_{\mathsf{U}_1}(u) p_{\mathsf{X}_1|\mathsf{U}_1}(x|u) p_{\mathsf{Y}\mathsf{Z}|\mathsf{X}}(y, z|x).$$

Let $(R_{0,2}, R_{1,2}, R_{e,2}) \in \mathcal{R}_2(p_{\mathsf{U}_2\mathsf{X}_2})$ be another rate triple satisfying the inequalities in (3.45) for some random variables U_2, X_2, Y_2, and Z_2 whose joint distribution satisfies

$$\forall (u, x, y, z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \qquad p_{\mathsf{U}_2\mathsf{X}_2\mathsf{Y}_2\mathsf{Z}_2}(u, x, y, z) = p_{\mathsf{U}_2}(u) p_{\mathsf{X}_2|\mathsf{U}_2}(x|u) p_{\mathsf{Y}\mathsf{Z}|\mathsf{X}}(y, z|x).$$

Our objective is to show that, for any $\lambda \in [0, 1]$, there exist random variables U_λ and X_λ such that

$$(\lambda R_{0,1} + (1-\lambda) R_{0,2}, \lambda R_{1,1} + (1-\lambda) R_{1,2}, \lambda R_{e,1} + (1-\lambda) R_{e,2}) \in \mathcal{R}_2(p_{\mathsf{U}_\lambda\mathsf{X}_\lambda}).$$

We introduce a random variable $\mathsf{Q} \in \{1, 2\}$ that is independent of all others such that

$$\mathsf{Q} \triangleq \begin{cases} 1 & \text{with probability } \lambda, \\ 2 & \text{with probability } 1-\lambda. \end{cases}$$

By construction, $\mathsf{Q} \to \mathsf{U}_\mathsf{Q} \to \mathsf{X}_\mathsf{Q} \to \mathsf{Y}_\mathsf{Q}\mathsf{Z}_\mathsf{Q}$ forms a Markov chain, and the joint distribution of U_Q, X_Q, Y_Q, and Z_Q satisfies

$$\forall (u, x, y, z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \qquad p_{\mathsf{U}_\mathsf{Q}\mathsf{X}_\mathsf{Q}\mathsf{Y}_\mathsf{Q}\mathsf{Z}_\mathsf{Q}}(u, x, y, z) = p_{\mathsf{U}_\mathsf{Q}}(u) p_{\mathsf{X}_\mathsf{Q}|\mathsf{U}_\mathsf{Q}}(x|u) p_{\mathsf{Y}\mathsf{Z}|\mathsf{X}}(y, z|x).$$

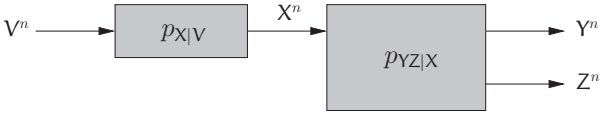

Figure 3.17 Addition of prefix channel to BCC.

Since $Q \to U_Q \to X_Q \to Y_Q Z_Q$ forms a Markov chain, notice that

$$\begin{aligned}
\mathbb{I}(X_Q; Y_Q|U_Q) &= \mathbb{I}(X_Q; Y_Q|U_Q Q),\\
\mathbb{I}(X_Q; Z_Q|U_Q) &= \mathbb{I}(X_Q; Z_Q|U_Q Q),\\
\mathbb{I}(U_Q; Y_Q) &\geqslant \mathbb{I}(U_Q; Y_Q|Q),\\
\mathbb{I}(U_Q; Z_Q) &\geqslant \mathbb{I}(U_Q; Z_Q|Q).
\end{aligned}$$

We set $U_\lambda \triangleq U_Q$, $X_\lambda \triangleq X_Q$, $Y_\lambda \triangleq Y_Q$, and $Z_\lambda \triangleq Z_Q$. Then,

$$\begin{aligned}
&\lambda(R_{0,1} + R_{1,1}) + (1-\lambda)(R_{0,2} + R_{1,2})\\
&\quad\leqslant \lambda(\mathbb{I}(X_1; Y_1|U_1) + \min(\mathbb{I}(U_1; Y_1), \mathbb{I}(U_1; Z_1)))\\
&\qquad + (1-\lambda)(\mathbb{I}(X_2; Y_2|U_2) + \min(\mathbb{I}(U_2; Y_2), \mathbb{I}(U_2; Z_2)))\\
&\quad= \lambda\mathbb{I}(X_1; Y_1|U_1) + (1-\lambda)\mathbb{I}(X_2; Y_2|U_2)\\
&\qquad + \lambda\min(\mathbb{I}(U_1; Y_1), \mathbb{I}(U_1; Z_1)) + (1-\lambda)\min(\mathbb{I}(U_2; Y_2), \mathbb{I}(U_2; Z_2))\\
&\quad\leqslant \mathbb{I}(X_Q; Y_Q|U_Q) + \min\big(\mathbb{I}(U_Q; Y_Q|Q), \mathbb{I}(U_Q; Z_Q|Q)\big)\\
&\quad\leqslant \mathbb{I}(X_Q; Y_Q|U_Q) + \min\big(\mathbb{I}(U_Q; Y_Q), \mathbb{I}(U_Q; Z_Q)\big)\\
&\quad= \mathbb{I}(X_\lambda; Y_\lambda|U_\lambda) + \min(\mathbb{I}(U_\lambda; Y_\lambda), \mathbb{I}(U_\lambda; Z_\lambda)).
\end{aligned}$$

Similarly, we can show,

$$\begin{aligned}
\lambda R_{e,1} + (1-\lambda)R_{e,2} &\leqslant \mathbb{I}(X_Q; Y_Q|U_Q) - \mathbb{I}(X_Q; Z_Q|U_Q)\\
&= \mathbb{I}(X_\lambda; Y_\lambda|U_\lambda) - \mathbb{I}(X_\lambda; Z_\lambda|U_\lambda).
\end{aligned}$$

Hence, for any $\lambda \in [0, 1]$, there exist U_λ and X_λ such that

$$(\lambda R_{0,1} + (1-\lambda)R_{0,2}, \lambda R_{1,1} + (1-\lambda)R_{1,2}, \lambda R_{e,1} + (1-\lambda)R_{e,2}) \in \mathcal{R}_2(p_{U_\lambda X_\lambda}) \subseteq \mathcal{R}'''.$$

Therefore, the convex hull of $\mathcal{R}_2(p_{U_1X_1}) \cup \mathcal{R}_2(p_{U_2X_2})$ is in $\bigcup_{p_{UX}} \mathcal{R}_2(p_{UX})$ and $\bigcup_{p_{UX}} \mathcal{R}_2(p_{UX})$ is convex.

Step 4. Addition of prefix channel

Consider an arbitrary DMC $(\mathcal{V}, p_{X|V}, \mathcal{X})$ that we append before the BCC $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$. This can be done using the source of local randomness. As illustrated in Figure 3.17, the concatenation defines a new BCC $(\mathcal{V}, p_{YZ|V}, \mathcal{Y}, \mathcal{Z})$ such that

$$\forall (v, y, z) \in \mathcal{V} \times \mathcal{Y} \times \mathcal{Z} \quad p_{YZ|V}(y, z|v) = \sum_{x \in \mathcal{X}} p_{YZ|X}(y, z|x) p_{X|V}(x|v).$$

The random-coding argument and outer code construction that led to the characterization of $\mathcal{R}_2(p_{\mathsf{UX}})$ can be reapplied to this new channel. Therefore, we conclude that

$$\bigcup_{p_{\mathsf{UVX}}} \left\{ (R_0, R_1, R_{\mathrm{e}}): \begin{array}{l} 0 \leqslant R_{\mathrm{e}} \leqslant R_1 \\ 0 \leqslant R_{\mathrm{e}} \leqslant \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) \\ 0 \leqslant R_0 + R_1 \leqslant \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \end{array} \right\} \subseteq \mathcal{R}^{\mathrm{BCC}}.$$

By construction, $\mathsf{U} \to \mathsf{V} \to \mathsf{X} \to \mathsf{YZ}$ forms a Markov chain. It remains to prove that we can restrict the cardinality of U and V to $|\mathcal{U}| \leqslant |\mathcal{X}| + 3$ and $|\mathcal{V}| \leqslant |\mathcal{X}|^2 + |\mathcal{X}| + 1$. This follows from Caratheodory's theorem, and we refer interested readers to [18, Appendix] for details.

At this point, the introduction of a prefix channel is somewhat artificial, but the converse part of the proof given in the next section shows that this additional randomization of the stochastic encoder is *required* in order to match the rate–equivocation region obtained in the converse proof. From a practical perspective, the prefix channel suggests that additional randomization may be required in the encoder; this might not be surprising, because the equivocation calculation in Step 2 relies on a *specific* stochastic encoding scheme that might not be optimal. Fortunately, Corollary 3.5 shows that this additional randomization is not necessary if the eavesdropper's channel is less capable than the main channel.

3.5.3 Converse proof for the broadcast channel with confidential messages

Consider an achievable rate triple $(R_0, R_1, R_{\mathrm{e}})$ and let $\epsilon > 0$. For n sufficiently large, there exists a $(2^{nR_0}, 2^{nR_1}, n)$ code $\mathcal{C}_n$ such that

$$\frac{1}{n}\mathbb{H}(\mathsf{M}_0|\mathcal{C}_n) \geqslant R_0, \qquad \frac{1}{n}\mathbb{H}(\mathsf{M}_1|\mathcal{C}_n) \geqslant R_1,$$
$$\frac{1}{n}\mathbf{E}(\mathcal{C}_n) \geqslant R_{\mathrm{e}} - \delta(\epsilon), \qquad \mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) \leqslant \delta(\epsilon).$$

In the remainder of this section, we drop the conditioning on $\mathcal{C}_n$ to simplify the notation. Using Fano's inequality, we obtain

$$\frac{1}{n}\mathbb{H}(\mathsf{M}_0\mathsf{M}_1|\mathsf{Y}^n) \leqslant \delta(\epsilon) \quad \text{and} \quad \frac{1}{n}\mathbb{H}(\mathsf{M}_0|\mathsf{Z}^n) \leqslant \delta(\epsilon).$$

Therefore,

$$\begin{aligned} R_{\mathrm{e}} &\leqslant \frac{1}{n}\mathbf{E}(\mathcal{C}_n) + \delta(\epsilon) \\ &= \frac{1}{n}\mathbb{H}(\mathsf{M}_1|\mathsf{Z}^n) + \delta(\epsilon) \\ &= \frac{1}{n}\mathbb{H}(\mathsf{M}_1|\mathsf{Z}^n\mathsf{M}_0) + \frac{1}{n}\mathbb{I}(\mathsf{M}_0;\mathsf{M}_1|\mathsf{Z}^n) + \delta(\epsilon) \\ &\leqslant \frac{1}{n}\mathbb{H}(\mathsf{M}_1|\mathsf{M}_0) - \frac{1}{n}\mathbb{I}(\mathsf{M}_1;\mathsf{Z}^n|\mathsf{M}_0) + \frac{1}{n}\mathbb{H}(\mathsf{M}_0|\mathsf{Z}^n) + \delta(\epsilon) \end{aligned}$$

$$\begin{aligned}
&\leqslant \frac{1}{n}\mathbb{H}(M_1|M_0) - \frac{1}{n}\mathbb{I}(M_1; Z^n|M_0) + \delta(\epsilon)\\
&= \frac{1}{n}\mathbb{I}(M_1; Y^n|M_0) - \frac{1}{n}\mathbb{I}(M_1; Z^n|M_0) + \frac{1}{n}\mathbb{H}(M_1|Y^n M_0) + \delta(\epsilon)\\
&\leqslant \frac{1}{n}\mathbb{I}(M_1; Y^n|M_0) - \frac{1}{n}\mathbb{I}(M_1; Z^n|M_0) + \delta(\epsilon). \qquad (3.46)
\end{aligned}$$

Single-letterizing (3.46) is more arduous than in Section 3.4.2 because the channel is not degraded. The solution to circumvent this difficulty is a standard technique from multi-user information theory, which consists of symmetrizing the expression by introducing the vectors

$$Y^{i-1} \triangleq (Y_1, \dots, Y_{i-1}) \qquad \text{and} \qquad \tilde{Z}^{i+1} \triangleq (Z_{i+1}, \dots, Z_n) \quad \text{for } i \in [\![1, n]\!],$$

with the convention that $Y^0 = 0$ and $\tilde{Z}^{n+1} = 0$. We introduce Y^{i-1} and $\tilde{Z}^{i+1}$ in $\mathbb{I}(M_1; Y^n|M_0)$ as follows:

$$\begin{aligned}
\mathbb{I}(M_1; Y^n|M_0) &= \sum_{i=1}^{n} \mathbb{I}\big(M_1; Y_i|M_0 Y^{i-1}\big)\\
&= \sum_{i=1}^{n} \Big(\mathbb{I}\big(M_1\tilde{Z}^{i+1}; Y_i|M_0 Y^{i-1}\big) - \mathbb{I}\big(\tilde{Z}^{i+1}; Y_i|M_0 M_1 Y^{i-1}\big)\Big)\\
&= \sum_{i=1}^{n} \Big(\mathbb{I}\big(M_1; Y_i|M_0 Y^{i-1}\tilde{Z}^{i+1}\big) + \mathbb{I}\big(\tilde{Z}^{i+1}; Y_i|M_0 Y^{i-1}\big)\\
&\qquad - \mathbb{I}\big(\tilde{Z}^{i+1}; Y_i|M_0 M_1 Y^{i-1}\big)\Big). \qquad (3.47)
\end{aligned}$$

Similarly, we introduce Y^{i-1} and $\tilde{Z}^{i+1}$ in $\mathbb{I}(M_1; Z^n|M_0)$ as follows:

$$\begin{aligned}
\mathbb{I}(M_1; Z^n|M_0) &= \sum_{i=1}^{n} \Big(\mathbb{I}\big(M_1; Z_i|M_0\tilde{Z}^{i+1}\big)\Big)\\
&= \sum_{i=1}^{n} \Big(\mathbb{I}\big(M_1 Y^{i-1}; Z_i|M_0\tilde{Z}^{i+1}\big) - \mathbb{I}\big(Y^{i-1}; Z_i|M_0 M_1\tilde{Z}^{i+1}\big)\Big)\\
&= \sum_{i=1}^{n} \Big(\mathbb{I}\big(M_1; Z_i|M_0 Y^{i-1}\tilde{Z}^{i+1}\big) + \mathbb{I}\big(Y^{i-1}; Z_i|M_0\tilde{Z}^{i+1}\big)\\
&\qquad - \mathbb{I}\big(Y^{i-1}; Z_i|M_0 M_1\tilde{Z}^{i+1}\big)\Big). \qquad (3.48)
\end{aligned}$$

The key observation to simplify these expressions is the following lemma

Lemma 3.5.

$$\sum_{i=1}^{n} \mathbb{I}\big(\tilde{Z}^{i+1}; Y_i|M_0 Y^{i-1}\big) = \sum_{j=1}^{n} \mathbb{I}\big(Z_j; Y^{j-1}|M_0\tilde{Z}^{j+1}\big),$$

$$\sum_{i=1}^{n} \mathbb{I}\big(\tilde{Z}^{i+1}; Y_i|M_0 M_1 Y^{i-1}\big) = \sum_{j=1}^{n} \mathbb{I}\big(Z_j; Y^{j-1}|M_0 M_1\tilde{Z}^{j+1}\big).$$

Proof. This result follows from the chain rule of mutual information using an appropriate change of indices:

$$\begin{aligned}\sum_{i=1}^{n}\mathbb{I}\big(\tilde{\mathsf{Z}}^{i+1};\mathsf{Y}_i|\mathsf{M}_0\mathsf{Y}^{i-1}\big) &= \sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{I}\big(\mathsf{Z}_j;\mathsf{Y}_i|\mathsf{M}_0\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{j+1}\big)\\ &= \sum_{j=1}^{n}\sum_{i=1}^{j-1}\mathbb{I}\big(\mathsf{Z}_j;\mathsf{Y}_i|\mathsf{M}_0\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{j+1}\big)\\ &= \sum_{j=1}^{n}\mathbb{I}\big(\mathsf{Z}_j;\mathsf{Y}^{j-1}|\mathsf{M}_0\tilde{\mathsf{Z}}^{j+1}\big). \end{aligned} \tag{3.49}$$

Similarly, one can show that

$$\sum_{i=1}^{n}\mathbb{I}\big(\tilde{\mathsf{Z}}^{i+1};\mathsf{Y}_i|\mathsf{M}_0\mathsf{M}_1\mathsf{Y}^{i-1}\big) = \sum_{j=1}^{n}\mathbb{I}\big(\mathsf{Z}_j;\mathsf{Y}^{j-1}|\mathsf{M}_0\mathsf{M}_1\tilde{\mathsf{Z}}^{j+1}\big). \tag{3.50}$$

□

Hence, on substituting (3.47) and (3.48) into (3.46), we obtain, with the help of Lemma 3.5,

$$R_{\mathrm{e}} \leqslant \frac{1}{n}\sum_{i=1}^{n}\Big(\mathbb{I}\big(\mathsf{M}_1;\mathsf{Y}_i|\mathsf{M}_0\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\big) - \mathbb{I}\big(\mathsf{M}_1;\mathsf{Z}_i|\mathsf{M}_0\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\big)\Big) + \delta(\epsilon). \tag{3.51}$$

The common message rate R_0 can be bounded in a similar manner as

$$\begin{aligned} R_0 \leqslant \frac{1}{n}\mathbb{H}(\mathsf{M}_0) &= \frac{1}{n}\mathbb{I}(\mathsf{M}_0;\mathsf{Y}^n) + \frac{1}{n}\mathbb{H}(\mathsf{M}_0|\mathsf{Y}^n)\\ &\leqslant \frac{1}{n}\mathbb{I}(\mathsf{M}_0;\mathsf{Y}^n) + \delta(\epsilon)\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(\mathsf{M}_0;\mathsf{Y}_i|\mathsf{Y}^{i-1}\big) + \delta(\epsilon)\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(\mathsf{M}_0\tilde{\mathsf{Z}}^{i+1};\mathsf{Y}_i|\mathsf{Y}^{i-1}\big) - \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(\tilde{\mathsf{Z}}^{i+1};\mathsf{Y}_i|\mathsf{Y}^{i-1}\mathsf{M}_0\big) + \delta(\epsilon). \end{aligned} \tag{3.52}$$

On substituting the simple bounds

$$\mathbb{I}\big(\tilde{\mathsf{Z}}^{i+1};\mathsf{Y}_i|\mathsf{Y}^{i-1}\mathsf{M}_0\big) \geqslant 0 \quad \text{and} \quad \mathbb{I}\big(\mathsf{M}_0\tilde{\mathsf{Z}}^{i+1};\mathsf{Y}_i|\mathsf{Y}^{i-1}\big) \leqslant \mathbb{I}\big(\mathsf{M}_0\tilde{\mathsf{Z}}^{i+1}\mathsf{Y}^{i-1};\mathsf{Y}_i\big)$$

into (3.52), we obtain

$$R_0 \leqslant \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(\mathsf{M}_0\tilde{\mathsf{Z}}^{i+1}\mathsf{Y}^{i-1};\mathsf{Y}_i\big) + \delta(\epsilon). \tag{3.53}$$

By repeating the steps above with the observation Z^n instead of Y^n we obtain a second bound for R_0:

$$\begin{aligned} R_0 &\leqslant \frac{1}{n}\mathbb{H}(M_0) \\ &\leqslant \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_0\tilde{Z}^{i+1}Y^{i-1};Z_i\big) - \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(Y^{i-1};Z_i|\tilde{Z}^{i+1}M_0\big) + \delta(\epsilon) \qquad (3.54) \\ &\leqslant \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_0\tilde{Z}^{i+1}Y^{i-1};Z_i\big) + \delta(\epsilon). \qquad (3.55) \end{aligned}$$

Finally, we bound the sum-rate $R_0 + R_1$ as follows:

$$\begin{aligned} R_1 + R_0 \leqslant \frac{1}{n}\mathbb{H}(M_0M_1) &= \frac{1}{n}\mathbb{H}(M_1|M_0) + \mathbb{H}(M_0) \\ &= \frac{1}{n}\mathbb{I}(M_1;Y^n|M_0) + \frac{1}{n}\mathbb{H}(M_1|Y^nM_0) + \mathbb{H}(M_0) \\ &\leqslant \frac{1}{n}\mathbb{I}(M_1;Y^n|M_0) + \frac{1}{n}\mathbb{H}(M_0) + \delta(\epsilon). \end{aligned}$$

From (3.52), we know that

$$\frac{1}{n}\mathbb{H}(M_0) \leqslant \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_0\tilde{Z}^{i+1};Y_i|Y^{i-1}\big) - \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(\tilde{Z}^{i+1};Y_i|Y^{i-1}M_0\big) + \delta(\epsilon),$$

and from (3.47) we have

$$\frac{1}{n}\mathbb{I}(M_1;Y^n|M_0) \leqslant \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_1;Y_i|M_0Y^{i-1}\tilde{Z}^{i+1}\big) + \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(\tilde{Z}^{i+1};Y_i|M_0Y^{i-1}\big).$$

On combining the two inequalities, we obtain

$$R_1 + R_0 \leqslant \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_0\tilde{Z}^{i+1};Y_i|Y^{i-1}\big) + \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_1;Y_i|M_0Y^{i-1}\tilde{Z}^{i+1}\big) + \delta(\epsilon). \qquad (3.56)$$

Similarly, using (3.54) in place of (3.52) and using Lemma 3.5, we obtain a second bound

$$R_1 + R_0 \leqslant \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_0Y^{i-1};Z_i|\tilde{Z}^{i+1}\big) + \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big(M_1;Y_i|M_0Y^{i-1}\tilde{Z}^{i+1}\big) + \delta(\epsilon). \qquad (3.57)$$

Let us now introduce the random variables

$$U_i \triangleq Y^{i-1}\tilde{Z}^{i+1}M_0 \quad \text{and} \quad V_i \triangleq U_iM_1.$$

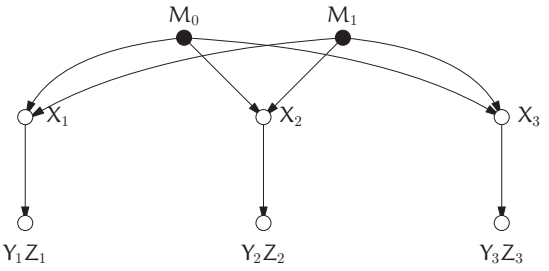

Figure 3.18 Functional dependence graph of random variables involved in the coding scheme for $n = 3$.

Using the functional dependence graph illustrated in Figure 3.18, one can verify that the joint distribution of U_i, V_i, X_i, Y_i, and Z_i is such that

$$\forall (u, v, x, y, z) \in \mathcal{U} \times \mathcal{V} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$$
$$p_{U_i V_i X_i Y_i Z_i}(u, v, x, y, z) = p_{U_i}(u) p_{V_i|U_i}(v|u) p_{X_i|U_i}(x|u) p_{YZ|X}(yz|x),$$

where $p_{YZ|X}$ are the transition probabilities of the BCC $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$. We now introduce a random variable Q that is uniformly distributed over $[\![1, n]\!]$ and independent of $M_0 M_1 X^n Z^n Y^n$, and we define

$$U \triangleq U_Q Q, \quad V \triangleq U M_1, \quad X \triangleq X_Q, \quad Y \triangleq Y_Q, \quad \text{and} \quad Z \triangleq Z_Q. \tag{3.58}$$

Note that $Q \to U \to X \to YZ$ forms a Markov chain. On substituting U, V, X, Y, and Z defined by (3.58) into (3.53) and (3.55), we obtain

$$\begin{aligned} R_0 &\leqslant \min\left(\frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\big(M_0 \tilde{Z}^{i+1} Y^{i-1}; Z_i\big), \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\big(M_0 \tilde{Z}^{i+1} Y^{i-1}; Y_i\big) \right) + \delta(\epsilon) \\ &= \min\left(\frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(U_i; Z_i), \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(U_i; Y_i) \right) + \delta(\epsilon) \\ &= \min(\mathbb{I}(U; Z|Q), \mathbb{I}(U; Y|Q)) + \delta(\epsilon) \\ &\leqslant \min(\mathbb{I}(U; Z), \mathbb{I}(U; Y)) + \delta(\epsilon), \end{aligned} \tag{3.59}$$

where the last inequality follows because $Q \to U \to YZ$ forms a Markov chain. Similarly, on substituting (3.58) into (3.51),

$$\begin{aligned} R_e &\leqslant \frac{1}{n} \sum_{i=1}^{n} \big(\mathbb{I}\big(M_1; Y_i | M_0 Y^{i-1} \tilde{Z}^{i+1}\big) - \mathbb{I}\big(M_1; Z_i | M_0 Y^{i-1} \tilde{Z}^{i+1}\big) \big) + \delta(\epsilon) \\ &= \frac{1}{n} \sum_{i=1}^{n} \left(\mathbb{I}(V_i; Y_i | U_i) - \mathbb{I}(V_i; Z_i | U_i) \right) + \delta(\epsilon) \\ &= \mathbb{I}(V; Y|U) - \mathbb{I}(V; Z|U) + \delta(\epsilon). \end{aligned} \tag{3.60}$$

Finally, on substituting (3.58) into (3.56) and (3.57),

$$R_0 + R_1 \leqslant \mathbb{I}(V; Y|U) + \min(\mathbb{I}(U; Z), \mathbb{I}(U; Y)) + \delta(\epsilon).$$

Since ϵ can be chosen arbitrarily small, we finally obtain the converse result

$$\mathcal{R}^{\text{BCC}} \subseteq \bigcup_{p_{UVX}} \left\{ (R_0, R_1, R_e): \begin{array}{l} 0 \leqslant R_e \leqslant R_1 \\ 0 \leqslant R_e \leqslant R_1 \leqslant \mathbb{I}(V; Y|U) - \mathbb{I}(V; Z|U) \\ 0 \leqslant R_0 + R_1 \leqslant \mathbb{I}(V; Y|U) + \min\left(\mathbb{I}(U; Y), \mathbb{I}(U; Z)\right) \end{array} \right\}.$$

3.6 Multiplexing and feedback

3.6.1 Multiplexing secure and non-secure messages

The expression of the secrecy-capacity region $\mathcal{C}^{\text{BCC}}$ in Corollary 3.2 tells us that, for a fixed distribution p_{UVX} on $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$ that factorizes as $p_U p_{V|U} p_{X|V}$, we can transmit a common message to Eve and Bob at a rate arbitrarily close to $\min(\mathbb{I}(U; Z), \mathbb{I}(U; Y))$ while simultaneously transmitting an individual secret message to Bob at a rate arbitrarily close to $\mathbb{I}(V; Y|U) - \mathbb{I}(V; Z|U)$. However, this result provides only a partial view of what can be transmitted over the channel, because we know from the proof of Theorem 3.3 that Alice can use a dummy message M_d as her source of local randomness; this message is decodable by Bob and is transmitted at a rate arbitrarily close to $\mathbb{I}(V; Z|U)$. Although message M_d is not decodable by Eve, it is not secure either. Hence, we call M_d a *public message* to distinguish it from the common message and the individual secret message. Therefore, for a BCC $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$, there exists a code that conveys three messages reliably:

- a *common message* for Bob and Eve at a rate close to $\min(\mathbb{I}(U; Z), \mathbb{I}(U; Y))$;
- a *secret message* for Bob at a rate close to $\mathbb{I}(V; Y|U) - \mathbb{I}(V; Z|U)$;
- a *public message* for Bob at a rate close to $\mathbb{I}(V; Z|U)$.

The total transmission rate to Bob, R_{tot}, is then arbitrarily close to

$$\min(\mathbb{I}(U; Z), \mathbb{I}(U; Y)) + \mathbb{I}(V; Y|U).$$

In particular, for the specific choice $U = 0$, we obtain a total rate to Bob on the order of $\mathbb{I}(V; Y)$, of which a fraction $\mathbb{I}(V; Y) - \mathbb{I}(V; Z)$ corresponds to a secure message. As illustrated in Figure 3.19, this allows us to interpret a wiretap code for the WTC as a means to create two parallel "pipes," a first pipe transmitting secure messages hidden from the eavesdropper and a second pipe transmitting public messages. From this perspective, note that transmitting secret messages incurs little rate penalty and comes almost "for free." For some specific WTCs, it is possible to show that secrecy comes exactly "for free."

Proposition 3.8. *Consider a WTC* $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$ *in which the eavesdroppers's channel* $\left(\mathcal{X}, p_{Z|X}, \mathcal{Z}\right)$ *with capacity* C_e *is noisier than the main channel* $\left(\mathcal{X}, p_{Y|X}, \mathcal{Y}\right)$ *with capacity* C_m. *Assume that both channels are weakly symmetric. Then, there exists a code that transmits simultaneously*

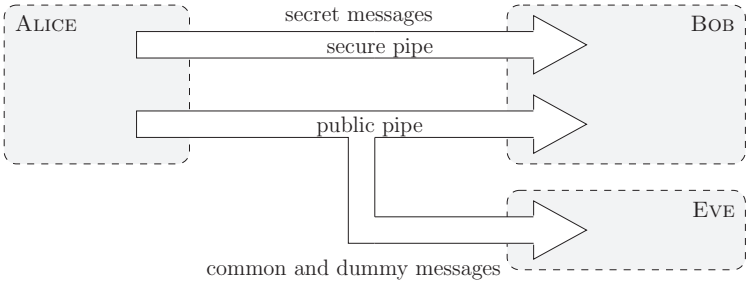

Figure 3.19 Communication over a broadcast channel with confidential messages viewed as parallel bit-pipes.

- *a secret message at a rate* R_s *arbitrarily close to* $C_s^{WT} = C_m - C_e$ *and*
- *a public message at a rate* R_p *arbitrarily close to* C_e.

In other words, it is possible to transmit a secret message at a rate arbitrarily close to the secrecy capacity C_s^{WT} *and still achieve a total reliable transmission rate arbitrarily close to the capacity of the main channel* C_m.

Proof. Let $\epsilon > 0$. For the specific choice $U = 0$ and $V = X$ in Section 3.5.2, we know that there exists a $(2^{nR_s}, n)$ wiretap code $\mathcal{C}_n$ with $R_s = \mathbb{I}(X;Y) - \mathbb{I}(X;Z) - \delta(\epsilon)$ that reliably transmits a dummy message at rate $R_d = \mathbb{I}(X;Z) - \delta(\epsilon)$. The total transmission rate is then

$$R_{tot} = R_s + R_d = \mathbb{I}(X;Y) - \delta(\epsilon).$$

In general, this does not imply that we can exhaust the capacity of the main channel and transmit at the secrecy capacity because the distribution maximizing $\mathbb{I}(X;Y)$ might not maximize $\mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ simultaneously. However, if all channels are weakly symmetric and the eavesdropper's channel is noisier than the main channel, the maximizing distribution is the same, and it is possible to transmit *simultaneously* a secure message at rate

$$\max_{p_X}(\mathbb{I}(X;Y) - \mathbb{I}(X;Z)) - \delta(\epsilon) = C_m - C_e - \delta(\epsilon)$$

and a public message at rate

$$\max_{p_X} \mathbb{I}(X;Z) - \delta(\epsilon) = C_e - \delta(\epsilon),$$

such that the total rate is arbitrarily close to the capacity of the main channel. □

3.6.2 Feedback and secrecy

Although feedback does not increase the channel capacity of a DMC, the situation is quite different for the secrecy capacity. In many situations, it is possible to show that *feedback increases the secrecy capacity*; however, a more precise statement hinges on additional assumptions regarding the nature of the feedback link. The most general approach to analyze feedback would be to consider a two-way communication channel, in which both

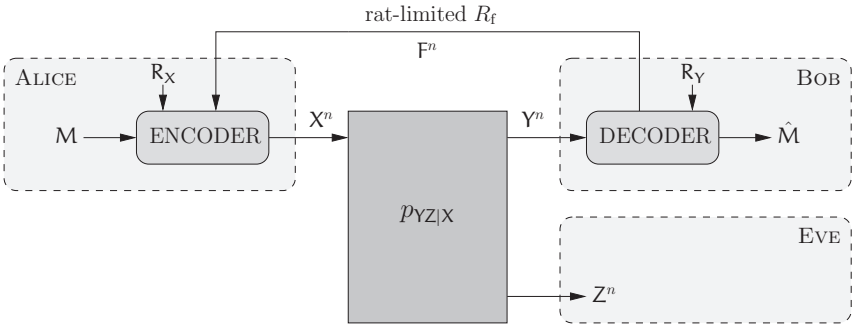

Figure 3.20 WTC with confidential rate-limited feedback.

the forward channel and the reverse channel are broadcast channels overheard by the eavesdropper. A specific instance of a two-way wiretap channel is analyzed in Chapter 8 in the context of multi-user wiretap channels, but a general solution remains elusive. Even if we simplify the model by considering extreme situations in which the feedback link is either *confidential* (unheard by the eavesdropper) or *public* (perfectly heard by the eavesdropper), the fundamental limit is unknown for arbitrary discrete memoryless channels.

In this section, we determine achievable full secrecy rates for a WTC with confidential but rate-limited feedback, which is illustrated in Figure 3.20. This model is a variation of the WTC, in which Bob has the opportunity to transmit confidential messages to Alice at a rate less than R_f. Although the model may seem overly optimistic, it provides valuable insight into how confidential feedback should be used for secrecy. The case of public feedback is studied in Chapter 4 in the context of secret-key agreement.

Definition 3.14. *A* $\left(2^{nR}, n\right)$ *code* $\mathcal{C}_n$ *for a WTC with confidential rate-limited feedback* R_f *consists of*

- *a message set* $\mathcal{M} = [\![1, 2^{nR}]\!]$*;*
- *a source of local randomness* $\left(\mathcal{R}_\mathcal{X}, p_{R_X}\right)$ *at the encoder;*
- *a source of local randomness* $\left(\mathcal{R}_\mathcal{Y}, p_{R_Y}\right)$ *at the receiver;*
- *a feedback alphabet* $\mathcal{F} = [\![1, F]\!]$ *such that* $\log F \leqslant R_f$*;*
- *a sequence of n encoding functions* $e_i : \mathcal{M} \times \mathcal{F}^{i-1} \times \mathcal{R}_\mathcal{X} \to \mathcal{X}$ *for* $i \in [\![1, n]\!]$*, which map a message m, the causally known feedback symbols* f^{i-1}*, and a realization* r_x *of the local randomness to a symbol* $x_i \in \mathcal{X}$*;*
- *a sequence of n feedback functions* $g_i : \mathcal{Y}^{i-1} \times \mathcal{R}_\mathcal{Y} \to \mathcal{F}$ *for* $i \in [\![1, n]\!]$*, which map past observations* y^{i-1} *and a realization* r_y *of the local randomness to a feedback symbol* $f_i \in \mathcal{K}$*;*
- *a decoding function* $g : \mathcal{Y}^n \times \mathcal{R}_\mathcal{Y} \to \mathcal{M} \cup \{\textbf{?}\}$*, which maps each channel observation* y^n *and realization of the local randomness* r_y *to a message* $\hat{m} \in \mathcal{M}$ *or an error message* **?**.

We restrict ourselves to characterization of the secrecy capacity. Note that it is not a priori obvious what the optimal way of exploiting the feedback is. Nevertheless, we can

obtain a lower bound for the secrecy capacity by studying a simple yet effective strategy based on the exchange of a secret key over the confidential feedback channel. In terms of the multiplexing of secure and non-secure messages discussed in Section 3.6.1, the idea is to use a wiretap code to create a secret and a public bit-pipe, and to protect part of the public messages by performing a one-time pad with the secret key obtained over the feedback channel.

Proposition 3.9 (Yamamoto)**.** *The secrecy capacity* $C_{\mathrm{s}}^{\mathsf{FB}}$ *of a WTC* $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$ *with confidential rate-limited feedback* R_{f} *satisfies*

$$C_{\mathrm{s}}^{\mathsf{FB}} \geqslant \max_{p_{\mathsf{V}\mathsf{X}}} \min(\mathbb{I}(\mathsf{V};\mathsf{Y}), \mathbb{I}(\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{V};\mathsf{Z}) + R_{\mathrm{f}}).$$

If the eavesdropper's channel is noisier than the main channel and both channels are weakly symmetric, the bound can be replaced by

$$C_{\mathrm{s}}^{\mathsf{FB}} \geqslant \min(C_{\mathrm{m}}, C_{\mathrm{m}} - C_{\mathrm{e}} + R_{\mathrm{f}}),$$

where C_{m} *is the main channel capacity and* C_{e} *the eavesdropper's channel capacity.*

Proof. Let $\epsilon > 0$, $B \in \mathbb{N}^*$, and $n \in \mathbb{N}^*$. We consider a block-Markov coding scheme over B blocks of length n, such that the selection of the transmitted codeword in each block $b \in [\![2, B]\!]$ is a function of the current message m_b and of an independent secret key k_{b-1} exchanged during the previous block over the secure feedback channel. The secret key allows us to encrypt otherwise insecure parts of the codewords with a one-time pad, and thus leads to higher secure communication rates.

Formally, the scheme operates as follows. We assume there exists a distribution p_{X} over $\mathcal{X}$ such that $\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) > 0$. We then consider a $(2^{nR}, 2^{nR_{\mathrm{d}}}, n)$ code $\mathcal{C}_n$ identified by the random coding argument in Section 3.4.1 with

$$R = \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) - \delta(\epsilon), \qquad R_{\mathrm{d}} = \mathbb{I}(\mathsf{X};\mathsf{Z}) - \delta(\epsilon)$$

and such that

$$\mathbf{P}_{\mathrm{e}}(\mathcal{C}_n) \leqslant \delta_\epsilon(n) \quad \text{and} \quad \frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon) + \delta_\epsilon(n).$$

We set $R_{\mathrm{o}} = \min(R_{\mathrm{f}}, R_{\mathrm{d}})$ and, for each $m \in [\![1, 2^{nR}]\!]$, we distribute the codewords $x^n(m, m_{\mathrm{d}})$ with $m_{\mathrm{d}} \in [\![1, 2^{nR_{\mathrm{d}}}]\!]$ in $\lceil 2^{nR_{\mathrm{o}}} \rceil$ bins $\mathcal{B}_m(i)$ with $i \in [\![1, 2^{nR_{\mathrm{o}}}]\!]$ and we relabel the codewords

$$x^n(i, j, k) \quad \text{with} \quad (i, j, k) \in [\![1, 2^{nR}]\!] \times [\![1, 2^{nR_{\mathrm{o}}}]\!] \times [\![1, 2^{nR_{\mathrm{d}} - nR_{\mathrm{o}}}]\!].$$

This binning procedure is the same as the one illustrated in Figure 3.11. The sub-binning is revealed to all parties and we consider the following encoding/decoding procedure over B blocks of length n.

- *Encoder for block 1.* Given m_1 and m'_1, transmit $x^n(m_1, m'_1, k) \in \mathcal{C}_n$, where k is chosen uniformly at random in $\mathcal{B}_{m_1}(m'_1)$. During the transmission, receive secret key k_1 uniformly distributed in $[\![1, 2^{nR_{\mathrm{o}}}]\!]$ over the feedback channel.
- *Encoder for block* $b \in [\![2, B]\!]$. Given m_b and m'_b, transmit $x^n(m_b, m'_b \oplus k_{b-1}, k) \in \mathcal{C}_n$, where $\oplus$ denotes modulo-$\lceil 2^{nR_{\mathrm{o}}} \rceil$ addition and k is chosen uniformly at random in

$\mathcal{B}_{m_b}(m'_b)$. During the transmission, receive secret key k_b uniformly distributed in $[\![1, 2^{nR_o}]\!]$ over the feedback channel.

- *Decoder for block* $b = 1$. Given y^n, use the decoding procedure for $\mathcal{C}_n$ described in Section 3.4.1.
- *Decoder for block* $b \in [\![2, B]\!]$. Given y^n, use the decoding procedure for $\mathcal{C}_n$ to retrieve m_b and $m'_b \oplus k_{b-1}$. Use k_{b-1} to retrieve m'_b.

The sub-binning together with the encoding and decoding procedures above defines a $(2^{nBR'}, nB)$ code $\tilde{\mathcal{C}}_{nB}$ of length nB for the WTC with secure feedback. We assume that messages M_b and M'_b for $b \in [\![1, B]\!]$ are uniformly distributed so that the rate of $\tilde{\mathcal{C}}_{nB}$ is

$$R' = R + R_o - \delta(n) = \min(\mathbb{I}(\mathsf{X};\mathsf{Y}), \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) + R_f) - \delta(n).$$

We ignore the fact that the distribution of the codewords of $\mathcal{C}_n$ may be slightly non-uniform because of the binning and we refer the reader to Section 3.4.1 for details on how to deal with this subtlety.

Because the secret keys k_b are perfectly known both to the encoder and to the decoder, the probability of error for $\tilde{\mathcal{C}}_{nB}$ is at most B times that of $\mathcal{C}_n$ and

$$\mathbf{P}_e(\tilde{\mathcal{C}}_{nB}) \leqslant B\mathbf{P}_e(\mathcal{C}_n) \leqslant B\delta_\epsilon(n).$$

For $b \in [\![1, b]\!]$, we let Z_b^n denote the eavesdropper's observation in block b. The information leaked $\mathbf{L}(\tilde{\mathcal{C}}_{nB})$ is then

$$\begin{aligned}\frac{1}{nB}\mathbf{L}(\tilde{\mathcal{C}}_{nB}) &= \frac{1}{nB}\mathbb{I}\big(\mathsf{M}_1 \dots \mathsf{M}_B \mathsf{M}'_1 \dots \mathsf{M}'_B; \mathsf{Z}_1^n \dots \mathsf{Z}_B^n | \tilde{\mathcal{C}}_{nB}\big) \\ &= \frac{1}{nB}\sum_{b=1}^{B}\mathbb{I}\big(\mathsf{M}_b\mathsf{M}'_b; \mathsf{Z}_b^n | \tilde{\mathcal{C}}_{nB}\big) \\ &= \frac{1}{B}\sum_{b=1}^{B}\left(\frac{1}{n}\mathbb{I}\big(\mathsf{M}_b; \mathsf{Z}_b^n | \tilde{\mathcal{C}}_{nB}\big) + \frac{1}{n}\mathbb{I}\big(\mathsf{M}'_b; \mathsf{Z}_b^n | \mathsf{M}_b\tilde{\mathcal{C}}_{nB}\big)\right),\end{aligned}$$

where the first equality follows because M_b and M'_b depend only on Z_b^n. For $b \geqslant 2$, message M'_b is protected with a one-time pad and the crypto lemma guarantees that $(1/n)\mathbb{I}\big(\mathsf{M}'_b; \mathsf{Z}_b^n | \mathsf{M}_b\tilde{\mathcal{C}}_{nB}\big) = 0$. For $b = 1$, we can use the upper bound

$$\frac{1}{n}\mathbb{I}\big(\mathsf{M}'_1\mathsf{Z}_1^n | \mathsf{M}_1\tilde{\mathcal{C}}_{nB}\big) \leqslant R_o + \delta(n).$$

In addition, the construction of $\tilde{\mathcal{C}}_{nB}$ guarantees that

$$\frac{1}{n}\mathbb{I}\big(\mathsf{M}_b; \mathsf{Z}_b^n | \tilde{\mathcal{C}}_{nB}\big) = \frac{1}{n}\mathbf{L}(\mathcal{C}_n) \leqslant \delta(\epsilon) + \delta_\epsilon(n) \quad \text{for } b \in [\![1, B]\!].$$

Therefore,

$$\begin{aligned}\frac{1}{nB}\mathbf{L}(\tilde{\mathcal{C}}_{nB}) &\leqslant \frac{1}{B}\left(R_o + \delta(n) + \sum_{b=2}^{B}(\delta_\epsilon(n) + \delta(\epsilon))\right) \\ &= \delta(\epsilon) + \delta(B) + \delta_\epsilon(n).\end{aligned}$$

Hence, the rate–equivocation pair $(R', R' - \delta(\epsilon) - \delta(B))$ is achievable. Since ϵ can be chosen arbitrarily small and B can be chosen arbitrarily large, we conclude that

$$C_{\mathrm{s}}^{\mathrm{FB}} \geqslant \max_{p_{\mathsf{X}}} \min(\mathbb{I}(\mathsf{X};\mathsf{Y}), \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) + R_{\mathrm{f}}).$$

One can also check that $\min(\mathbb{I}(\mathsf{X};\mathsf{Y}), R_{\mathrm{f}})$ is an achievable rate if $\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}) = 0$ for all X. As in Section 3.5.2, we can introduce a prefix channel $\left(\mathcal{V}, p_{\mathsf{X}|\mathsf{V}}, \mathcal{X}\right)$ before the WTC $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$ to obtain the desired lower bound.

If all channels are weakly symmetric and the eavesdropper's channel is noisier than the main channel, we can remove the prefix channel as in the proof of Corollary 3.5, and the symmetry guarantees that both $\mathbb{I}(\mathsf{X};\mathsf{Y})$ and $\mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z})$ are maximized with X uniformly distributed. □

If the eavesdropper's channel is physically degraded with respect to the main channel, the pragmatic feedback strategy used in Proposition 3.9 turns out to be optimal. The proof of optimality is established in Section 4.4 on the basis of results about the secret-key capacity.

Theorem 3.4 (Ardestanizadeh *et al.*)**.** *The secrecy capacity of a DWTC* $(\mathcal{X}, p_{\mathsf{YZ}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z})$ *with confidential rate-limited feedback* R_{f} *is*

$$C_{\mathrm{s}}^{\mathrm{DWTC}} = \max_{p_{\mathsf{X}}} \min(\mathbb{I}(\mathsf{X};\mathsf{Y}), \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}) + R_{\mathrm{f}}).$$

Note that, even if $\mathsf{Z} = \mathsf{Y}$ and the secrecy capacity without feedback is zero, Proposition 3.9 guarantees a non-zero secure communication rate. This is not surprising because the feedback link is secure and can always be used to transmit a secure key and perform a one-time pad encryption; however, we will see in Chapter 4 and Chapter 8 that even public feedback enables more secure communications.

3.7 Conclusions and lessons learned

We conclude this chapter by summarizing the lessons learned from the analysis of secure communication over noisy DMCs. Most importantly, we proved the existence and identified the structure of codes that guarantee reliability and security simultaneously over wiretap channels. The specificity of these codes is that they map a given message to a *bin* of codewords, of which one is selected randomly by the encoder. Intuitively, the role of the randomness in the encoder is to "confuse" the eavesdropper by compensating for the information leaked about the codeword during transmission.

Unfortunately, as exemplified by Corollary 3.5, secure communications at a non-zero rate seem to be possible only when the legitimate receiver has a "physical advantage" over the eavesdropper, which we formalized with the notion of "noisier" channels. Although there are practical applications for which this condition is met, in particular if the Gaussian wiretap channel studied in Chapter 5 is a realistic model (near-field communications, RFID transmissions; see Chapter 7), it is fair to acknowledge that requiring an explicit advantage over the eavesdropper is a weakness in the model.

From the results of this chapter alone, one would even be inclined to conclude that the information-theoretic approach considered in this chapter does not improve on standard cryptography, and might merely be an alternative solution for which different trade-offs are made. Cryptography imposes restrictions on the computational power of the eavesdropper to relax assumptions on the communication channel (it considers the worst situation with a noiseless channel to the eavesdropper), whereas information-theoretic security seems to require an advantage at the physical layer in order to avoid restrictive assumptions on the abilities of the eavesdropper. Nevertheless, the somewhat unsatisfactory results obtained thus far hinge on the simplicity of the communication schemes, which do not account for powerful codes based on interactive communications. We will see in subsequent chapters that information-theoretically secure communications are sometimes possible without an explicit advantage at the physical layer.

We close this section with a review of the explicit and implicit assumptions used in the previous sections. Understanding these assumptions and their relevance is crucial to assess the potential of physical-layer security from a cryptographic perspective.

- **Knowledge of channel statistics.** In general, the lower bound of the eavesdropper's equivocation guaranteed by Theorem 3.2 is valid only if the code is tailored to the channel, which requires knowledge of the channel statistics *both* for the main channel and for the eavesdropper's channel. The assumption that the main channel is perfectly known is usually reasonable, since Alice and Bob can always cooperate to characterize their channel; however, the assumption that the eavesdropper's channel statistics are also known is more questionable. Nevertheless, as discussed in Section 3.5.1, this stringent requirement can be somewhat alleviated for a class of stochastically degraded channels.
- **Authentication.** The wiretap channel model assumes implicitly that the main channel is authenticated and, therefore, wiretap codes do not provide any protection against man-in-the-middle attacks; however, this assumption is not too restrictive, since authentication mechanisms can be implemented in the upper layers of the protocol stack. We shall see in Chapter 7 that it is possible to ensure unconditionally secure authentication with a negligible cost in terms of the secure communication rate, provided that a short secret key is available initially.
- **Passive attacker.** The scope of the results developed in this chapter is strictly restricted to eavesdropping strategies. Additional techniques are required for situations in which the adversary tampers with the channels, for instance by jamming the channel.
- **Perfect random-number generation.** The proofs developed in this chapter rely on the availability of a perfect random-number generator at the transmitter. The reader can check that the equivocation *decreases* if the entropy of the random-number generator is not maximum.
- **Weak secrecy.** As discussed in Section 3.3, weak secrecy is likely not an appropriate cryptographic metric. Therefore, the relevance of the results derived in this chapter will be fully justified in Chapter 4 when we show that the results do not change if the weak secrecy criterion is replaced by the strong secrecy criterion.

3.8 Bibliographical notes

The one-time pad was developed by Vernam during World War I for the Signal Corps of the US Army [19]. Although no rigorous analysis of the cipher was performed, the need for a random key as long as the message to be encrypted had already been identified. Shannon's cipher system formalized the problem of secure communications over noiseless channels in an information-theoretic framework [1]. Shannon's work includes an analysis of the one-time pad, but the crypto lemma in its most general form is due to Forney [20]. The degraded wiretap channel model and the notion of secrecy capacity were first introduced by Wyner in his seminal paper [21]. The simple expression of the secrecy capacity for symmetric channels was established by Leung-Yan-Cheong shortly after [22], and a simplified proof of Wyner's result was proposed by Massey [23]. The extension of Wyner's results to broadcast channels with confidential messages is due to Csiszár and Körner [18]. The proofs presented in this chapter are based on typical-set decoding in the spirit of [21], and can be easily combined with standard random-coding techniques for multi-user channels [3, 6]; however, this approach yields only weak secrecy results, and additional steps are required in order to strengthen the secrecy criterion. These steps are based on results from secret-key agreement that are presented in Section 4.5. Alternative proofs based on more powerful mathematical tools, such as graph-coloring techniques [24] or information-spectrum methods [25, 26], can be used to derive the secrecy capacity with strong secrecy directly.

Although stochastic encoders and codes with a binning structure are required for secure communications over memoryless channels, this need not be the case for other models. For instance, Dunn, Bloch, and Laneman investigated the possibility of secure communications over parallel timing channels [27] and showed that deterministic codes can achieve non-zero secure rates.

The various notions of partial ordering of channels were introduced by Körner and Marton [28], and the characterization of the relation "being noisier" in terms of the concavity of $\mathbb{I}(X;Y) - \mathbb{I}(X;Z)$ was obtained by Van Dijk [29]. The examples illustrating the notions of "noisier" and "less capable" channels given in Section 3.5.1 are due to Nair [30]. The combination of wiretap coding and one-time pad was proposed by Yamamoto [31], and the wiretap channel with a secure feedback link was investigated by Leung-Yan-Cheong [32], Ahlswede and Cai [33], and Ardestanizadeh, Franceschetti, Javidi, and Kim [34].

All of the results presented in this chapter assume a uniformly distributed source of messages; nevertheless, Theorem 3.2 and Theorem 3.3 can be generalized to account for arbitrarily distributed sources [18, 21], in which case a separation result holds: no rate is lost by first compressing the source and then transmitting it over a broadcast channel with confidential messages.

Although we mentioned that equivocation and probability of error are fundamentally different quantities, the relation between them was investigated more precisely by Feder and Merhav [35]. In particular, they provide lower bounds for the conditional entropy as a function of the probability of block error, which confirm that evaluating security

on the basis of equivocation is a stronger requirement than simply requiring a decoding error for the eavesdropper.

An important issue that we have not addressed in this chapter is the numerical computation of the secrecy capacity. The presence of auxiliary random variables in Corollary 3.4 makes this computation difficult, and no efficient generic algorithm is known. Nevertheless, headway has been made in certain cases. For cases in which the eavesdropper's channel is noisier than the main channel, Yasui, Suko, and Matsushima proposed a Blahut–Arimoto-like algorithm to compute the secrecy capacity [36], which was also shown to be useful when the eavesdropper's channel is less capable than the main channel by Gowtham and Thangaraj [37].